

CARACTERIZACIÓN, CLASIFICACIÓN Y FILOGENIA DE ADENILIL Y GUANILIL CICLASAS DE CIANOBACTERIAS

Jesús A. G. Ochoa de Alda
Ana del Pico
Andrés Pedraza
Universidad SEK
Jean Houmard
Ecole Normale Supérieure
jochoa@sekmail.com

Resumen

*Las adenilil (AC) y guanilil ciclasas (GC) sintetizan los nucleótidos cíclicos cAMP y cGMP, respectivamente. Por ello juegan un papel fundamental en la transmisión de señales ambientales tanto en las células procariotas como en las eucariotas. A diferencia de la mayor parte de bacterias, algunas cianobacterias contienen tanto cGMP como cAMP y sus concentraciones intracelulares fluctúan en función de la disponibilidad de oxígeno, nutrientes, y las condiciones lumínicas. La disponibilidad de genomas de cianobacterias completamente secuenciados revela la existencia de un número creciente de ACs y GCs de clase III en este grupo de procariotas fotosintéticos. En este trabajo, describimos la totalidad de ciclasas de nucleótidos purínicos encontradas actualmente en los genomas de cianobacterias. Para ello analizamos la secuencia primaria de las ACs y GCs mediante alineamiento múltiple, estudiamos su arquitectura molecular y empleamos métodos filogenéticos para inferir su origen. Como regla general, observamos que las ACs y GCs son más abundantes: i) en cianobacterias de agua dulce que en cianobacterias marinas, ii) en cianobacterias filamentosas que en unicelulares, y iii) en cianobacterias diazotrofas que en las no fijadoras de nitrógeno. Esto sugiere una estrecha relación entre la ecofisiología de estas bacterias y el sistema de transmisión de señales ambientales dependiente de los nucleótidos cíclicos que emplean. La mayor diversidad de ACs y GCs se observa en la cianobacteria filamentosa marina fijadora de nitrógeno *Trichodesmium erythraeum* que contiene 11 posibles ACs y 2 posibles GCs. A partir del análisis bioinformática sugerimos una nomenclatura sistemática para las ciclasas de nucleótidos purínicos de cianobacterias.*

Esta clasificación requiere una caracterización bioquímica posterior de los principales representantes de ACs y GCs (en particular los más divergentes) con el fin de confirmar la función predicha en este trabajo. El análisis filogenético indica que el acervo de ACs y GCs en las cianobacterias es el resultado de duplicaciones y deleciones genéticas, de fusiones y fisiones de genes y de mutaciones puntuales en el centro activo que han convertido algunas ACs en GCs. Como consecuencia, es frecuente observar diferentes isoformas de ACs así como el centro catalítico de estos enzimas formando parte del extremo carboxilo terminal en arquitecturas moleculares muy diferentes, entre las que se observan polipéptidos multienzimáticos. En conjunto, los resultados presentados sugieren que las cianobacterias son el organismo modelo más simple en el que se puede estudiar la complejidad del sistema que constituyen las rutas de señalización dependientes de nucleótidos cíclicos en los seres vivos.

Palabras clave: Genómica; Cianobacterias; Adenilil ciclasas; Gualilil ciclasas.

Summary

*Adenylyl (ACs) and guanylyl cyclases (GCs) synthesize the cyclic nucleotides cAMP and cGMP, playing a key role in signal transduction both in prokaryotes and eukaryotes. Some cyanobacteria produce cAMP and cGMP and their concentrations fluctuate depending on nutrients and oxygen availability, as well as on light conditions. Advances in cyanobacterial genome-sequencing projects reveal the existence of an increasing number of putative ACs and GCs which all belong to the so-called universal Class III of AC/GCs. We describe the complete sets of these purine nucleotide cyclases found in the available genomes, and analyze them by primary sequence alignment, molecular architecture and phylogenetic methods. As a rule, ACs and GCs are more abundant i) in fresh water than in marine cyanobacteria, ii) in filamentous than in unicellular cyanobacteria and iii) in diazotrophic than in non-diazotrophic ones. The filamentous diazotrophic marine cyanobacterium *Trichodesmium erythraeum* IMS101 contains 11 putative ACs and 2 putative GCs. This suggests a relationship between the ecophysiology of this bacteria and the cyclic nucleotide transduction pathway that they contain. Our analysis allows us to derive a systematic nomenclature for cyanobacterial purine nucleotide cyclases that will require a further precise biochemical characterization, in particular for the most divergent sequences, to confirm their putative functions. In cyanobacteria, the repertoire of ACs and GCs is likely the result of deletions, gene duplication, protein fusion and fission events, as well as mutation in the residues that determine the specificity for ATP and GTP. As a consequence, AC isoforms are frequently present in cyanobacteria as well as the presence of ACs and GCs domains at the C-terminal of multienzyme polypeptides. Altogether, the results suggest that cyanobacteria could be the simplest model organisms to study the system complexity of cyclic nucleotide signaling pathways.*

Key Words: Genomic; Cyanobacteria, Adenylyl cyclase, Guanylyl cyclase.

* * * * *

Introducción^{*/**/**}

Una de las principales características de los seres vivos es su capacidad de adaptación a las variables condiciones ambientales de la biosfera (Koshland, 2002). En el nivel molecular, esta adaptación requiere de un intrincado sistema de señalización que integra los diferentes estímulos ambientales y permite una respuesta apropiada a las necesidades de cada momento. Los sistemas de señalización contienen frecuentemente segundos mensajeros, moléculas que median entre el estímulo ambiental (i.e. luz, temperatura, nutrientes) y la adaptación.

Los sistemas de señalización dependientes de los segundos mensajeros AMP cíclico (adenosina-3', 5'-monofosfato) y GMP cíclico (guanosina-3',5'-monofosfato), cAMP y cGMP respectivamente, se encuentran en los tres dominios en los que se agrupan los seres vivos y sus funciones abarcan desde el control de la patogénesis, el metabolismo, la movilidad y la reproducción de microorganismos (Baker and Kelly, 2004) al control de la polinización de plantas (Moutinho *et al.*, 2001) y la visión y transmisión nerviosa en animales superiores (Essayan, 2001; Sunahara and Taussig, 2002), entre otras.

Los niveles intracelulares de cAMP y cGMP están controlados por el balance que se establece entre su síntesis, su degradación y su excreción. El cAMP es sintetizado por la enzima adenilil ciclasa, AC, (EC 4.6.1.1) y el cGMP

* Abreviaturas. AC, adenilil ciclasa; Ala: Alanina; Asp: Aspártico; Asn: Asparragina; Arg: Arginina; ATP, adenosina-5'-trifosfato; CYC: Dominio catalítico de las adenilil y guanilil ciclasas de clase III; Cys: Cisteína; cAMP, adenosina-3', 5'-monofosfato; cGMP, guanosina-3',5'-monofosfato; GC, guanilil ciclasas; GI: Identificador del registro en el banco de genes; Gly: Glicina; Glu: Glutámico; Gln: Glutamina; GTP, guanosina-5'-trifosfato; IUPAC-IUB, International Union of Pure and Applied Chemistry- International Union of Biochemistry; Leu: Leucina; NCBI, National Center for Biotechnology Information; Phe: Fenilalanina; Ser: Serina; Thr: Treonina; Val: Valina

** Contribuciones de los autores. Jean Houmard y Jesús Ochoa de Alda han desarrollado conceptualmente el trabajo. Jesús Ochoa de Alda ha dirigido, supervisado y redactado el trabajo. Ana de Pico y Jesús AG Ochoa de Alda han analizado las bases de datos, realizado los alineamientos, estudiado las arquitecturas y la filogenia. Andrés Pedraza ha estudiado los dominios sensores, la filogenia de la arquitectura A. Ana del Pico y Andrés Pedraza han preparado las figuras.

*** Agradecimientos. El trabajo presentado ha sido financiado por la Consejería de Educación y Cultura de la Junta de Castilla y León, referencia SEK7/04. Este estudio ha sido posible gracias al trabajo de todos aquellos investigadores que han depositado en el banco de genes, antes de su publicación, las secuencias de los genomas de cianobacterias y a aquellas instituciones que, a través de sus servidores permiten la utilización de numerosos programas bioinformáticos como en *Institut Pasteur* y el *European Bioinformatics Institut*. El estudio presentado es parte del Trabajo Fin de Carrera defendido por Ana María del Pico Ausín en enero de 2003.

por la enzima guanilil ciclasa, GC, (EC 4.6.1.2). Ambos nucleótidos son degradados por fosfodiesterasas (EC 3.1.4.17 y EC 3.1.4.35) con diferente grado de especificidad por los nucleótidos cíclicos (Bosgraaf *et al.*, 2002; Essayan, 2001). El efecto de la variación intracelular de nucleótidos cíclicos se ejerce sobre receptores que, dependiendo de los organismos, pueden controlar la actividad de enzimas y/o la transcripción de genes (Barnstable *et al.*, 2004; Ochoa de Alda and Houmard, 2000).

Las ACs y GCs descritas hasta el momento se agruparon inicialmente en 3 clases (Danchin, 1993) que han aumentado progresivamente hasta 7 (Cotta *et al.*, 1998; Danchin, 1993; Ludidi and Gehring, 2003; Moutinho *et al.*, 2001; Sismeiro *et al.*, 1998; Tellez-Sosa *et al.*, 2002). Las diferencias en la secuencia primaria de cada una de las clases sugieren un origen evolutivo independiente para cada una de las clases.

La clase I de ACs, también llamada de enterobacterias, está representada por la de *Escherichia coli*. En esta AC (P00936, GI:32172418), el centro catalítico se sitúa entre los residuos 82 y 340 (Holland *et al.*, 1988).

La clase II de ACs, también llamada tóxica, está representada por la de *Bordetella pertussis* (Y00545.1, GI:396665). En esta AC el centro catalítico se sitúa en los 400 residuos amino terminales (Yahr *et al.*, 1998). Este dominio se encuentra formando parte de diferentes arquitecturas moleculares en ACs del genero *Bacillus* del *Phylum Firmicutes*, ii) del genero *Bordetella* de la clase B-Proteobacteria, iii) de los generos *Yersinia* y *Pseudomonas* de la clase Proteobacteria, perteneciendo estas últimas clases al *Phylum Proteobacteria*. La AC de clase II mejor caracterizada es la de *B. anthracis* que ha sido recientemente cristalizada (Drum *et al.*, 2002).

La clase III, también llamada universal, comprende la mayor parte de las ACs y GCs descritas hasta el momento (Baker and Kelly, 2004; Danchin, 1993; Shenroy and Visweswariah, 2004). Se encuentran homólogos de estas ciclasas en los 3 dominios de los seres vivos, principalmente en eucariotas y bacterias. La arquitectura molecular de estas ciclasas es muy variada. Los representante mejor estudiado son la AC de membrana de mamíferos, el centro catalítico de una de cuyas 11 isoformas ya ha sido cristalizado (Zhang *et al.*, 1997), la AC CyaC de la cianobacteria *Spirulina platensis* (Steebhorn *et al.*, 2004) y la AC Rv1900c de *Mycobacterium tuberculosis* (Sinha *et al.*, 2005). Esto ha permitido identificar los aminoácidos implicados en la actividad catalítica de ciclación del nucleótido purínico trifosfato y en la especificidad bien por el ATP (en

las ACs de clase III) o por el GTP (en las GCs de esta clase) (Liu *et al.*, 1997; Sinha *et al.*, 2005; Steegborn *et al.*, 2004). La especificidad por el sustrato está determinada por 2-3 residuos que, cuando son adecuadamente mutados, convierten una GC en una AC (Kasahara *et al.*, 2001; Linder *et al.*, 2000; Tucker *et al.*, 1998) y una AC en una GC (Kasahara *et al.*, 2001). No obstante se ha propuesto que otros residuos también podrían estar implicados en la catálisis y en la selectividad por el sustrato (Hannenhalli and Russell, 2000; Linder and Schultz, 2003). La arquitectura molecular de las ciclasas de clase III es muy variable. En general, el dominio ciclasa se encuentra en el extremo amino terminal de la proteína y está unido a dominios sensores que controlan la actividad enzimática. La actividad requiere la formación de homodímeros, heterodímeros o de homodominios (Baker and Kelly, 2004; Cann, 2004; Shenroy and Visweswariah, 2004). El miembro más divergente de esta clase lo constituye una GC recientemente identificada en plantas superiores (Ludidi and Gehring, 2003). El extremo N-terminal de esta proteína muestra similitudes con algunos residuos implicados en la catálisis y la especificidad por el GTP de algunas GCs de clase III.

La clase IV se identificó inicialmente en arqueobacterias hipertermófilas (Sismeiro *et al.*, 1998) y está representada por la AC CyaB de *Aeromonas hydrophyla*. Homólogos de esta proteína se encuentran en los 3 dominios de los seres vivos y se agrupan en el dominio estructural denominado CYTH/CyaB cuyos miembros mejor caracterizados son la AC de *A. hydrophyla* y la enzima timina trifosfatasa (Iyer and Aravind, 2002). La falta de estudios funcionales y estructurales del dominio CYTH/CyaB dificulta la asignación por homología de una de estas dos funciones enzimáticas a una secuencia primaria de proteína.

La clase V de ACs se ha identificado únicamente en la bacteria anaerobia del rumen *Prevotella ruminicola* (Cotta *et al.*, 1998).

La clase VI de ACs se identificó inicialmente en *Rhizobium etli* (Tellez-Sosa *et al.*, 2002) y se encuentra también en otros miembros de la clase α -Proteobacteria y en las especies *Leptospira interrogans* L1-130 y en *Chloroflexus aurantiacus* de los Phyla *Spirochaetes* y *Chloroflexi*, respectivamente

La clase VII corresponde a una AC de plantas superiores (Moutinho *et al.*, 2001). Contiene un dominio N terminal de unión a ATP que podría estar implicado en la catálisis.

Aunque en general los seres vivos contienen una única clase de ciclasas, algunos organismos como *Rhizobium* y *Pseudomonas* disponen de varias cla-

ses de ACs en su genoma (Tellez-Sosa et al., 2002) y (Yahr et al., 1998).

Uno de los organismos en los que se ha encontrado mayor diversidad de ACs y GCs son las cianobacterias (Baker and Kelly, 2004; Cann, 2004; Shenroy and Visweswariah, 2004), un anciano y diverso grupo de bacterias que aparecieron en la tierra hace más de 3500 millones de años. La característica común de este grupo de procariotas es que llevan a cabo una fotosíntesis de tipo oxigénico, con un indudable origen monofilético, que es análoga a la de los cloroplastos de las plantas superiores (Tandeau de Marsac and Houmard, 1993). Todas las ACs y GCs de cianobacterias descritas hasta el momento son de la clase III y muestran una arquitectura molecular compleja en la que el dominio catalítico está fusionado a diversos dominios reguladores (Baker and Kelly, 2004; Cann, 2004; Ochoa de Alda et al., 2000; Shenroy and Visweswariah, 2004). Dentro de las ciclasas bacterianas de clases III, las ACs de cianobacterias son las mejor caracterizadas desde el punto de vista bioquímico, genético y fisiológico (Kanacher et al., 2002; Kasahara and Ohmori, 1999; Kasahara et al., 2001; Masuda and Ono, 2004; Steegborn et al., 2004; Tada et al., 2001). La única GC procariota descrita hasta el momento pertenece al *Phylum Cyanobacteria* (Ochoa de Alda et al., 2000).

El aumento continuo de tipos de ACs y GCs dentro del *Phylum Cyanobacteria* exige una clasificación sistemática que mejore su comprensión. En *Nostoc* PCC 7120 se han descrito 6 ACs que han sido denominadas CyaA, CyaB1, CyaB2, CyaC, CyaD, CyaE (Kasahara et al., 2001). En *Synechocystis* PCC 6083 se han descrito 3 ACs y GCs que han sido denominadas Cya1, Cya2, Cya3 (Ochoa de Alda et al., 2000; Terauchi and Ohmori, 1999). En *Spirulina platensis* se han clonado 3 ACs Cya A (Yashiro et al., 1996), CyaC (Kasahara et al., 1997), CyaG (Kasahara et al., 2001). La nomenclatura de las ACs y GCs no siempre está relacionada con la actividad enzimática que catalizan ni con su arquitectura molecular. De hecho, CyaA de *Spirulina* y *Nostoc* solo son similares en el centro catalítico y no en su arquitectura molecular. Por otra parte, si se sigue la nomenclatura recomendada por la IUPAC para multienzimas (IUPAC-IUB, 1977, 1989), la denominación de Cya1, Cya2, Cya3 debería emplearse para isoformas y no para enzimas con arquitectura molecular tan diferente.

El estudio de los seres vivos como sistema (entendiendo por sistema un conjunto de partes o variables conectadas unas con otras que trabajan juntas) requiere comprender la estructura del sistema (sus componentes), su dinámica (como interactúan estos componentes), los métodos de control (como se

regulan sus interacciones) y el diseño (Kitano, 2002). El objetivo de este trabajo es caracterizar uno de los componentes de señalización (las ACs y GCs) de las cianobacterias. La arquitectura molecular de las ciclasas identificadas nos permite inferir aspectos sobre la dinámica del sistema además de su clasificación. Las cianobacterias pueblan en la actualidad la mayor parte de los ecosistemas, tanto acuáticos como terrestres, y presentan una gran capacidad de adaptación a condiciones ambientales muy diferentes (Tandeau de Marsac and Houmard, 1993). La disponibilidad de numerosos genomas de cianobacterias ya secuenciados o en fase de secuenciación permite comparar la abundancia de ACs y GCs en las diferentes especies en función de sus características ecofisiológicas, es decir podemos observar las pautas de diseño del sistema de señalización a lo largo de la evolución.

Materiales y métodos

Para llevar a cabo este estudio se ha consultado la base de datos no redundante de proteínas y la de genomas microbianos del NCBI (Wheeler *et al.*, 2005) así como los programas disponibles a través de internet detallados en la Tabla 1.

Los datos analizados en este trabajo son aquellos depositados en el NCBI en junio de 2005. Entre ellos se encuentran 354 genomas microbianos de los cuales 16 corresponden a cianobacterias. De éstos, 9 han sido secuenciados y ensamblados (*Gloeobacter violaceus* PCC 7421, *Nostoc sp.* PCC 7120, *Prochlorococcus marinus str.* MIT 9313, *Prochlorococcus marinus subsp. marinus str.* CCMP1375, *Prochlorococcus marinus subsp. pastoris str.* CCMP1986, *Synechococcus elongatus* PCC 6301, *Synechococcus sp.* WH 8102, *Synechocystis sp.* PCC 6803, *Thermosynechococcus elongatus* BP-1), mientras que 7 están aun en fase de secuenciación (*Anabaena variabilis* ATCC 2941, *Crocospaera watsonii* WH 8501, *Nostoc punctiforme* PCC 73102, *Synechococcus elongatus* PCC 7942, *Synechococcus sp.* PCC 7002, *Trichodesmium erythraeum* IMS101, *Synechococcus sp.* CC9311).

Inicialmente se interrogó directamente la base de datos de proteínas empleando los términos ((adenyl OR guanylyl OR adenylate OR guanylate) AND cyclase AND cyanobacteri*). La búsqueda dio como resultado numerosos registros anotados automáticamente, que no contenían ningún dominio catalítico de AC o GC y que por consiguiente habían sido erróneamente ano-

Bases de Datos	URL	Referencia
Bases de datos no redundante de proteínas (nr)	http://www.ncbi.nlm.nih.gov/	(Wheeler <i>et al.</i> , 2005)
Genomas microbianos	http://www.ncbi.nlm.nih.gov/genom_table.cgi	(Wheeler <i>et al.</i> , 2005)
RDP Ribosomal Database Project-II Release 9	http://rdp.cmc.msu.edu/index.jsp	(Cole <i>et al.</i> , 2005)
PROGRAMAS		
BLAST y PSI-BLAST	http://www.ncbi.nlm.nih.gov/BLAST/	(Altschul <i>et al.</i> , 1997)
SMART-PFAM	http://smart.cml-heidberg.de/	(Letunic <i>et al.</i> , 2004)
CLUSTALW	http://www.cbi.ac.uk/clustalw/	(Thompson <i>et al.</i> , 1994)
ConPred II	http://bioinfo.si.hirosaki-u.ac.jp/~ConPred2/	(Arai <i>et al.</i> , 2004)
T-Coffee	http://igs-server.cnrs-mrs.fr/Tcoffee/tcoffee_egi/index.cgi	(Notredame <i>et al.</i> , 2000)
BLOCKS	http://molevol.ibmb.csic.es/Gblocks_server/index.html	(Castresana, 2000)
Phylip, Quarted puzzle	http://bioweb.pasteur.fr/#log	(Schmidt <i>et al.</i> , 2002), (Felsenstein, 1996)

Tabla 1. URL de las bases de datos y programas empleados para la realización de este estudio.

tados como ACs o GCs. Esta asignación se debe a un fenómeno denominado transitividad mediante el cual se asigna una identidad enzimática a una fase codante por su similitud con la parte reguladora de un enzima, no por la parte catalítica.

Por ello, la base de datos no redundante de proteínas del NCBI se interrogó empleando el programa PSI-BLAST (Altschul *et al.*, 1997), usando como cebo la secuencia primaria del representante de cada clase de AC y GC mencionado en la introducción. Cuando se conocía la secuencia primaria correspondiente al centro catalítico de la clase buscada, la base de datos se interrogaba exclusivamente con ésta para evitar la identificación de falsos positivos por transitividad. Las secuencias se iteraban hasta la convergencia. En general, con 5 iteraciones era suficiente. Finalmente, el resultado se filtraba para seleccionar las secuencias de cianobacterias que eran guardadas en formato FASTA. Para analizar el genoma de la cianobacteria *Synechococcus* PCC 7002 y *Synechococcus sp.* CC9311, cuyo repertorio de proteínas no ha sido depositado aún en la base de datos no redundante, empleamos tBLASTn. El borrador de los genomas de estas cianobacterias se obtuvo de la base de datos de genomas microbianos del NCBI. La traducción conceptual de las secuencias de nucleótidos de estos dos últimos genomas se llevó a cabo empleando el programa ORFfinder del NCBI.

La arquitectura molecular de las posibles ACs y GCs de clase III identificadas mediante PSI-BLAST y tBLASTn se determinó empleando los programas SMART, PFAM (Letunic *et al.*, 2004) y CDD (Wheeler *et al.*, 2005). El dominio SMART CYCc (COG 2114) corresponde a las secuencias primarias de aminoácidos homólogas a las ACs y GCs de clase III. La posición de posibles regiones transmembranares y su topología se determinó con el programa ConPred II (Arai *et al.*, 2004).

La búsqueda de ACs y GCs de Clase III con arquitectura molecular similar a las observadas en cianobacterias y otros organismos se realizó empleando PSI-BLAST y consultando la base de datos no redundante de proteínas. La arquitectura de los homólogos encontrados se verificó empleando SMART.

Aquellas pautas de lectura abierta que contenían dominios CYCc se alinearon empleando los programas CLUSTAL W (Thompson *et al.*, 1994) y T-Coffee (Notredame *et al.*, 2000). Se alinearon las secuencias completas y no los dominios CYCc debido a que el extremo C-terminal de este último está mal acotado en el programa SMART. Dado que el alineamiento incluía la, recientemente cristalizada, AC CyaC, (Steegborn *et al.*, 2004), ésta se tomo como referencia para identi-

car los determinantes de catálisis y especificidad en cada una de las secuencias (Ochoa de Alda *et al.*, 2000). Las regiones hipervariables del alineamiento se eliminaron con el programa Gblocks 0.91b (Castresana, 2000). Como resultado se obtuvieron bloques que contenían un máximo de 115 residuos por secuencia. Estos bloques contenían todos los determinantes de CyaC de *Spirulina platensis* implicados en la actividad catalítica y el reconocimiento del ATP.

Para la realización de los árboles filogenéticos se emplearon los métodos de: i) Máxima Verosimilitud del programa Puzzle 5.2 (Schmidt *et al.*, 2002), ii) Máxima Parsimonia, y iii) Distancia de Proteínas/NeigbordJoint del paquete de programas Phylip 3.6a3 (Felsenstein, 1996). Para evaluar la topología de los árboles por el método de Máxima Verosimilitud se reconstruyeron 1000 puzzles y para evaluar la topología obtenida por los métodos de máxima parsimonia y distancia de proteínas se realizaron 1000 remuestreos. En todos los casos se empleó el modelo de sustitución JTT (Jones *et al.*, 1992).

Las relaciones evolutivas de las cianobacterias estudiadas se estimaron a partir de la filogenia de la subunidad ribosomal 16S. Las secuencias, ya alineadas, fueron obtenidas de la Base de Datos Ribosomal RDB (Cole *et al.*, 2005), las regiones hipervariables se eliminaron con el programa Gblocks 0.91b (Castresana, 2000) y el árbol se construyó con el método de Máxima Verosimilitud del programa Puzzle 5.2 (Schmidt *et al.*, 2002) empleando el modelo de sustitución de HKY (Hasegawa *et al.*, 1985), dos velocidades de mutación y reconstruyendo 1000 puzzles. Las secuencias de RDB empleadas fueron las siguientes: S000112996 (*Spirulina platensis*), S000427667 (*Trichodesmium erythraeum*), S000427190 (*Thermosynechococcus elongatus* BP-1), S000451897 (*Crocospaera watsonii* WH 8501), S000117544 (*Synechocystis* sp. PCC 6803), S000005917 (*Synechococcus* sp. PCC 7002), S000334410 (*Synechococcus* sp. WH 8102), S000399807 (*Synechococcus* sp. CC9311), S000387808 (*Synechococcus elongatus*; PCC 6301), S000387853 (*Synechococcus elongatus* PCC 7942), S000390024 (*Anabaena cylindrica*), S000112413 (*Anabaena variabilis*), S000391438 (*Nostoc* sp. PCC 7120), S000427839 (*Nostoc punctiforme* PCC 73102), S000334436 (*Prochlorococcus marinus* subsp. *pastoris* str. CCMP1986; MED4), S000334439 (*Prochlorococcus marinus* str. MIT 9313; MIT9313), S000386760 (*Prochlorococcus marinus* subsp. *marinus* str. CCMP1375; CCMP1375; SS120), S000387822 (*Gloeobacter violaceus* PCC 7421) S000433496 (*Thermotoga maritima* SL7). La secuencia de *Thermotoga maritima* se empleó como grupo externo y no se muestra en la figura. La secuencia de *Synechococcus elongatus* PCC 6301 (gi

56684969) y PCC 7942 (gi 5771435) son 100 % idénticas, por eso ocupan la misma posición en el árbol filogenético.

Resultados y Discusión

Clases de AC y GC presentes en las cianobacterias

El análisis comenzó interrogando toda la base de datos de proteínas empleado como cebo los representantes de cada una de las clases (I, II, III, IV, V, VI y VII) descritos en la introducción. Esto nos permitió tener una visión general de la distribución de cada clase en los seres vivos, que pudo ser contrastada con las revisiones más recientes (Baker and Kelly, 2004; Cann, 2004; Iyer and Aravind, 2002; Shenroy and Visweswariah, 2004) y, de esta forma, pudimos valorar la bondad del método empleado. En términos generales, los resultados de la distribución de las diferentes clases de AC y GC en los seres vivos sigue las pautas descritas en los trabajos citados (Resultados no mostrados). No encontramos ningún homólogo de las ACs y GCs de Clase I, II, V, VI y VII en el *Phylum Cyanobacteria*. La mayoría de los homólogos de ACs/GCs que encontramos en este *Phylum* eran de clase III. Excepcionalmente, en el extremo C-terminal de la fase codante gll0897 (gi | 37520048) perteneciente a la cianobacteria *G. violaceus* PCC 7421 encontramos un dominio CYTH/CyaB homólogo a las ACs de Clase IV. Concretamente, este dominio CYTH/CyaB de gll0897 mostraba una identidad del 22 % y una similitud del 35 % con la AC CyaB de *A. hydrophila*. La falta de estudios de estructura-función relativos a las ACs de Clase IV impidió asignar una función a gll0897 en base a su secuencia primaria, ya que el dominio CYTH/CyaB ha sido relacionado con numerosas actividades enzimáticas del metabolismo de nucleótidos y polifosfatos (Iyer and Aravind, 2002). Para averiguar cual era la actividad enzimática más relacionada con gll0897 comparamos el dominio CYTH/CyaB de gll0897 con la base de datos no redundante de proteínas empleando el programa PSI-BLAST e iterando hasta la convergencia (8 iteraciones). Los resultados indican que gll0897 es homóloga a las GTP piroquinasas ($E=10^{-40}$) y no es probablemente una AC o una GC.

Por consiguiente, es probable que las cianobacterias tengan únicamente ACs y GCs de Clase III.

Frecuencia de las adenilil y guanilil ciclasa de Clase III en las cianobacterias

La búsqueda de ACs y GCs de clase III en las bases de datos del NCBI (junio 2005) dio como resultado 56 secuencias de proteínas de cianobacterias que contenían un dominio CYCc. De ellas, 7 correspondían a organismos cuyo genoma no ha sido completamente secuenciado como *S. platenses* (3 CYCcs; gi|2575805, gi|11990887, gi|7470942), *A. cylindrica* (1 CYCc; gi|1169150) y *Anabaena sp.* (2 CYCcs; gi|1754640; gi|1754642 y un fragmento gi|600251) y 50 correspondían a secuencias encontradas en los 16 genomas que han sido completamente secuenciados (Fig. 1).

La frecuencia del dominio CYCc en estos genomas es muy variable. Así, no se encuentra ningún homólogo CYCc en el genoma de las cianobacterias unicelulares marinas *P. marinus subsp. marinus str.* CCMP1375, *P. marinus subsp. pastoris str.* CCMP1986 y *Synechococcus sp.* WH 8102. En el genoma de la cianobacteria costera *Synechococcus sp.* PCC 7002 se encontró una secuencia homóloga a CYCc que, probablemente, no es funcional ya que se observó fragmentada. Cuando se unieron las dos fases codificantes fragmentadas de esta proteína se obtuvo una secuencia cuya arquitectura molecular era similar a CyaD de *Nostoc* PCC 7120. Otros genomas de cianobacterias unicelulares marinas como los de *Synechococcus sp.* CC9311 y *P. marinus str.* MIT 9313 contenían 1 (contig: 1047078002995) y 2 (gi|33864416, gi|33864410) homólogos de CYCc, respectivamente, mientras que el de la cianobacteria marina unicelular diazotrofa *C. watsonii* WH 8501 contenía 3 (gi|53734915, gi|46118802, gi|46119054). Dentro del grupo de las cianobacterias marinas el mayor número de homólogos de CYCc se encontró en la cianobacteria diazotrofa filamentosa *T. erythraeum* IMS101 que contenía 13 secuencias homólogas a las ACs/GCs de Clase III (gi|48893394, gi|48894764, gi|4889041, gi|48890832, gi|48892879, gi|48892878, gi|48895515, gi|48893696, gi|48895220, gi|48892865, gi|48891058, gi|48891559, gi|48893689).

Todos los genomas de cianobacterias de agua dulce contenían algún homólogo de CYCc. El menor número se observó en las cianobacterias unicelulares como *G. violaceus* PCC 7421 (1 CYCc; gi|37520048), *T. elongatus* BP-1 (2 CYCcs; gi|22299823, gi|22299953), *S. elongatus* PCC 6301 (2 CYCcs; gi|56750874, gi|56751911) *S. elongatus* PCC 7942 (2 CYCcs; gi|46130360, gi|53762857) y *Synechocystis sp.* PCC 6803 (3 CYCcs; gi|16330472,

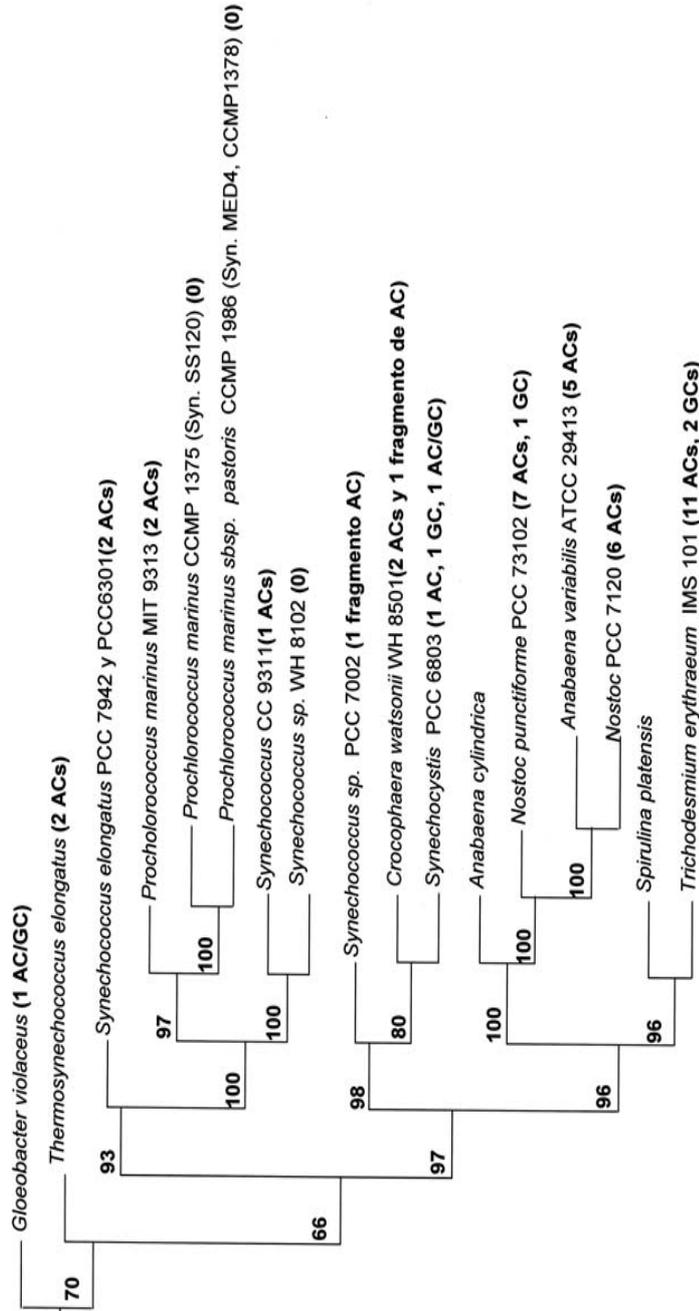


Figura 1: Relaciones evolutivas de las cianobacterias mencionadas en el texto. El árbol filogenético se elaboró a partir de las secuencias de la subunidad 16S ribosomal. Junto al nombre de cada una de las especies cuyo genoma ha sido completamente secuenciado se indica entre paréntesis y en negrita la abundancia de adenilil (AC) y guanilil (GC) ciclasas que se ha observado de este estudio. Junto al nombre de algunas especies se indican los sinónimos (Syn.) con el que también son conocidas. Las dos cepas, PCC 7942 y PCC 6301, de *Synechococcus elongatus* se sitúan contiguas porque están muy próximas evolutivamente. La abundancia de estas enzimas en las especies cuyo genoma no ha sido completamente secuenciado se describe en el texto. El número situado junto a cada nudo del árbol representa el porcentaje de veces que se observaba ese agrupamiento durante la reconstrucción de los posibles árboles filogenéticos y es una medida de la fiabilidad de la topología mostrada. El grupo externo empleado para ubicar la raíz del árbol filogenético fue la secuencia 16 S de *Thermotoga maritima* SL7.

gi|16329561, gi|16330418). Dentro de las cianobacterias de agua dulce, el mayor número de homólogos de CYCc se encontró en las cianobacterias filamentosas que diferencian heterocistos como *A. variabilis* (5 CYCcs; gi|45509234, gi|53765262, gi|53764651, gi|45509458, gi|45506866), *Nostoc* PCC 7120 (6 CYCcs; gi|17228613, gi|17229758, gi|17229396, gi|17232455, gi|17228238, gi|17228157) y *N. punctiforme* PCC 73102 (8 CYCcs; gi|23127996, gi|23127822, gi|23126551, gi|23130559, gi|23126594, gi|23126862, gi|23125697, gi|23127008).

Los resultados sobre frecuencia de dominios CYCc en los genomas completos sugieren que, como regla general, este dominio es: i) más frecuente en cianobacterias de agua dulce que en cianobacterias marinas; ii) más frecuente en cianobacterias filamentosas que en unicelulares y iii) más frecuente en cianobacterias diazotrofas que en las no fijadoras de nitrógeno. Además, parece no ser fundamental para llevar a cabo la fotosíntesis oxigénica, característica común de todas las cianobacterias, ya que se encuentra ausente en algunas especies. De acuerdo con esta última observación la inactivación de la AC Cya1 de *Synechocystis* PCC 6803 disminuye la concentración intracelular de cAMP hasta niveles indetectables y no modifica los parámetros de crecimiento de la cianobacteria (Terauchi and Ohmori, 1999).

Arquitectura molecular de los homólogos de adenilil y guanilil ciclasa de Clase III

De las 56 ACs/GCs de Clase III detectadas eliminamos para estudios posteriores 1 correspondiente a un fragmento (gi|600251) y estudiamos la arquitectura de las 55 restantes. Entre éstas, se pudieron encontrar 20 arquitecturas moleculares diferentes que contenían entre otros: dominios transmembranales, dominios sensores y uno o varios dominios catalíticos (Figura 2). Ninguna de las secuencias mostraba más de un dominio CYCc. El dominio CYCc se encontraba en el extremo C-terminal de las proteínas salvo cuando éste abarcaba la proteína completa.

A efectos de clasificar las arquitecturas y con el fin de agrupar y nombrar el gran número de arquitecturas moleculares observadas se consideraron los dominios que abarcaban más de 35 residuos. De esta forma no se emplearon para definir una nueva arquitectura ni el número de regiones transmembra-

na, ni la presencia de repeticiones de tetratricopéptidos. Las arquitecturas moleculares de ACs y GCs observadas en cada una de las cianobacterias se muestran en la Figura 2 y se describen a continuación

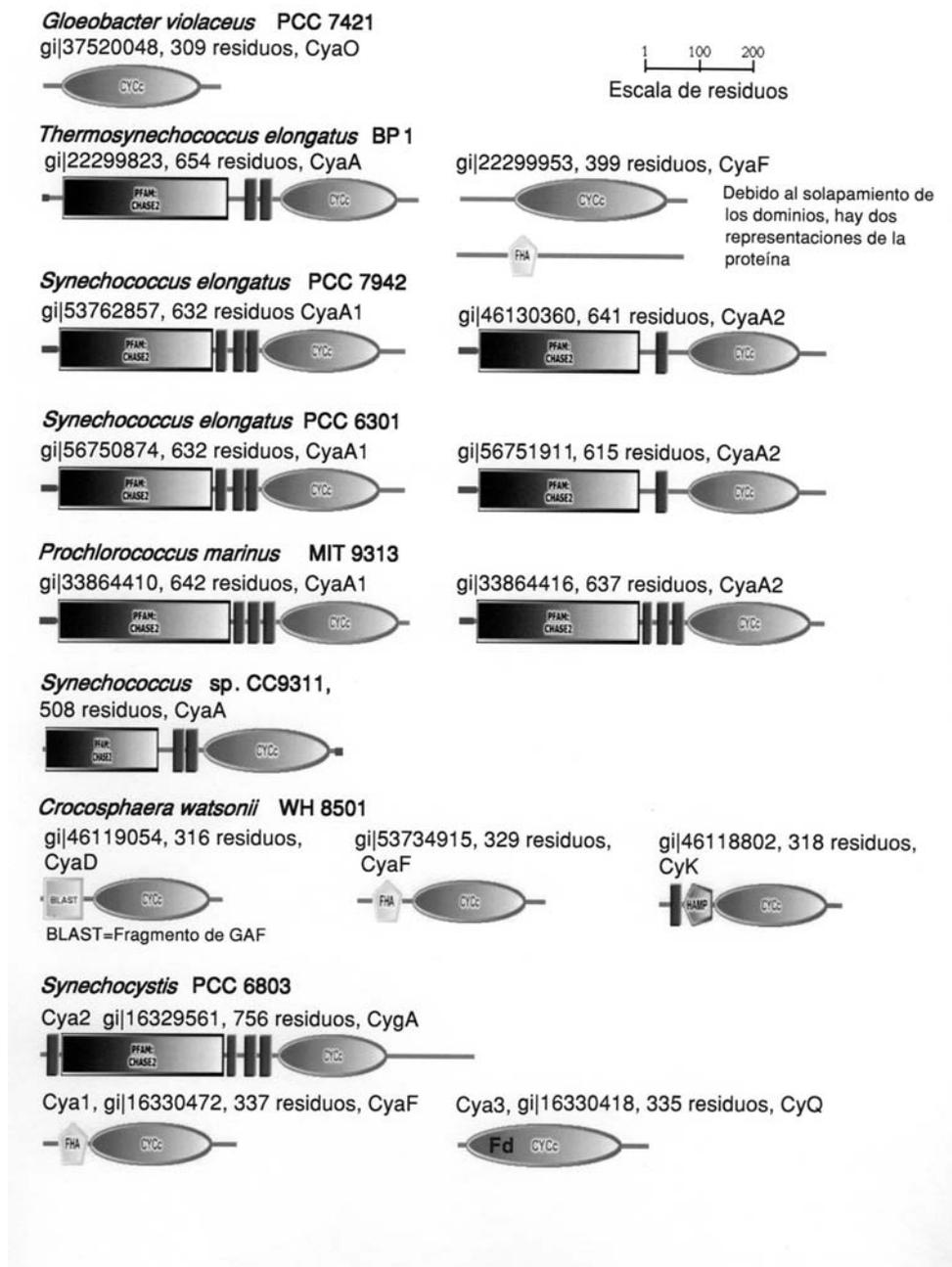
Arquitectura A

Recibe el nombre de la arquitectura de la AC de membrana CyaA de *Nostoc* PCC 7120 (Katayama and Ohmori, 1997). Esta compuesta de un dominio CHASE2 flanqueado, en general, por 1 dominio transmembranal en el extremo N-terminal y 3 en el extremo C-terminal, al que se une el dominio CYCc. Esta arquitectura es la más frecuente entre las ACs/GCs de cianobacterias, encontrándose en 14 de las 55 estudiadas. En algunos casos (gi|33864410, gi|46130360, gi|53762857, gi|56750874 y gi|56751911) la región transmembranal N-terminal se solapa con un posible péptido señal. Esto sugiere que la proteína podría ser secretada a través de la membrana citoplásmica o la tilacoidal. Empleando el programa ConPredII se observa que la topología más frecuente es aquella en la que el dominio CHASE2 está en el exterior de la membrana (periplasma o lumen) y el dominio CYCc en el citoplasma. CHASE2 es probablemente un dominio sensor de membrana observado en numerosos organismos tanto en la forma de secuencia aislada como formando parte de la región sensora de adenilil ciclasas, diguanilato ciclasas y proteínas de la quimiotaxis (Zhulin *et al.*, 2003).

Esta arquitectura molecular *A* de ACs/GCs no es exclusiva del *Phylum Cyanobacteria* sino que también se puede observar en los *Phyla Proteobacteria* y *Spirochaetes* (Shenroy and Visweswariah, 2004). Esto sugiere que se pudiera haber originado antes de la divergencia de las cianobacterias del resto de las bacterias o que la arquitectura haya surgido más de una vez en eventos independientes. Es remarcable que en algunas cianobacterias como *P. marinus* MIT 9313, *S. elongatus*, *T. erythraeum* IMS101 se observan duplicaciones de la arquitectura *A*, igual que en la δ -*Proteobacteria Myxococcus xanthus*. En esta bacteria las dos adenilil ciclasas de arquitectura *A* se han relacionado con la osmoregulación (Kimura *et al.*, 2005).

Arquitectura B

Recibe el nombre de la arquitectura de las ACs solubles CyaB1 y CyaB2 de *Nostoc* PCC 7120 (Katayama and Ohmori, 1997). Esta arquitectura esta compuesta de dos dominios GAF (GAF-A y GAF-B) situados en el



Anabaena cylindrica

gij1169150, 502 residuos, CyaJ



Nostoc punctiforme PCC 73102

gij23127822, 742 residuos, CyaA



gij23127996, 546 residuos, CyaD



gij23126551, 1177 residuos, CyaC



gij23125697, 858 residuos, CyaB1



gij23126594, 149 residuos, CyaO



gij23127008, 861 residuos, CyaB2



gij23130559, 616 residuos, CygL



gij23126862, 433 residuos, CyaT



Nostoc PCC 7120

CyaA, gij17228613, 735 residuos, CyaA



CyaD, gij17228238, 546 residuos, CyaD



CyaE, gij17228157, 472 residuos, CyaE



Debido al solapamiento de los dominios, hay dos representaciones de la proteína

CyaB1, gij17229758, 859 residuos, CyaB1



CyaB2, gij17229396, 860 residuos, CyaB2



CyaC, gij17232455, 1155 residuos, CyaC



Anabaena sp.

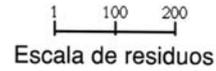
CyaA, gi|53765262, 735 residuos, CyaA



CyaB2, gi|1754640, 860 residuos, CyaB2



CyaC, gi|1754642, 1155 residuos, CyaC



Anabaena variabilis ATCC 29413

CyaA, gi|53765262, 735 residuos, CyaA



CyaD, gi|45509234, 546 residuos, CyaD



CyaB1, gi|45509458, 859 residuos, CyaB1



Debido al solapamiento de los dominios, hay dos representaciones de la proteína

CyaB2, gi|45506866, 860 residuos, CyaB2



CyaC, gi|53764651, 1156 residuos, CyaC

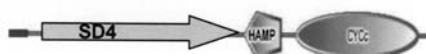


Spirulina platensis

gi|2575805, 494 residuos, CyaN



CyaG, gi|11990887, 671 residuos, CyaG



CyaC, gi|7470942, 1202 residuos, CyaC



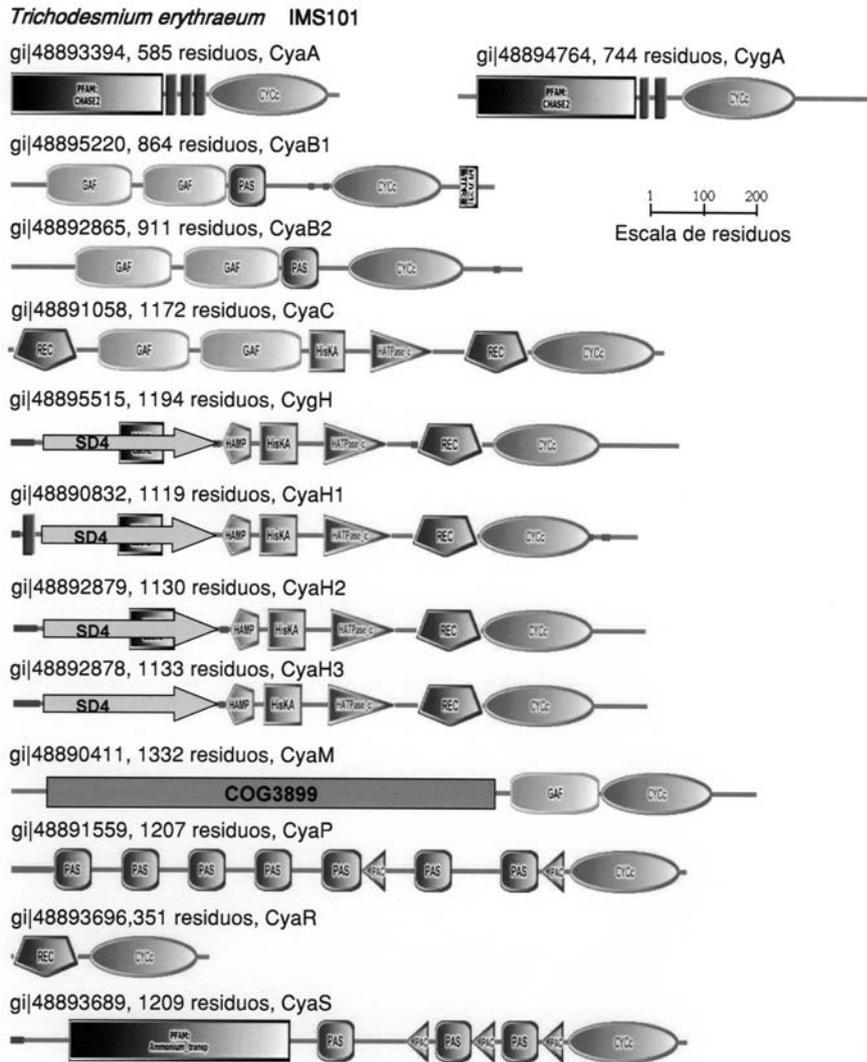


Figura 2: Representación gráfica de la arquitectura molecular SMART, PFAM, PSI-BLAST y CDD de las adenilil y guanilil ciclasas de cianobacterias. Sobre cada arquitectura se muestra la denominación actual de la proteína (si ha sido descrita anteriormente), seguida del identificador del gen (gi) en el Genbank, el número de residuos de la proteína, y la denominación propuesta en este estudio. Las arquitecturas se muestran guardando la proporcionalidad según se muestra en la escala de residuos. La barra horizontal representa a la secuencia de aminoácidos. Las líneas de color rojo sobre esta barra representan las secuencias correspondientes a posibles péptidos señal. Las barras verticales de color azul, indican regiones transmembrana. Las líneas verdes se corresponden con regiones coiled-coil. El nombre del resto de los dominios se muestra en el interior de los dibujos y se describen más detalladamente en el texto.

extremo N-terminal, junto a un dominio PAS al que se une el dominio CYCc. La arquitectura *B* la observamos en 9 de las 55 ACs/GCs estudiadas y se encontró únicamente en cianobacterias filamentosas fijadoras de nitrógeno. En estas cianobacterias, cuyo genoma ha sido secuenciado por completo, detectamos dos isoformas de CyaB, CyaB1 y CyaB2. En *Nostoc* PCC 7120 ambas ACs son enzimas homodiméricas activadas por cAMP, en CyaB1 a través de la unión de este ligando a los dominios GAF-B de los homodímeros y en CyaB2 a través de la unión del cAMP a los dominios GAF-A y GAF-B de ambos homodímeros (Bruder *et al.*, 2005; Kanacher *et al.*, 2002; Tada *et al.*, 2001). La función del dominio PAS de CyaB no ha sido aún descrita. El dominio PAS pertenece a una familia de dominios de proteínas sensores relacionados con la percepción de oxígeno, luz, potencial redox, entre otros, e involucrados en la transmisión de señales en una amplia variedad de organismos. Todos los dominios PAS parecen tener una estructura y una flexibilidad conformacional similar (Vreede *et al.*, 2003) que les permite regular una gran variedad de dominios proteicos adyacentes (Galperin, 2004). En consecuencia, es probable que el dominio PAS de CyaB también participe en la modulación, a través de la unión de un ligando o la recepción de una señal, de la actividad del dominio CYCc.

Una extensa búsqueda en la base de datos no redundante de proteínas empleando PSI-BLAST indica que, al contrario que la arquitectura *A*, la arquitectura *B* parece ser exclusiva de cianobacterias.

Arquitectura C

Recibe el nombre de la arquitectura de la AC soluble CyaC de *Nostoc* PCC 7120 (Katayama and Ohmori, 1997). Esta arquitectura está compuesta por un dominio REC-A (en el extremo N-terminal) al que se une, por este orden, dos dominios GAF, un dominio HisKA, un dominio HATPasec, otro dominio REC-B y, en el extremo C-terminal, un dominio CYCc. Los dominios REC, HisKA y HATPasec constituyen el eje central de los sistemas de transducción de señales de dos componentes (Stock *et al.*, 2000). Esta arquitectura fue observada en 6 de las 55 ACs/GCs estudiadas y se encontró únicamente en las cianobacterias filamentosas estudiadas. La caracterización bioquímica indica que CyaC es realmente una proteína bifuncional en la que el dominio HATPasec hidroliza ATP y fosforila al dominio HisKA que, a su vez, fosforila al dominio REC-B. La fosforilación de REC-B aumenta la actividad AC

(Kasahara and Ohmori, 1999). Se desconoce la función de los dominios GAF pero, por homología con otros sistemas (Galperin, 2004), es probable que los dominios GAF funcionen como sensores que regulan la actividad del dominio ATPasa (HATPasec) a través de la unión de un ligando. Como los dominios REC controlan la actividad de los dominios que tienen unidos en el extremo C-terminal (Stock *et al.*, 2000), es probable que el dominio REC-A controle al sensor GAF. La fosforilación de REC-A, que finalmente controla la actividad AC, depende de una proteína similar al fitocromo que se activa con luz visible de 720 nm (Okamoto *et al.*, 2004).

Al igual que la arquitectura molecular B, la arquitectura C parece que también es exclusiva de cianobacterias.

Arquitectura D

Recibe el nombre de la arquitectura de la AC soluble CyaD de *Nostoc* PCC 7120 (Katayama and Ohmori, 1997). Está compuesta por un dominio FHA N-terminal al que se une un dominio GAF seguido de un dominio CYCc. Esta arquitectura es exclusiva de cianobacterias y fue observada únicamente en las 3 cianobacterias heterocistadas cuyo genoma ha sido completamente secuenciado, *Nostoc* PCC 7120, *N. punctiforme* PCC 73102 y *A. variabilis* ATCC 29413. En la proteína gi|46119054 de la cianobacteria unicelular *C. watsonii* WH 8501 el dominio CYCc se une a un fragmento de un dominio GAF. En el genoma esta secuencia está intercalada por dos transposones lo que sugiere que ha podido ser inactivada por transposición. Como se muestra más adelante, la posición en el árbol filogenético de esta AC indica que podría ser una forma degenerada de CyaD. En la cianobacteria *Synechococcus* PCC 7002 la secuencia de una AC similar está interrumpida (cambia de fase en la pauta de lectura) al inicio del dominio CYCc, pudiéndose tratar también de una forma degenerada señalada en la Figura 5 como CyaD'.

Esta AC no ha sido caracterizada bioquímicamente pero se tiene evidencias genéticas de su actividad AC a través de la complementación funcional de un mutante de *E. coli* carente de AC (Katayama and Ohmori, 1997). Presumiblemente la actividad AC es regulada por los dominios FHA y GAF. El dominio GAF podría regular el dominio CYCc a través de la unión de un ligando. El dominio FHA reconoce fosfoproteínas (Hammet *et al.*, 2003) integrando los eventos intracelulares de la transmisión de señales a través de la formación de complejos multiprotéicos con sustratos de proteínas serín-treonín

quinasas. (Molle *et al.*, 2004; Yaffe and Smerdon, 2004). Esto sugiere que la arquitectura *D* podría permitir el cruce de los sistemas señalización de ACs/GCs y serín-treonín quinasas.

Arquitectura E

Recibe el nombre de la arquitectura de la AC CyaE de *Nostoc* PCC 7120 (Kasahara *et al.*, 2001). En el extremo N-terminal se sitúan dos hélices transmembranales que flanquean una secuencia larga extracitoplásmica de 160 residuos. La hélice N-terminal es identificada como péptido señal con el programa SMART, por eso la proteína CyaE de *Nostoc* PCC 7120 (gi|17228157), se representa con una sola hélice transmembrana en la figura 2. En el extremo C-terminal se localiza el dominio CYCc.

La interrogación de la base de datos no redundante de proteínas indicó que esta arquitectura se encuentra únicamente en *Nostoc* PCC 7120. Resulta significativo que no se observen arquitecturas similares en otros seres vivos, ni siquiera en cianobacterias evolutivamente próximas como *N. punctiforme* o *A. variabilis*.

Por analogía con las GCs eucariotas, es probable que el largo *loop* externo de 160 residuos tenga una función sensora. La interrogación de la base de datos no redundante de proteínas con PSI-BLAST, empleando la secuencia externa de 160 aminoácidos como cebo, indica que esta secuencia externa muestra homología con regiones de proteínas que unen ligandos como la región similar a la quinasa del péptido natriurético (gi|2118323), la secuencia implicada en la unión al ATP o el GTP de la enzima glutatión sintetasa denominada *ATP-grasp* (gi|16130848) y proteínas que unen Leu o Val (gi|40889813). La comparación de los alineamientos con las estructuras cristalinas de estas proteínas (registros del banco de datos de proteínas 1GSA, 2LIV, 1USK) permiten observar que los residuos implicados en la unión de los ligandos, ADP y Leucina, no están conservados en la región externa de CyaE.

En conjunto, los resultados sugieren que el péptido externo es un dominio sensor que nosotros denominamos SD3, aún no incluido en las bases de datos de dominios protéicos, con un plegamiento característico de proteínas que unen ligandos. Este dominio podría controlar la actividad AC membranal.

Arquitectura F

Esta arquitectura es característica de Cya1 de PCC6803 (Terauchi and Ohmori, 1999). Se compone de un dominio FHA N-terminal al que se une un

dominio CYCc (Ochoa de Alda and Houmard, 2000). En *Synechocystis* PCC 6803, esta enzima es responsable de los niveles intracelulares de cAMP, es necesaria para la movilidad celular, es específica para el ATP e inhibida por bicarbonato (Masuda and Ono, 2004; Masuda and Ono, 2005; Terauchi and Ohmori, 1999, 2004). Esta arquitectura se observa en 3 de las 55 ACs/GCs de cianobacterias estudiadas. Todas ellas pertenecen a cianobacterias unicelulares. Como se ha citado anteriormente el dominio FHA también forma parte de la arquitectura *D* pero no forma parte de arquitecturas moleculares de ACs o GCs de otros seres vivos.

La denominación Cya1 no es correcta ya que según la IUPAC (IUPAC-IUB, 1989) los números deben emplearse para diferenciar las isoformas de enzimas. Por ello, como se indicará posteriormente, proponemos el cambio de denominación a CyaF.

Arquitectura G

Recibe el nombre de la arquitectura de la AC CyaG (gi|11990887) de *S. platenses* (Kasahara *et al.*, 2001), única cianobacteria en la que se encuentra. Se compone de un dominio sensor de unos 310 aminoácidos, que nosotros denominamos SD4, flanqueado de 2 hélices transmembranales a cuyo extremo C-terminal se une un dominio HAMP que conecta el dominio CYCc. El dominio sensor muestra similitudes con los dominios sensores de numerosas histidina quinazas del sistema de dos componentes, diguanilato ciclasas y quimiorreceptores (resultados no mostrados) y contiene en su interior un dominio CACHE (PFAM 02743) entre los residuos 200 y 250 que ya ha sido descrito en canales de Ca²⁺ y quimiorreceptores (Anantharaman and Aravind, 2000). El dominio HAMP se encuentra frecuentemente en la membrana plasmática uniendo un dominio sensor externo con un dominio de señalización citoplásmico y juega un papel importante en la regulación por fosforilación o metilación de receptores homodiméricos mediante la transmisión de cambios conformacionales desde dominios periplásmicos de unión a ligando a dominios aceptores de señales citoplásmicos (Linder *et al.*, 2004; Ma *et al.*, 2005). En consecuencia, estos dominios HAMP podrían actuar modulando la actividad AC (Linder *et al.*, 2004).

El dominio sensor unido al dominio HAMP (los 430 aminoácidos N-terminales) muestra una similitud del 57 % y una identidad del 36 % con la misma región de las ACs/GCs de arquitectura *H* gi|48890832, gi|48892878,

gi|48892879, gi|48895515 de *T. erythraeum* IMS101. A diferencia de la arquitectura *G*, en las ACs/GCs de arquitectura *H*, el dominio CYCs esta precedido de una dominio histidina quinasa y un regulador de respuesta.

La arquitectura *G* se encuentra en proteínas de otros organismos evolutivamente distantes a las cianobacterias como en la AC gi|48834441 de la *Proteobacteria Magnetococcus* sp. MC-1 y la AC gi|15073188 de *Sinorhizobium meliloti*. La similitud de los dominios sensores de estas proteínas con el de CyaG de *S. platenses* es del 30 %. El dominio sensor no debe estar unido necesariamente al dominio HAMP para su funcionalidad ya que se encuentra en otras ACs unido directamente al dominio CYCc como en el caso de la AC1 de *Legionella pneumophila* subsp. *pneumophila* (gi|52841553).

Arquitectura *H*

La arquitectura *H* se compone de un extremo N-terminal similar al de la arquitectura *G* (un dominio sensor de unos 310 aminoácidos flanqueado por dos hélices transmembranales a cuyo extremo C-terminal se une un dominio HAMP) y un extremo C-terminal similar al de la arquitectura *C* (la secuencia de un dominio HisKA, un dominio HATPasec y un dominio REC al que está unido el dominio CYCc). Es probablemente una proteína bifuncional, histidina quinasa y AC/GC.

Esta arquitectura se observa únicamente en 4 proteínas (H1, H2, H3 y H4) de la cianobacteria marina *T. erythraeum* IMS101. Al igual que en la arquitectura *G*, el programa ConPredII predice en los 4 casos de arquitectura *H* la presencia de 2 hélices transmembranales, éstas hélices están enmascaradas en la arquitectura SMART por un péptido señal (en el extremo N-terminal) y el dominio HAMP (Figura 2).

Una arquitectura similar, pero sin el dominio CYCc, se encuentra en proteínas de *Nostoc* PCC 7120 (Ohmori *et al.*, 2001) y *N. punctiforme* PCC 73102 (Resultados no mostrados).

Arquitectura *J*

La arquitectura *J* corresponde exclusivamente a una AC descrita en *A. cylindrica* (Katayama *et al.*, 1995). Se compone de dos hélices transmembranales, entre las que se sitúa una secuencia externa de 150 residuos, y un dominio HAMP, que incluye una de las hélices, que conecta el dominio CYCc C-terminal. Se ha probado su actividad AC y su localización en la membrana tilacoidal

(Katayama *et al.*, 1995). En este contexto, el programa ConPredII predice que el dominio CYCc se encuentra en el citoplasma y el *loop* de 150 residuos que conecta las dos hélices en el lumen. Este dominio intratilacoidal podría corresponder al dominio sensor, que nosotros denominamos SD1, ya que se encuentran numerosos homólogos de esta secuencia en ACs, receptores de sistemas de dos componentes y de quimiotaxis (resultados no mostrados). Este dominio sensor aún no ha sido incluido en las bases de datos de dominios proteicos. Como se ha comentado anteriormente, el dominio HAMP podría modular la acción del receptor. La arquitectura J no es exclusiva de cianobacterias ya que se encuentran ACs con arquitecturas homólogas en *L. pneumophila subsp. Pneumophila* (gi|52841364 y gi|54294071)].

Arquitectura K

La arquitectura K se compone únicamente de un dominio HAMP que une una hélice transmembranal N-terminal y un dominio CYCc C-terminal. Esta arquitectura es similar a la observada en *Leptospira interrogans* (Shenroy and Visweswariah, 2004). Entre las cianobacterias esta arquitectura se encuentra únicamente en una proteína (gi|46118802) de *C. watsonii* WH 8501.

Arquitectura L

La arquitectura L se compone de un dominio transmembranal (en la arquitectura SMART identificado como péptido señal) seguido de 109 residuos que unen el dominio CYCc C-terminal. Entre las cianobacterias, esta arquitectura solo se observó en la proteína gi|23130559 de *N. punctiforme* PCC 73102. Se observaron proteínas similares en las proteobacterias *Rhodospseudomonas palustris* (gi|39648715), *Bradyrhizobium japonicum* (gi|27355907) y *Silicibacter* sp. TM1040 (gi|52009161). Éstas muestran similitudes incluso en la secuencia de 109 aminoácidos lo que sugiere que este podría ser el dominio sensor, que nosotros denominamos SD2 que regula la actividad de dominio CYCc. Este dominio sensor aún no ha sido incluido en las bases de datos de dominios proteicos.

Arquitectura M

La arquitectura M se compone de un dominio similar a ATPasas (COG3899) seguido de un dominio GAF al que se une un dominio CYCc. Aunque pueda parecer una proteína bifuncional, ATPasa, AC/GC, la actividad

ATPasa es dudosa ya que carece de la secuencia Walter B , implicado en la catálisis, dentro del *P-Loop* característico de las ATPasas (resultados no mostrados). Como ya se ha comentado previamente (Noubir *et al.*, 2002) este dominio ATPasa podría estar implicado en el ensamblaje y desensamblaje de proteínas. En este contexto podría regular el dominio CYCc tal y como se ha sugerido para la proteína RcaD (Noubir *et al.*, 2002). El dominio GAF podría funcionar como un sistema de control adicional.

Aunque en las proteínas de señalización de cianobacterias la fusión de un dominio ATPasa (COG3899) y GAF es común, la arquitectura *M* completa solo se observa en la proteína gi|48890411 de *T. erythraeum* IMS101. También se encuentra completa en proteínas de otros organismos como la AC gi|41398775 de *M. avium* subsp. *paratuberculosis* str. *k10*. El dominio COG3899 es muy común en el extremo C-terminal (no en el N-terminal como la arquitectura *M*) de numerosas ACs de clase III (Shenroy and Visweswariah, 2004).

Arquitectura *N*

La arquitectura *N* se compone de un péptido señal unido a un dominio CHASE3 al que se une una hélice transmembranal que conecta el dominio intracitoplásmico C-terminal. Se observa únicamente en la proteína gi|2575805 de *S. platenses*. El dominio CHASE3 es un dominio sensor extracelular presente en numerosos receptores transmembranales (Zhulin *et al.*, 2003). La señal que percibe es aun desconocida. Una extensiva búsqueda de arquitecturas similares empleando PSI-BLAST indica que esta AC se encuentra exclusivamente en *S. platenses*.

Arquitectura *O*

La arquitectura *O* se compone exclusivamente del dominio CYCc. Entre las ACs/GCs de cianobacterias, esta arquitectura se observa en la proteína gi|23126594 de *N. punctiforme* PCC 73102 y en la proteína gi|37520048 de *G. violaceus*. Esta arquitectura ya se ha observado en ACs/GCs de otras bacterias (Shenroy and Visweswariah, 2004). En este caso, la regulación del dominio CYCc no puede ser inferida de su arquitectura

Arquitectura *P*

La arquitectura *P* se compone de una secuencia de 7 dominios PAS unidos a un dominio CYCc C-terminal. El extremo C-terminal de dos dominios PAS lleva unido dos dominios PAC. Entre las cianobacterias estudiadas,

esta estructura solo se observa en la proteína gi|48891559 de *T. erytraeum* IMS101. Los dominios PAS pertenecen a una familia de dominios de proteínas sensoras involucradas en la traducción de señales en una amplia variedad de organismos. Entre las señales que perciben se encuentran variaciones en la tensión de oxígeno, potencial redox y cambios en el espectro luminoso. El dominio PAC forma parte del extremo C-terminal de los dominios PAS y se ha propuesto que contribuye al plegamiento de PAS (Galperin, 2004). La sucesión y diversidad de los dominios PAS de esta arquitectura sugiere que estos podrían controlar el dominio CYCc integrando varias señales.

Aunque la sucesión de dominios PAS ya se ha observado en otras proteínas de señalización de cianobacterias (Ohmori *et al.*, 2001) este es el único caso en el que se encuentran unidos a un dominio CYCc y es exclusivo de esta proteína soluble (gi|48891559) de *T. erytraeum* IMS101. No ha sido observada en ningún otro organismo

Arquitectura Q

La arquitectura *Q* es característica de la AC/GC Cya3 (gi|16330418) de *Synechocystis* PCC 6083. Se compone de un dominio CYCc en el interior de cuyo extremo N-terminal se sitúa probablemente un grupo sulfoferro (Ochoa de Alda *et al.*, 2000), que probablemente controla el dominio CYCc. Aunque esta arquitectura forma parte de proteínas de ACs/GCs de otros organismos (Shenroy and Visweswariah, 2004) es el único caso en el que no tiene algún dominio adicional.

Arquitectura R

La arquitectura *R* se compone de un dominio REC N-terminal al que se une un dominio CYCc. Esta arquitectura ya ha sido descrita en una AC de la *Proteobacteria* *B. japonicum* (Shenroy and Visweswariah, 2004) y la *Myxobacteria* *Stigmatella aurantiaca* (Coudart-Cavalli *et al.*, 1997). Como se ha comentado en la arquitectura *C*, la fosforilación del dominio REC podría controlar la actividad del dominio CYCc a través de un sistema de dos componentes.

En los genomas de las cianobacterias estudiadas esta arquitectura *R* solo se observa en la proteína gi|48893696 de *T. erytraeum* IMS101.

Arquitectura S

La arquitectura *S* se compone de 11 hélices transmembraneles que

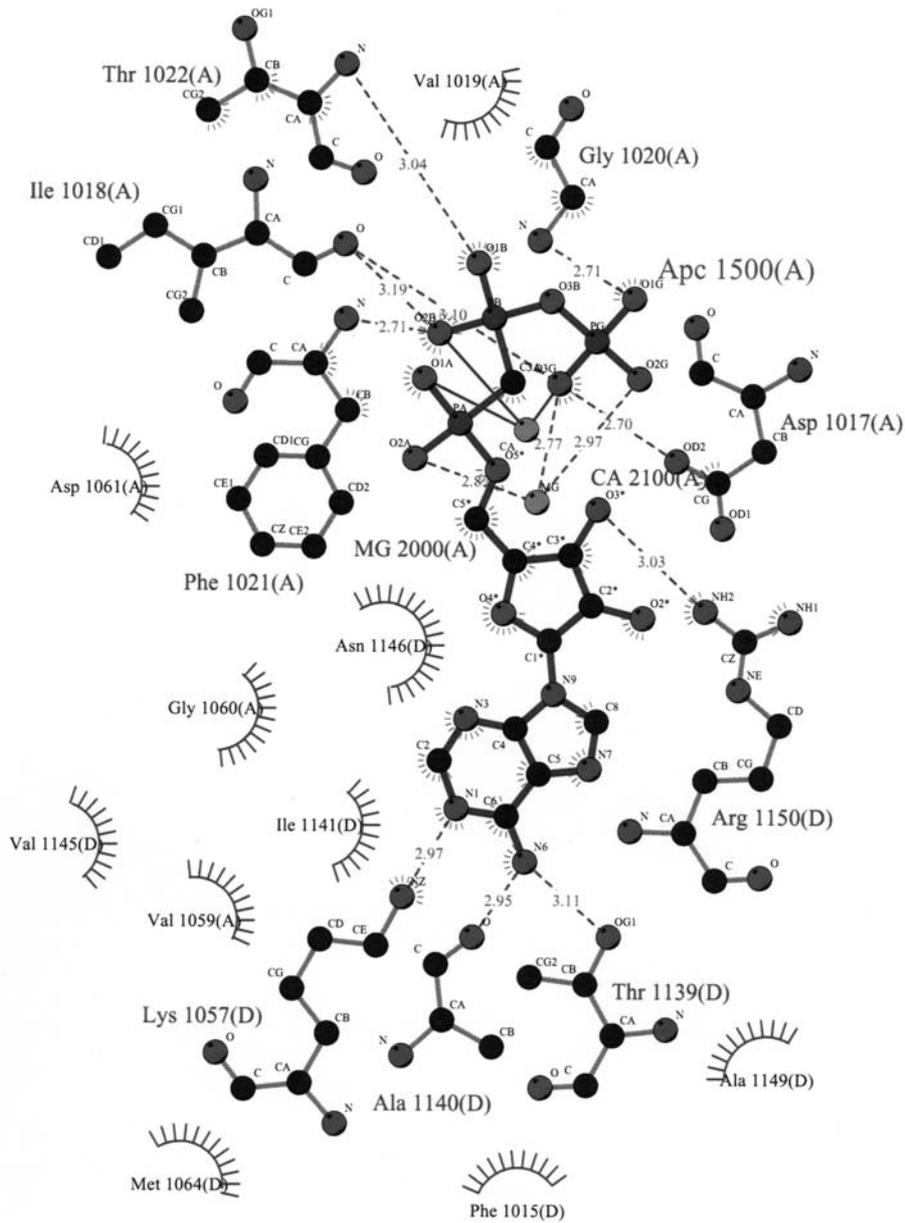


Figura 3: Mapa de interacciones del ATP (LIGPLOT) con los residuos de la proteína CyaC de *Spirulina platensis*. El nombre de cada aminoácido se abrevia con tres letras y junto a ellas se especifica su posición en la secuencia primaria de la proteína y, entre paréntesis la cadena del homodímero (A o D) al que pertenece. Los datos se obtuvieron de la estructura 1wc5 del banco de datos de proteínas (PDB)

comprenden un dominio similar a un transportador de amonio (PFAM00909.11) que está unido a tres secuencias consecutivas de los dominios PAS-PAC, ya descritos en la arquitectura *P*, y un dominio CYCc en el extremo C-terminal. El alineamiento del dominio similar al transportador de amonio con el transportador de amonio AmtB de *Escherichia coli* (Zheng *et al.*, 2004) muestra que el dominio de la arquitectura *S* contiene todos los determinantes necesarios para el transporte y para la especificidad por el amonio de AmtB (Resultados no mostrados). Por tanto, se trata probablemente de una proteína bifuncional AC/GC y transportadora de amonio. Esta arquitectura se observa únicamente en la proteína gi|48893689 de *T. erytraeum* IMS101.

Aunque la fusión de un dominio CYCc a un transportador ya se ha descrito en una GC de protozoos (Linder *et al.*, 1999), es la primera vez que se observa un transportador de amonio en la región reguladora de una AC/GC. El dominio transportador de amonio podría tener una función reguladora de la actividad CYCc ya que en *E. coli* la oligomerización del transportador controla su actividad (Ludewig *et al.*, 2003) y esta oligomerización está controlada por la unión, a AmtB, de la proteína PII, que es la que percibe el estado nitrogenado celular (Javelle *et al.*, 2004). Esto relacionaría el sistema de transmisión de señales con nucleótidos cíclicos con el metabolismo de nitrógeno.

Arquitectura T

La arquitectura *T* se compone de 5 hélices transmembranales situadas en el extremo N terminal y constituyen un dominio homólogo a las ATPasas transportadoras de cationes (PFAM 00689). Este dominio de membrana está unido a un dominio CYCc citoplásmico que alcanza hasta el extremo C-terminal. Al igual que en la GC de protozoos (Linder *et al.*, 1999), tal y como se ha sugerido para la arquitectura *S*, el dominio transmembranal podría regular la actividad ciclasa.

Entre las cianobacteras, esta arquitectura molecular solo se observa la proteína gi|23126862 de *N. punctiforme* PCC 73102.

Alineamiento de los homólogos de adenilil y guanilil ciclasa de Clase III

Las 55 secuencias de las posibles AC/GCs de clase III fueron alineadas y de ese alineamiento múltiple se seleccionaron los bloques de secuencias

conservadas en el alineamiento, tal y como se indica en materiales y métodos. La comparación de cada una de las secuencias con la secuencia de la AC CyaC de *S. platensis* (gi|7470942) cuyo centro activo ha sido cristalizado en presencia de análogos de ATP, calcio y bicarbonato (Steebhorn *et al.*, 2004) permite reconocer en las 54 secuencias restantes aquellos residuos que están probablemente implicados en la catálisis o transición de estado, y en el reconocimiento del sustrato ATP. Los residuos de CyaC implicados en las diferentes interacciones con los fosfatos, la ribosa y la purina del ATP pueden obtenerse de la estructura cristalina 1cw5 depositada en el banco de datos de proteínas (Figura 3). La identificación de las GCs se realizó comparando las secuencias alineadas con la GC Cya2 (gi|16329561) de *Synechocystis* PCC 6803 (Ochoa de Alda *et al.*, 2000).

Análisis de los residuos implicados en la catálisis

En CyaC, los residuos implicados en la transición de estado que permite la catálisis son la Asn1146 y la Arg1150, marcadas con un asterisco (*) en la figura 4 (Steebhorn *et al.*, 2004). De las 55 secuencias de AC/GCs estudiadas 52 conservan estos aminoácidos y 3, en lugar del par Asn/Arg tienen los residuos Lys/Phe (gi|46118802), Ala/Leu (gi|16330418) y Val/Arg (gi|37520048) (Figura 4). Eso significa probablemente que las 52 secuencias que contienen el par Asn/Arg son capaces de catalizar la reacción de ciclación de un nucleótido trifosfato. De las otras 3 no se puede concluir que no puedan ciclar un nucleótido trifosfato ya que, aunque no correspondan con ninguna de las variantes del par Asn/Arg descritas hasta el momento (Shenroy and Visweswariah, 2004), se ha mostrado recientemente que numerosos mutantes dirigidos hacia este par Asn/Arg conservan la actividad (Sinha *et al.*, 2005). Concretamente, en la AC Rv1900c de *Mycobacterium tuberculosis* el par Ala/Arg es activo (Sinha *et al.*, 2005). Por ello es probable que, al menos, el par Val/Arg de la proteína (gi|37520048) mantenga la capacidad de ciclación.

Para llevar a cabo la catálisis, las ACs y GCs deben interaccionar con un metal divalente (Magnesio o manganeso) a través de dos residuos de Asp (En CyaC, Asp1017 y Asp1061) u otro aminoácido ácido como el Glu (Shenroy and Visweswariah, 2004). Entre las 55 secuencias estudiadas solamente dos de las anteriormente mencionadas (gi|46118802, gi|16330418) carecen de estos residuos de Asp (Figura 4). Esto indica que si estas proteínas mantienen la actividad catalítica ésta debe llevarse a cabo según un mecanismo desconocido actualmente.

Análisis de los residuos implicados en la especificidad por el ATP o el GTP

En general, las ACs y GCs presentan una gran especificidad por el sustrato reconociendo el ATP y el GTP a través de las diferencias que muestran estas moléculas en el anillo púrico. Este reconocimiento se ejerce a través de 2-3 aminoácidos situados en el centro catalítico de la proteína. Así, en la AC CyaC, la Lys1057 establece un puente de hidrógeno con el N1 de la adenina. En algunas ACs este residuo lo ocupan los aminoácidos Arg, Gln, Asn, Asp y Glu (Katayama and Ohmori, 1997; Sinha *et al.*, 2005) no perdiendo por ello la especificidad por el ATP. Otros aminoácidos que determinan la especificidad en CyaC son la Ala1140 y la Thr1139 que establecen puentes de hidrógeno con el N6 de la adenina (Figura 3). En algunas ACs este último puente de hidrógeno se establece a través de un residuo de Asp (Liu *et al.*, 1997) o Ser. A diferencia de las ACs, en las GCs el lugar que ocupa la Lys en las ACs lo ocupa un residuo de Glu y el que ocupa la Thr o el Asp lo ocupa un residuo pequeño como la Gly (Ochoa de Alda *et al.*, 2000) o la Cys (Liu *et al.*, 1997). Estos residuos que determinan la especificidad por el sustrato se muestran recuadrados en la figura 4. Como se observa en el alineamiento, en el lugar correspondiente a los residuos que determinan la especificidad, la mayor parte de las secuencias estudiadas contienen el par Lys/Thr o Lys/Asp característico de la ACs. Algunas secuencias como las gi|23127822, gi|17228613 y gi|53765262, muestran el par Glu/Thr que ya ha sido descrito en una AC (Katayama and Ohmori, 1997).

Solamente 4 de las 55 secuencias de las estudiadas (gi|16329561, gi|48894764, gi|23130559 y gi|48895515) muestran los determinantes Glu/Gly característicos de GCs.

En dos secuencias (gi|46118802 y gi|16330418) la especificidad por el GTP o el ATP no puede ser inferida a partir del alineamiento de la figura 4.

Análisis de los residuos implicados en la activación por bicarbonato

Algunas ACs de los seres vivos pueden ser activadas (Cann *et al.*, 2003; Chen *et al.*, 2000) o inhibidas (Masuda and Ono, 2005) por bicarbonato. En la AC CyaC de *Spirulina*, los residuos implicados en la activación por bicarbonato parecen ser la Lys1057, Thr1139 y el residuo 1133 (Cann *et al.*, 2003; Steegborn *et al.*, 2004). A partir del alineamiento de las secuencias observamos

gi	7470942	1012	◆ ▲ ▲ ▲	◆ ◆ ◆ ◆	□	TILFSDIVGFTRMNSN (10) LNEYLGEMTRAVFENQGTVDKRFVGDAIMALY (8) VRRAIATA
gi	46119054	129	TILFTDLVGYTPLAS (10) LNQYLEEMTAVIFHYGGTVDKRFIGDAVVALFGAP (8) VYRAIAVA			
gi	48892878	879	AIFFSDIRSFTTISE (10) VNGYLKEVCPDIRDRNGLIIFLFGDGMVAVF --- (6) ---ALEAA			
gi	48895515	888	TILTSDLRGFTSLSE (10) LNLYLSSMADVIISFQGTIDEFMGDGILVLF (9) ---AIACA			
gi	17229758	600	TVLFSDIRGYTTLTE (10) LNQYFETMVEAVFVYEGTLDKRFIGDALMAVFGAP (7) AWQAVQSA			
gi	17229396	600	SILFSDIRGYTTLTE (10) LNEYFESMVEAVFKHKGTLDKYIGDAIMAVFGSP (6) AWMAVKTS			
gi	48894764	443	TILTSDLRGFTATSE (10) LNLYLGYMADVITEYQGTIDEFMGDGILVLF (6) PTRAVACA			
gi	16329561	442	TILTSDLRGFTSTSE (10) LNLYFGKMADVITHHGGTIDEFMGDGILVLF (6) ALRAVACG			
gi	23127822	441	TVLVSDLRGFTSAISE (10) LNLYLGTMTDVIINQYKGTINEFMGDGIFVIF (6) SQRAIACA			
gi	23130559	310	TILTSDLRGFTALSE (10) LNLYLEYMADVINHYYQGTIDEFMGDGILVLF (6) ALRAVTCA			
gi	48893394	393	TILFTDLKNFSTISE (10) LNQYLSAMTDIVINHNGIVNKFTGDMVAVFGIP (12) AQNAVNCA			
gi	23127996	360	SILFADLVGFTALSE (10) LNNLFEEMLKEVFTCGGTLDKYIGDCIMAFFGAP (6) ADRAVTAA			
gi	16330472	131	TVLVADMRNFTGMAQ (10) IGWFRQAAGHILREAGSWVDKYIGDAVMAIWFHG (8) IIQILHAV			
gi	48893689	1031	TVLFADLVGFTQLAS (10) LNEIFSKFDLLAEKYGLEKIKIKITIGDAYMVVGG (6) AEAIALIA			
gi	23126862	255	TVLVADLVGFTQLAS (10) LNTIFCLFDQLAELHKKVEKIKIKITIGDAYMAVAGLP (6) ALAIANMA			
gi	48893696	276	TVLVADLVGFTKQVSA (10) LNMIFSCFDQLADKYGLEKIKIKITIGDAYMVVGG (6) AEAIAEMA			
gi	11990887	487	TILFADLVGFTPLSA (10) LNEMFSIFDNLADKHKLEKIKIKITIGDAYMVVGG (6) AEAIADMA			
gi	17228157	292	TVMFADLVGFTDL (10) LNQIFSRFDLLAESYNLEKIKIKITIGDAYMVVGG (6) AIAIAHMG			
gi	48891559	1032	TILFADLVGFTSFA (10) LNQVSEFDKITEOHNLEKIKIKITIGDSYMASGLP (6) AEAIAKMA			
gi	46118802	133	TVLSANIVGFDEECE (10) LNDIITAFDDAAQRHDVEKISTVGTSYLAASGLF (6) DKAVIELA			
gi	33864410	458	SVLFADLVGFTALTR (10) LNHYLDEMRSVVWDHGGFLDKRFIGDAVMAVFG (8) AQSAICCA			
gi	16330418	121	ALLMANIRGATNFDE (10) MGRYFSQMQKIEKYGGTISNVMGRKMLVAFGLK (6) AERATWAA			
gi	37520048	109	FVMSIDIRRSSTELML (11) ISSLCKDLKEVITKNYGIKFKFTGDKGILAFFPEF (7) -IHAVQSA			
gi	23126594	1	----- (9) LNSIFSTFDRLTEQHGLKRIKIKITINDAYMVVGG (6) TQAIAL--			

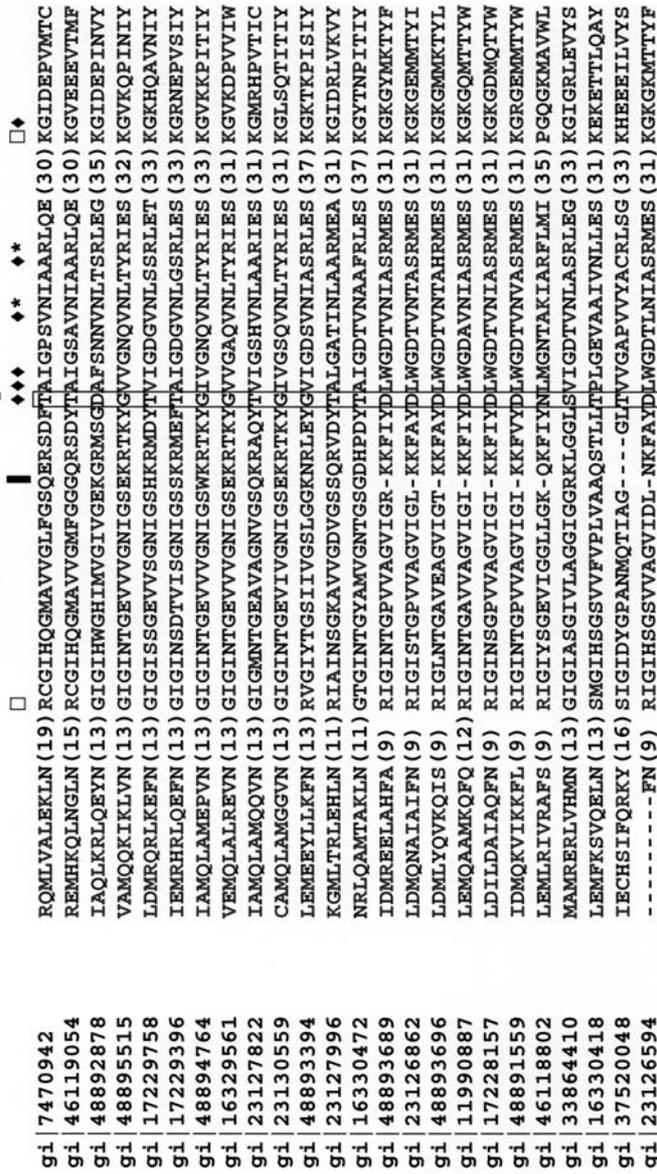


Figura 4: Alineamiento múltiple de las secuencias primarias del centro catalítico de las adenilil y guanilil ciclasas más representativas de cianobacterias. Sobre la secuencia de la proteína cristalizada CyaC de *Spirulina platensis* (gi|7470942) se indican los residuos de esta proteína (Fig.3) que interaccionan con el fosfato (□), la ribosa (◆) o la adenina (▲) del ATP, con el metal (▲), o que son necesarios para la activación por bicarbonato (■) o la transición de estado (*) durante la catálisis. Los recuadros indican los principales residuos implicados en la especificad por el sustrato. Las secuencias se nombran con el identificador del gen (columna izquierda). El número junto a esta secuencia indica la posición del primer residuo. Los números entre paréntesis representan al número de residuos de cada secuencia que han sido eliminados en esa posición debido a que no están conservados en la mayor parte de las secuencias.

36 de las ACs de cianobacterias contienen estos residuos y, por lo tanto, podrían ser activadas por bicarbonato

Filogenia de los homólogos de adenilil y guanilil ciclase de Clase III

Las ACs/GCs de las cianobacterias pueden provenir del ancestro común de todas ellas y/o haber surgido a lo largo de la evolución de las cianobacterias a través de duplicaciones y/o haber sido obtenidas mediante transferencia genética horizontal. Para estudiar la importancia de cada una de estas posibilidades realizamos un análisis filogenético de las secuencias correspondientes al centro catalítico de las ACs/GCs. Para eliminar los artefactos debidos a regiones hipervariables éstas regiones fueron suprimidas, como se indica en materiales y métodos, resultando bloques conservados de 115 residuos. Si bien esto disminuye la información filogenético de las secuencias completas de los centros catalíticos, aumenta la estabilidad topológica de los árboles obtenidos. Como se observa en la figura 5 algunas secuencias se agrupan con una alta probabilidad independientemente del procedimiento empleado para la construcción del árbol filogenético, máxima verosimilitud, distancia de proteínas y máxima parsimonia. Muchos de estos grupos corresponden a secuencias con la misma arquitectura molecular. Así, las arquitecturas *B*, *C*, *D* y *F* tienen un origen común. La topología de las ramas que contienen las arquitecturas *C*, *D* y *F* así como cada una de las isoformas de la arquitectura *B*, (*B1* y *B2*) siguen la misma topología que el árbol filogenético de la subunidad 16S (Figura 1) de cada una de las especies estudiadas. Esto indica que cada una de estas arquitecturas constituye un grupo de ortólogos y, por consiguiente es probable que los elementos de cada grupo tengan la misma función. Curiosamente estas arquitecturas no se observan en otros *Phyla* de bacterias lo que sugiere que aparecieron y evolucionaron durante el proceso de especiación de las cianobacterias.

En algunos casos, como en las secuencias con arquitectura *A*, no se observa un origen común. Esto, más que deberse a un origen polifilético de la arquitectura *A* podría ser consecuencia de la escasa información filogenética de las secuencias. Un estudio más detallado de la evolución de la arquitectura *A* en las bacterias indicó que todas las de cianobacterias tenían un origen común (A. Pedraza, sin publicar) y formaban un grupo monofilético. Esto es, el ancestro de todas las cianobacterias estudiadas tenía una AC de arquitectura

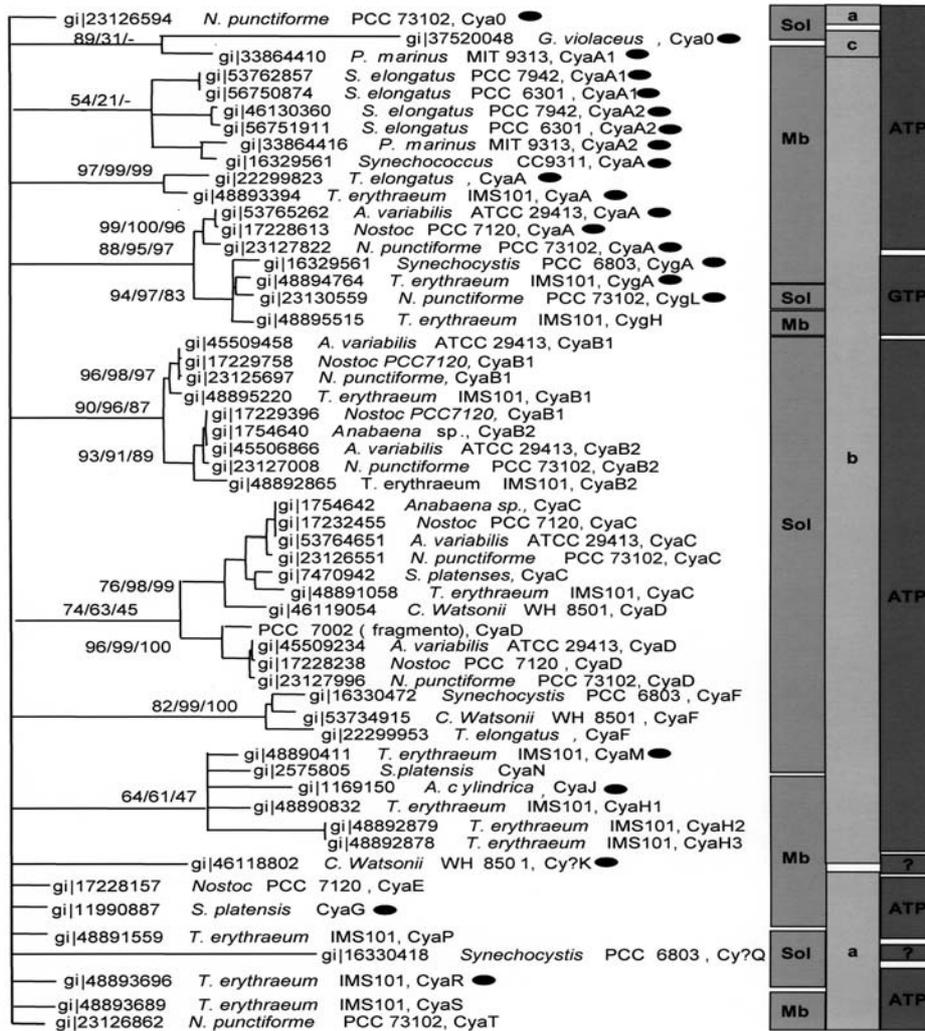


Figura 5: Árbol filogenético de Máxima verosimilitud del dominio catalítico de las adenilil y guanilil ciclasas de las cianobacterias. El árbol se realizó empleando el alineamiento de la figura 4. Para el análisis se emplearon también los métodos de Máxima verosimilitud/Distancia de proteínas/Máxima parsimonia. El porcentaje de confianza obtenido por cada uno de estos métodos se muestra junto a cada nudo. Se muestra el árbol de Máxima verosimilitud. Junto al número de identificación de cada uno de los genes (gi) se indica el organismo al que pertenece y el nombre que proponemos en este estudio. Los puntos negros junto a los nombres señalan aquellas arquitecturas que se observan en otras *Phyla* distintos al de *Cyanobacteria*. La barra vertical izquierda identifica aquellas proteínas solubles (sol) o de membrana (mb); la barra del medio identifica la subclase (a, b, o d) a la que pertenece la AC; la barra vertical de la derecha identifica la especificidad por el sustrato, ATP o GTP, de cada una de las ciclasas.

A. Esta arquitectura se encuentra en numerosos grupos de bacterias, principalmente en los *Phyla Proteobacteria* y *Spirochaetes*. Esto indica que esta arquitectura se encontraba probablemente en el ancestro común de estos *Phyla* y el *Phylum Cyanobacteria*. La presencia de isoformas de esta arquitectura *A* en algunas cianobacterias sugiere que son el resultado de duplicaciones genéticas durante la diversificación de las cianobacterias. Por otro lado la irregular distribución de las diferentes ACs y GCs en las especies de cianobacterias estudiadas indica que las deleciones de los genes que codifican estas proteínas también son frecuentes.

Los cuatro representantes que muestran especificidad por el GTP (se trata por tanto de GCs) formar un grupo dentro de las ACs de arquitectura *A* lo que sugiere que proviene de ACs en las que se ha modificado la especificidad por el sustrato. De las cuatro GCs observadas dos muestran una arquitectura *A* (gi|16329561 y gi|48894764), lo que confirma su origen común al de las ACs de esta arquitectura, y dos tienen arquitecturas diferentes, gi|23130559 y gi|48895515, lo que sugiere que son el resultado de fusiones de GCs de arquitectura *A* con otros módulos de señalización que han dado lugar a las arquitectura *L* y *H* respectivamente. Dado que la arquitectura *L* también se observa en proteobacterias y que otras ACs con arquitectura *H* (H1, H2 y H3) se ubican en otro grupo independiente del árbol, estas fusiones podrían ser recurrentes a lo largo de la evolución de las ACs/GCs y no necesariamente tener un origen común. Las fusiones de dominios que dan lugar a la enorme variedad de arquitecturas moleculares parecen ser frecuentes como muestra el agrupamiento significativo en el árbol filogenético de ACs con arquitecturas diferentes como las *M*, *N*, *J* y *H*.

Finalmente, un numeroso grupo de arquitecturas observadas en otros *Phyla* (*R*, *G*, *K*, *O*) no forman un grupo consistente en el árbol filogenético lo que sugiere que pudieran haber sido heredadas del ancestro común de todas las bacterias. Esto significaría que el ancestro común de algunos *Phyla* bacterianos contenía numerosas ACs, igual que las cianobacterias, y que podría tener unos sistemas de transmisión de señal muy sofisticados. Esto se interpreta actualmente como un indicador de que el ancestro común de estas bacterias vivía en un medio ambiente inestable (Galperin, 2005).

Clasificación de los homólogos de adenilil y guanilil ciclasa de Clase III

Las ACs y GCs se han clasificado clásicamente en dos tipos: solubles y de membrana. Esa clasificación es insuficiente a juzgar por la enorme variedad de arquitectura moleculares que se observan en ACs y GCs solubles y de membrana. En consecuencia, recientemente, se ha propuesto una clasificación en cuatro subclases (a, b, c y d) según las características del centro activo (Linder and Schultz, 2003). Como se observa en la figura 5, estas dos clasificaciones no reflejan la diversidad de ACs y GCs encontradas en las cianobacterias. Por ello, nosotros proponemos una nomenclatura en la que primero se indique la posible especificidad por el sustrato. Así, las denominamos Cya si son específicas por el ATP, Cyg si tiene especificidad por el GTP y Cy si no se puede inferir su especificidad. A este nombre le sigue la letra que designa la arquitectura molecular (CyaA, CyaB, CygH, etc) y un número en el caso de que haya isozimas (CyaB1, CyaB2, etc). Esta nomenclatura no modifica gran parte de la nomenclatura empleada actualmente, amplía y recoge la diversidad de arquitecturas moleculares de estos enzimas y se ajusta a las directrices de la IUPAC (IUPAC-IUB, 1977, 1989).

Conclusiones

El estudio de la arquitectura molecular de las AC/GCs de cianobacterias permite mostrar la gran variabilidad de estos enzimas que existe en este grupo de bacterias. Es significativo que no haya ningún módulo de transmisión común a todas las cianobacterias estudiadas, es decir, ninguna AC o GC se observa en todas las cianobacterias. Esto implica que, muy probablemente, las ACs/GCs de cianobacterias, y por consiguiente los niveles intracelulares de cAMP/cGMP, responden a diferentes estímulos en cada una de ellas. El 45% de las arquitecturas moleculares no contenían dominios transmembranales lo que indica que las ACs/GCs de cianobacterias pueden ser tanto solubles como de membrana y en consecuencia este sistema de señalización es empleado por las cianobacterias tanto para transmitir señales intracelulares como señales externas al interior celular. Las arquitecturas moleculares encontradas indican que en algunas cianobacterias los sistemas de señalización dependientes de nucleótidos cíclicos se entrecruzan con los sistemas de señalización de dos

componentes, con los sistemas de señalización de serín-treonín kinasas y el sistema de señalización que controla el metabolismo del nitrógeno. La cianobacteria estudiada con el sistema de señalización más complejo es *T. erythraeum* IMS101.

La diversidad de ACs en las cianobacterias parece ser el resultado de i) la herencia de un ancestro común (para aquellas observadas en otros *Phyla*), ii) duplicaciones, iii) fusiones y deleciones de genes. Nuestros resultados indican que las GCs de cianobacterias provienen de las ACs a través de cambios en los residuos que determinan la especificidad por el ATP. Diferentes fusiones posteriores dieron lugar a la diversidad de GCs observada en este grupo de bacterias.

Las cianobacterias son por el momento el único grupo procariota en el que se ha descrito la presencia de ACs y GCs en la misma célula. Actualmente, el organismo biológico modelo más simple en el que se estudian simultáneamente los sistemas de transmisión de señales dependiente de cAMP y cGMP es el moho deslizante *Dictyostelium discoideum* (Roelofs *et al.*, 2003), un organismo fundamental en el estudio de la toma de decisiones a través de los sistemas de transmisión de señales en células eucariotas. La abundancia de ACs y GCs que hemos observado nosotros en las cianobacterias sugieren que son el organismo modelo más simple en el que se puede estudiar la complejidad del sistema que constituyen las rutas de señalización dependientes de nucleótidos cíclicos en los seres vivos.

BIBLIOGRAFÍA

- ALTSCHUL, S.F., MADDEN, T.L., SCHAFER, A.A., ZHANG, J., ZHANG, Z., MILLER, W., AND LIPMAN, D.J. (1997) GAPPED BLAST AND PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402.
- ANANTHARAMAN, V., AND ARAVIND, L. (2000): Cache - a signaling domain common to animal Ca(2+)-channel subunits and a class of prokaryotic chemotaxis receptors. *Trends Biochem Sci* 25: 535-537.
- ARAI, M., MITSUKE, H., IKEDA, M., XIA, J.X., KIKUCHI, T., SATAKE, M., AND SHIMIZU, T. (2004): ConPred II: a consensus prediction method for obtaining transmembrane topology models with high reliability. *Nucleic Acids Res* 32: W390-393.
- BAKER, D.A., AND KELLY, J.M. (2004): Structure, function and evolution of microbial adenylyl and guanylyl cyclases. *Mol Microbiol* 52: 1229-1242.
- BARNSTABLE, C.J., WEI, J.Y., AND HAN, M.H. (2004): Modulation of synaptic function by cGMP and cGMP-gated cation channels. *Neurochem Int* 45: 875-884.
- BOSGRAAF, L., RUSSCHER, H., SNIPPE, H., BADER, S., WIND, J., AND VAN HAASTERT, P.J.M. (2002): Identification and characterization of two unusual cGMP-stimulated phosphodiesterases in Dictyostelium. *Molecular Biology Of The Cell* 13: 3878-3889.
- BRUDER, S., LINDER, J.U., MARTINEZ, S.E., ZHENG, N., BEAVO, J.A., AND SCHULTZ, J.E. (2005): The cyanobacterial tandem GAF domains from the cyaB2 adenylyl cyclase signal via both cAMP-binding sites. *Proc Natl Acad Sci U S A*.
- CANN, M.J., HAMMER, A., ZHOU, J., AND KANACHER, T. (2003): A defined subset of adenylyl cyclases is regulated by bicarbonate ion. *J Biol Chem* 278: 35033-35038.
- CANN, M.J. (2004): Signalling through cyclic nucleotide monophosphates in cyanobacteria. *New Phytologist* 161: 23-34.
- CASTRESANA, J. (2000): Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17: 540-552.
- COLE, J.R., CHAI, B., FARRIS, R.J., WANG, Q., KULAM, S.A., MCGARRELL, D.M., GARRITY, G.M., AND TIEDJE, J.M. (2005): The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. *Nucleic Acids Res* 33 Database Issue: D294-296.
- COTTA, M.A., WHITEHEAD, T.R., AND WHEELER, M.B. (1998): Identification of a novel adenylyl cyclase in the ruminal anaerobe, *Prevotella ruminicola* D31d. *FEMS Microbiol Lett* 164: 257-260.
- COUDART-CAVALLI, M.P., SISMEIRO, O., AND DANCHIN, A. (1997): Bifunctional structure of two adenylyl cyclases from the myxobacterium *Stigmatella aurantiaca*. *Biochimie* 79: 757-767.

- CHEN, Y.Q., CANN, M.J., LITVIN, T.N., IOURGENKO, V., SINCLAIR, M.L., LEVIN, L.R., AND BUCK, J. (2000): Soluble adenylyl cyclase as an evolutionarily conserved bicarbonate sensor. *Science* 289: 625-628.
- DANCHIN, A. (1993): Phylogeny of adenylyl cyclases. *Adv Second Messenger Phosphoprotein Res* 27: 109-162.
- DRUM, C.L., YAN, S.Z., BARD, J., SHEN, Y.Q., LU, D., SOELAIMAN, S., GRABAREK, Z., BOHM, A., AND TANG, W.J. (2002): Structural basis for the activation of anthrax adenylyl cyclase exotoxin by calmodulin. *Nature* 415: 396-402.
- ESSAYAN, D.M. (2001): Cyclic nucleotide phosphodiesterases. *J Allergy Clin Immunol* 108: 671-680.
- FELSENSTEIN, J. (1996): Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. *Methods Enzymol* 266: 418-427.
- GALPERIN, M.Y. (2004): Bacterial signal transduction network in a genomic perspective. *Environ Microbiol* 6: 552-567.
- GALPERIN, M.Y. (2005): A census of membrane-bound and intracellular signal transduction proteins in bacteria: Bacterial IQ, extroverts and introverts. *BMC Microbiol* 5: 35.
- HAMMET, A., PIKE, B.L., MCNEES, C.J., CONLAN, L.A., TENIS, N., AND HEIERHORST, J. (2003): FHA domains as phospho-threonine binding modules in cell signaling. *IUBMB Life* 55: 23-27.
- HANNENHALLI, S.S., AND RUSSELL, R.B. (2000): Analysis and prediction of functional sub-types from protein sequence alignments. *J Mol Biol* 303: 61-76.
- HASEGAWA, M., IIDA, Y., YANO, T., TAKAIWA, F., AND IWABUCHI, M. (1985): Phylogenetic relationships among eukaryotic kingdoms inferred from ribosomal RNA sequences. *J Mol Evol* 22: 32-38.
- HOLLAND, M.M., LEIB, T.K., AND GERLT, J.A. (1988): Isolation and characterization of a small catalytic domain released from the adenylate cyclase from *Escherichia coli* by digestion with trypsin. *J Biol Chem* 263: 14661-14668.
- IUPAC-IUB (1977): Nomenclature of multiple forms of enzymes. Recommendations (1976): IUPAC-IUB Commission on Biochemical Nomenclature (CBN). *J Biol Chem* 252: 5939-5941.
- IUPAC-IUB (1989): Nomenclature Committee Of The International Union Of Biochemistry (Nc-Iub) - Nomenclature For Multienzymes - Recommendations 1989. *European Journal Of Biochemistry* 185: 485-486.
- IYER, L.M., AND ARAVIND, L. (2002): The catalytic domains of thiamine triphosphatase and CyaB-like adenylyl cyclase define a novel superfamily of domains that bind organic phosphates. *BMC Genomics* 3: 33. Print 2002 Nov 2027.
- JAVELLE, A., SEVERI, E., THORNTON, J., AND MERRICK, M. (2004): Ammonium sensing in *Escherichia coli*. Role of the ammonium transporter AmtB and AmtB-

- GlnK complex formation. *J Biol Chem* 279: 8530-8538.
- JONES, D.T., TAYLOR, W.R., AND THORNTON, J.M. (1992): The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8: 275-282.
- KANACHER, T., SCHULTZ, A., LINDER, J.U., AND SCHULTZ, J.E. (2002): A GAF-domain-regulated adenylyl cyclase from *Anabaena* is a self-activating cAMP switch. *Embo Journal* 21: 3672-3680.
- KANACHER, T., SCHULTZ, A., LINDER, J.U., AND SCHULTZ, J.E. (2002): A GAF-domain-regulated adenylyl cyclase from *Anabaena* is a self-activating cAMP switch. *Embo J* 21: 3672-3680.
- KASAHARA, M., YASHIRO, K., SAKAMOTO, T., AND OHMORI, M. (1997): The *Spirulina platensis* adenylate cyclase gene, *cyaC*, encodes a novel signal transduction protein. *Plant Cell Physiol* 38: 828-836.
- KASAHARA, M., AND OHMORI, M. (1999): Activation of a cyanobacterial adenylate cyclase, *CyaC*, by autophosphorylation and a subsequent phosphotransfer reaction. *J Biol Chem* 274: 15167-15172.
- KASAHARA, M., AND OHMORI, M. (1999): Activation of a cyanobacterial adenylate cyclase, *CyaC*, by autophosphorylation and a subsequent phosphotransfer reaction. *J Biol Chem* 274: 15167-15172.
- KASAHARA, M., UNNO, T., YASHIRO, K., AND OHMORI, M. (2001): *CyaG*, a novel cyanobacterial adenylyl cyclase and a possible ancestor of mammalian guanylyl cyclases. *J Biol Chem* 276: 10564-10569.
- KASAHARA, M., UNNO, T., YASHIRO, K., AND OHMORI, M. (2001): *CyaG*, a novel cyanobacterial adenylyl cyclase and a possible ancestor of mammalian guanylyl cyclases. *J Biol Chem* 276: 10564-10569.
- KATAYAMA, M., WADA, Y., AND OHMORI, M. (1995): Molecular-Cloning Of The Cyanobacterial Adenylate-Cyclase Gene From The Filamentous Cyanobacterium *Anabaena-Cylindrica*. *J Bacteriol* 177: 3873-3878.
- KATAYAMA, M., AND OHMORI, M. (1997): Isolation and characterization of multiple adenylate cyclase genes from the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 179: 3588-3593.
- KIMURA, Y., OHTANI, M., AND TAKEGAWA, K. (2005): An Adenylyl Cyclase, *CyaB*, Acts as an Osmosensor in *Myxococcus xanthus*. *J Bacteriol* 187: 3593-3598.
- KITANO, H. (2002): Systems biology: a brief overview. *Science* 295: 1662-1664.
- KOSHLAND, D.E., JR. (2002) Special essay. The seven pillars of life. *Science* 295: 2215-2216.
- LETUNIC, I., COPLEY, R.R., SCHMIDT, S., CICCARELLI, F.D., DOERKS, T., SCHULTZ, J., PONTING, C.P., AND BORK, P. (2004): SMART 4.0: towards genomic data integration. *Nucleic Acids Res* 32: D142-144.
- LINDER, J.U., ENGEL, P., REIMER, A., KRUGER, T., PLATTNER, H., SCHULTZ, A., AND

- SCHULTZ, J.E. (1999): Guanylyl cyclases with the topology of mammalian adenylyl cyclases and an N-terminal P-type ATPase-like domain in Paramecium, Tetrahymena and Plasmodium. *Embo J* 18: 4222-4232.
- LINDER, J.U., HOFFMANN, T., KURZ, U., AND SCHULTZ, J.E. (2000): A guanylyl cyclase from Paramecium with 22 transmembrane spans. Expression of the catalytic domains and formation of chimeras with the catalytic domains of mammalian adenylyl cyclases. *J Biol Chem* 275: 11235-11240.
- LINDER, J.U., AND SCHULTZ, J.E. (2003): The class III adenylyl cyclases: multi-purpose signalling modules. *Cell Signal* 15: 1081-1089.
- LINDER, J.U., HAMMER, A., AND SCHULTZ, J.E. (2004): The effect of HAMP domains on class IIIb adenylyl cyclases from Mycobacterium tuberculosis. *Eur J Biochem* 271: 2446-2451.
- LIU, Y., RUOHO, A.E., RAO, V.D., AND HURLEY, J.H. (1997): Catalytic mechanism of the adenylyl and guanylyl cyclases: Modeling and mutational analysis. *Proc Natl Acad Sci U S A*. 94: 13414-13419.
- LUDEWIG, U., WILKEN, S., WU, B., JOST, W., OBRDLIK, P., EL BAKKOURY, M., MARINI, A.M., ANDRE, B., HAMACHER, T., BOLES, E., VON WIREN, N., AND FROMMER, W.B. (2003): Homo- and hetero-oligomerization of ammonium transporter-1 NH4 uniporters. *J Biol Chem* 278: 45603-45610.
- LUDIDI, N., AND GEHRING, C. (2003): Identification of a novel protein with guanylyl cyclase activity in Arabidopsis thaliana. *J Biol Chem* 278: 6490-6494. Epub 2002 Dec 6412.
- MA, Q., JOHNSON, M.S., AND TAYLOR, B.L. (2005): Genetic analysis of the HAMP domain of the Aer aerotaxis sensor localizes flavin adenine dinucleotide-binding determinants to the AS-2 helix. *J Bacteriol* 187: 193-201.
- MASUDA, S., AND ONO, T. (2004): Biochemical characterization of the major adenylyl cyclase, Cya1, in the cyanobacterium Synechocystis sp PCC 6803. *Febs Letters* 577: 255-258.
- MASUDA, S., AND ONO, T.A. (2004): Biochemical characterization of the major adenylyl cyclase, Cya1, in the cyanobacterium Synechocystis sp. PCC 6803. *FEBS Lett* 577: 255-258.
- MASUDA, S., AND ONO, T.A. (2005): Adenylyl Cyclase Activity of Cya1 from the Cyanobacterium Synechocystis sp. Strain PCC 6803 Is Inhibited by Bicarbonate. *J Bacteriol* 187: 5032-5035.
- MOLLE, V., SOULAT, D., JAULT, J.M., GRANGEASSE, C., COZZONE, A.J., AND PROST, J.F. (2004): Two FHA domains on an ABC transporter, Rv1747, mediate its phosphorylation by PknF, a Ser/Thr protein kinase from Mycobacterium tuberculosis. *FEMS Microbiol Lett* 234: 215-223.
- MOUTINHO, A., HUSSEY, P.J., TREWAVAS, A.J., AND MALHO, R. (2001): cAMP acts as a

- second messenger in pollen tube growth and reorientation. *Proc Natl Acad Sci U S A* 98: 10481-10486. Epub 12001 Aug 10421.
- NOTREDAME, C., HIGGINS, D.G., AND HERINGA, J. (2000): T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302: 205-217.
- NOUBIR, S., LUQUE, I., OCHOA DE ALDA, J.A.G., PEREWOSKA, I., DE MARSAC, N.T., COBLEY, J.G., AND HOUMARD, J. (2002): Co-ordinated expression of phycobiliprotein operons in the chromatically adapting cyanobacterium *Calothrix* PCC 7601: a role for RcaD and RcaG. *Molecular Microbiology* 43: 749-762.
- OCHOA DE ALDA, J.A., AJLANI, G., AND HOUMARD, J. (2000): *Synechocystis* strain PCC 6803 *cya2*, a prokaryotic gene that encodes a guanylyl cyclase. *J Bacteriol* 182: 3839-3842.
- OCHOA DE ALDA, J.A., AND HOUMARD, J. (2000): Genomic survey of cAMP and cGMP signalling components in the cyanobacterium *Synechocystis* PCC 6803. *Microbiology* 146: 3183-3194.
- OHMORI, M., IKEUCHI, M., SATO, N., WOLK, P., KANEKO, T., OGAWA, T., KANEHISA, M., GOTO, S., KAWASHIMA, S., OKAMOTO, S., YOSHIMURA, H., KATOH, H., FUJISAWA, T., EHIRA, S., KAMEI, A., YOSHIHARA, S., NARIKAWA, R., AND TABAT, S. (2001): Characterization of genes encoding multi-domain proteins in the genome of the filamentous nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res* 8: 271-284.
- OKAMOTO, S., KASAHARA, M., KAMIYA, A., NAKAHIRA, Y., AND OHMORI, M. (2004): A Phytochrome-like Protein AphC Triggers the cAMP Signaling Induced by Far-red Light in the Cyanobacterium *Anabaena* sp. Strain PCC7120 paragraph sign. *Photochem Photobiol* 80: 429-433.
- ROELOFS, J., SMITH, J.L., AND VAN HAASTERT, P.J. (2003): cGMP signalling: different ways to create a pathway. *Trends Genet* 19: 132-134.
- SCHMIDT, H.A., STRIMMER, K., VINGRON, M., AND VON HAESLER, A. (2002): TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* 18: 502-504.
- SHENROY, A.R., AND VISWESWARIAH, S.S. (2004): Class III nucleotide cyclases in bacteria and archaeobacteria: lineage-specific expansion of adenylyl cyclases and a dearth of guanylyl cyclases. *FEBS Lett* 561: 11-21.
- SINHA, S.C., WETTERER, M., SPRANG, S.R., SCHULTZ, J.E., AND LINDER, J.U. (2005): Origin of asymmetry in adenylyl cyclases: structures of *Mycobacterium tuberculosis* Rv1900c. *Embo J.*
- SISMEIRO, O., TROTOT, P., BIVILLE, F., VIVARES, C., AND DANCHIN, A. (1998): *Aeromonas hydrophila* adenylyl cyclase 2: a new class of adenylyl cyclases with thermophilic properties and sequence similarities to proteins from hyperthermophilic archaeobacteria. *J Bacteriol* 180: 3339-3344.

- STEEGBORN, C., LITVIN, T.N., LEVIN, L.R., BUCK, J., AND WU, H. (2004): Bicarbonate activation of adenylyl cyclase via promotion of catalytic active site closure and metal recruitment. *Nat Struct Mol Biol*.
- STOCK, A.M., ROBINSON, V.L., AND GOUDREAU, P.N. (2000): Two-component signal transduction. *Annu Rev Biochem* 69: 183-215.
- Sunahara, R.K., and Taussig, R. (2002): Isoforms of mammalian adenylyl cyclase: multiplicities of signaling. *Mol Interv* 2: 168-184.
- TADA, T., SEKIMOTO, H., AND OHMORI, M. (2001): Biochemical characterization of an adenylyl cyclase, CyaB1, in the cyanobacterium *Anabaena* sp strain PCC 7120. *Journal Of Plant Research* 114: 387-394.
- TANDEAU DE MARSAC, N.T., AND HOUMARD, J. (1993): Adaptation of Cyanobacteria to Environmental Stimuli - New Steps Towards Molecular Mechanisms. *FEMS Microbiol Rev* 104: 119-189.
- TELLEZ-SOSA, J., SOBERON, N., VEGA-SEGURA, A., TORRES-MARQUEZ, M.E., AND CEVALLOS, M.A. (2002): The *Rhizobium etli* cyaC product: characterization of a novel adenylyl cyclase class. *J Bacteriol* 184: 3560-3568.
- TERAUCHI, K., AND OHMORI, M. (1999): An adenylyl cyclase, Cya1, regulates cell motility in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol* 40: 248-251.
- TERAUCHI, K., AND OHMORI, M. (2004): Blue light stimulates cyanobacterial motility via a cAMP signal transduction system. *Molecular Microbiology* 52: 303-309.
- THOMPSON, J.D., HIGGINS, D.G., AND GIBSON, T.J. (1994): CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680.
- TUCKER, C.L., HURLEY, J.H., MILLER, T.R., AND HURLEY, J.B. (1998): Two amino acid substitutions convert a guanylyl cyclase, RetGC-1, into an adenylyl cyclase. *Proc Natl Acad Sci U S A*. 95: 5993-5997.
- VREEDE, J., VAN DER HORST, M.A., HELLINGWERF, K.J., CRIELAARD, W., AND VAN AALTEN, D.M. (2003): PAS domains. Common structure and common flexibility. *J Biol Chem* 278: 18434-18439.
- WHEELER, D.L., BARRETT, T., BENSON, D.A., BRYANT, S.H., CANESE, K., CHURCH, D.M., DI CUCCIO, M., EDGAR, R., FEDERHEN, S., HELMBERG, W., KENTON, D.L., KHOVAYKO, O., LIPMAN, D.J., MADDEN, T.L., MAGLOTT, D.R., OSTELL, J., PONTIUS, J.U., PRUITT, K.D., SCHULER, G.D., SCHRIML, L.M., SEQUEIRA, E., SHERRY, S.T., SIROTKIN, K., STARCHENKO, G., SUZEK, T.O., TATUSOV, R., TATUSOVA, T.A., WAGNER, L., AND YASCHENKO, E. (2005): Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 33 Database Issue: D39-45.

- YAFFE, M.B., AND SMERDON, S.J. (2004): The use of in vitro peptide-library screens in the analysis of phosphoserine/threonine-binding domain structure and function. *Annu Rev Biophys Biomol Struct* 33: 225-244.
- YHR, T.L., VALLIS, A.J., HANCOCK, M.K., BARBIERI, J.T., AND FRANK, D.W. (1998): ExoY, an adenylate cyclase secreted by the *Pseudomonas aeruginosa* type III system. *Proc Natl Acad Sci U S A* 95: 13899-13904.
- YASHIRO, K., SAKAMOTO, T., AND OHMORI, M. (1996): Molecular characterization of an adenylate cyclase gene of the cyanobacterium *Spirulina platensis*. *Plant Molecular Biology* 31: 175-181.
- ZHANG, G.Y., LIU, Y., RUOHO, A.E., AND HURLEY, J.H. (1997): Structure of the adenylate cyclase catalytic core. *Nature* 386: 247-253.
- ZHENG, L., KOSTREWA, D., BERNECHE, S., WINKLER, F.K., AND LI, X.D. (2004): The mechanism of ammonia transport based on the crystal structure of AmtB of *Escherichia coli*. *Proc Natl Acad Sci USA* 101: 17090-17095.
- ZHULIN, I.B., NIKOLSKAYA, A.N., AND GALPERIN, M.Y. (2003): Common extracellular sensory domains in transmembrane receptors for diverse signal transduction pathways in bacteria and archaea. *J Bacteriol* 185: 285-294.