The structures of the PII proteins from the cyanobacteria *Synechococcus* sp. PCC 7942 and *Synechocystis* sp. PCC 6803

Yibin Xu,[a]† Paul D. Carr,[b]† Paula Clancy,[a] Mario Garcia-Dominguez,[b,c] Karl Forchhammer,[d] Francisco Florencio,[c] Nicole Tandeau de Marsac,[e] Subhash G. Vasudevan[a,f] and David L. Ollis[b]*

[a]Department of Biochemistry and Molecular Biology, James Cook University, Townsville, Queensland 4811, Australia, [b]Research School of Chemistry, Australian National University, GPO Box 414, Canberra, ACT 2601, Australia, [c]Instituto de Bioquímica Vegetal y Fotosíntesis, Universidad de Sevilla-Consejo Superior de Investigaciones Científicas, Américo Vespucio s/n, E-41092 Sevilla, Spain, [d]Institut für Mikrobiologie und Molekularbiologie, Justus-Liebig-Universität Giessen, Germany, [e]Unité des Cyanobactéries, CNRS-URA 2172, Département de Microbiologie Fondamentale et Médicale, Institut Pasteur, 28 Rue du Docteur Roux, 75724 Paris CEDEX 15, France, and [f]Novaris Institute for Tropical Diseases Pte Ltd, 1 Science Park Road, #04-14 The Capricorn, Singapore Science Park II, Singapore 117528, Singapore

† These two authors contributed equally to this paper.

Correspondence e-mail: ollis@rsc.anu.edu.au

The PII proteins from the cyanobacteria *Synechococcus* sp. PCC 7942 and *Synechocystis* sp. PCC 6803 have been crystallized and high-resolution structures have been obtained using X-ray crystallography. The core of these new structures is similar to that of the PII proteins from *Escherichia coli*, although the structures of the T- and C-loops differ. The T-loop of the *Synechococcus* protein is ordered, but appears to be stabilized by crystal contacts. The same loop in the *Synechocystis* protein is disordered. The C-terminus of the *Synechocystis* protein is stabilized by hydrogen bonding to the same region of a crystallographically related molecule. The same terminus in the *Synechococcus* protein is stabilized by coordination with a metal ion. These observations are consistent with the idea that both the T-loop and the C-terminus of PII proteins are flexible in solution and that this flexibility may be important for receptor recognition. Sequence comparisons are used to identify regions of the sequence unique to the cyanobacteria.
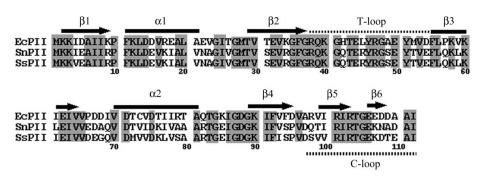
1. Introduction

PII is a signal-transduction protein that is best known for its role in the regulation of nitrogen assimilation by enteric bacteria. However, the protein is found and highly conserved in a variety of organisms covering all the three kingdoms of life (Ninfa & Atkinson, 2000; Arcondeguy *et al.*, 2001). In several of these organisms there are genes that code for two or even three PII-like proteins. In *Escherichia coli* there are two such proteins, known as GlnB and GlnK. Whilst these two proteins share ∼67% similarity, they do not appear to be redundant in their function (van Heeswijk *et al.*, 1995; Atkinson & Ninfa, 1999). The role of the PII-like proteins in other organisms is only now being elucidated, but it is already clear that they have different functions. In *E. coli*, the PII paralogues are uridylylated and modulate the activity of glutamine synthetase (GS) and the transcription of the corresponding gene *glnA* (Merrick & Edwards, 1995). In cyanobacteria, the PII protein is reversibly phosphorylated on a seryl residue (Forchhammer & de Marsac, 1995*a*). The specific phospho-PII phosphatase, a member of the PP2C family, was identified recently (Irmler & Forchhammer, 2001). The cyanobacterial PII protein is involved in the regulation of nitrate and nitrite uptake (Forchhammer & de Marsac, 1995*b*; Lee *et al.*, 1998, 2000), bicarbonate transport (Hisbergues *et al.*, 1999) and is required for the activation of NtcA-dependent gene expression under conditions of nitrogen starvation (Aldehni *et al.*, 2003).

The structures of *E. coli* GlnB (EcPII) and GlnK have both been solved (Cheah *et al.*, 1994; Xu *et al.*, 1998). More recently,

the structure of the PII protein from *Herbaspirillum seropedicae* (HsPII), a member of the same order (proteobacteria) but a different subgroup to *E. coli*, has been solved (Benelli *et al.*, 2002). In all three cases, signal transduction occurs through modification of the Tyr51 residue found on the T-loop which extends out from the core of the molecule. This loop has been found to have different conformations in the *E. coli* proteins and is disordered in the complexes formed with ATP (Xu *et al.*, 1998, 2001). The T-loop is also largely disordered in HsPII. It is thought that this loop is flexible in solution and that this flexibility allows recognition to occur with several receptors. Apart from the T-loop, the C-terminus of the PII protein appears to be capable of taking on multiple conformations and may be involved in recognition. The C-terminal loop adopts one of two topologies: either a pair of short antiparallel β-strands or a 3_{10}-helix. This loop is observed to exhibit rigid-body movements relative to the core of the molecule when known structures are overlaid. The *E. coli* proteins require ATP in order to carry out their functions. Crystallographic studies have shown that ATP binds in a highly conserved cleft on the side of the molecule.

Like their *E. coli* counterparts, the PII proteins from cyanobacteria form trimers of 12 kDa subunits. For *Synechococcus* sp. PCC 7942, it has been shown that the PII protein (SnPII) can form with zero, one, two or three phosphorylated serine residues. The extent of phosphorylation depends upon the availability of carbon and nitrogen sources (Forchhammer & de Marsac, 1995*b*; Lee *et al.*, 1998). At the molecular level, the phosphorylation state of PII responds to central carbon metabolites, in particular 2-oxoglutarate, reflecting the metabolic state of the cells (Ruppert *et al.*, 2002). The serine residue in question is located in the T-loop at position 49 (Forchhammer & de Marsac, 1995*a*). When expressed in *E. coli*, the tyrosine at position 51 can be uridylylated, but it does not modulate either GS activity or the transcription of its gene (Forchhammer, 2003). This suggests that the T-loop in SnPII differs from that found in the *E. coli* paralogues. Similarly, studies on PII from *Synechocystis* sp. PCC 6803 (SsPII) have shown that the same seryl residue at position 49 is phosphorylated (Garcia-Dominguez & Florencio, unpublished work). The sequences of the T-loops in the SnPII and SsPII proteins are identical and differ from that of EcPII at 5 of the 19 residues (Fig. 1). The sequence identities between the SnPII and SsPII proteins and the EcPII

protein are 65 and 55%, respectively. This compares to 86% identity between the two cyanobacterial sequences. Although the sequences of the cyanobacterial PII proteins are very similar to those of the *E. coli* PII-like proteins, they are differently modified and have different target receptors. The cyanobacterial proteins are more closely related to the PII proteins found in plants. The sequence of the PII from *Arabidopsis thaliana* is more similar to that of the cyanobacteria than it is to *E. coli* (Smith *et al.*, 2003). In addition, the plant proteins do not have a tyrosine in the T-loop but, like the cyanobacterial proteins, have a serine that appears to be phosphorylated. We have determined the structures of the SsPII protein and a mutant form of the SnPII protein. The mutant had an alanine in place of the functionally important Ser49 and mimics the unphosphorylated form of the protein (Lee *et al.*, 2000). We comment on the functional significance of structural comparisons made between the PII proteins from *E. coli* and cyanobacteria.

2. Materials and methods

2.1. Bacterial strains and growth conditions

E. coli DH5α was used for plasmid construction and replication. *E. coli* BL21 was used for the expression of SsPII protein from *Synechocystis* sp. PCC 6803 (hereafter referred to as *Synechocystis*). The vector expressing the mutant form of the PII protein (SnPIIA) from *Synechococcus* sp. PCC 7942 (hereafter referred to as *Synechococcus*) was transformed into RB9065 cells. *E. coli* strains were grown in Luria Broth (LB) medium as described by Sambrook *et al.* (1989) supplemented with 100 µg ml^{-1} ampicillin when required.

2.2. Expression and purification

For expression of the SsPII protein in *E. coli*, a plasmid, pMAB11, was constructed as a derivative of the expression vector pET-3a (Novagen). A DNA fragment of about 400 bp encompassing the *Synechocystis glnB* gene was obtained by PCR amplification with oligonucleotides gbp1 (5′-GTACCA-CATATGAAAAAAG-3′) and gbp2 (5′-CTTGTCTGGATC-CGCCCAAC-3′). Oligonucleotides gbp1 and gbp2 harbour *Nde*I and *Bam*HI sites, respectively. The PCR product was checked by sequencing, digested with *Nde*I and *Bam*HI and cloned into the *Nde*I–*Bam*HI sites of pET-3a. Oligonucleotide gbp1 substitutes the natural translation starting codon of the *Synechocystis glnB* gene (TTGl; Garcia-Dominguez & Florencio, 1997) by a standard ATG codon, included in the *Nde*I site.

pMAB11-transformed *E. coli* BL21 strain was grown in 3 l of M9 minimal medium with ampicillin (100 µg ml^{-1}) to an optical density at 600 nm of 0.5; for production of the SsPII protein, isopropyl-β-D-thiogalactopyranoside (IPTG) was then added to a final concentration of 1 mM and the culture

Figure 1
Alignment of two cyanobacterial PII sequences with EcPII.

was further incubated for 4 h before harvesting cells by centrifugation.

Cells were resuspended in 15 ml T buffer (50 mM Tris–HCl pH 8.0, 50 mM NaCl, 5 mM MgCl$_2$, 1 mM EDTA), disrupted by sonication (20 kHz, 75 W) for 2 min and centrifuged at 40 000 g for 30 min. Cell-free extract was fractionated between 35 and 50% ammonium sulfate. After dialysis with T buffer, the protein was heated at 338 K for 5 min and centrifuged; the supernatant was then applied to a DEAE-cellulose column. The flowthrough was collected, concentrated with 50% ammonium sulfate and desalted in a Sephadex G-50 (Pharmacia) column equilibrated in T buffer. A 150 mg fraction of SsPII protein was yielded and adjusted to 10 mg ml^{-1} with T buffer.

The gene for the SnPIIA protein was obtained from plasmid pPM308 (Lee *et al.*, 2000). This plasmid was used to produce a PCR fragment with an *Nde*I and *Kpn*I site at either end. The PCR primers used for this purpose were SynPII-FP (5′-TATATT**CATATG**AAGAAGATTGAGGCGATTATTC-3′) and SynPII-RP (5′-TATATT**GGTACC**TTAGATCGTGTCG-GCGTTTTTC-3′). The PCR product was digested with *Nde*I and *Kpn*I restriction endonucleases and ligated with pND707 (Love *et al.*, 1996) that had been similarly cut. The resulting clone was checked by nucleotide-sequence analysis. The SnPIIA protein was overexpressed in RB9065 cells (UTase$^-$ mutant). The cells were grown in a 5 l fermenter of LB medium with ampicillin (100 µg ml^{-1}) at 303 K to an OD$_{595}$ of 0.6; the temperature was then rapidly shifted to 315 K and the cells were grown for a further 3 h. The cells were harvested by centrifugation, resuspended in HEPES buffer A (20 mM HEPES, 1 mM EDTA, 1 mM β-mercaptoethanol) and disrupted using a French press. The cells were pressed twice and PMSF (0.5 mM) was added before and after pressing to prevent proteolysis. Nucleic acid contaminants were removed from the sample using streptomycin sulfate [1.5%(w/v)] precipitation, followed by fractionation using ammonium sulfate (between 25 and 35%). The sample was dialysed against 3 × 2 l changes of HEPES buffer A and applied to a Cibacron Blue matrix column, washed with the same buffer containing 100 mM NaCl and finally eluted with a salt gradient in HEPES buffer (1–0 M NaCl, 5 mM ATP). The resulting protein solution was >95% in purity based on a 15% SDS–PAGE gel and was concentrated to 20 mg ml^{-1} for crystallization trials.

2.3. Crystallization, data collection and structure determination

Crystals of the SsPII protein were grown using the hanging-drop method at 293 K. Equal volumes of the protein solution (10 mg ml^{-1}) and a reservoir solution containing 20% 2-propanol, 100 mM CaCl$_2$, 100 mM sodium acetate pH 4.7 were used. Small crystals appeared after 3 d and continued to grow for about two weeks. Crystals were prepared for data collection by transfer to a solution of the reservoir buffer containing 25% glycerol.

The SnPIIA protein was crystallized using the hanging-drop method at 293 K by mixing equal volumes of the protein solution (20 mg ml^{-1}) with a reservoir solution containing 1.0 M lithium sulfate, 0.1 M Tris buffer pH 8.5 and 10 mM nickel chloride hexahydrate. Crystals were prepared for data collection by transfer to a solution of the reservoir buffer.

All data were collected using a Rigaku R-AXIS IIc image plate mounted on a rotating-anode generator running at 50 kV, 100 mA. Data were processed with the *HKL* program (Otwinowski & Minor, 1997). For SsPII, data were collected at 100 K. The crystals were found to belong to space group R3, with unit-cell parameters $a = b = 129.79$, $c = 74.14$ Å, and contained four subunits in the asymmetric unit. For the SnPII S49A mutant, data were collected at 277 K. The crystals were found to belong to space group $I4_122$, with unit-cell parameters $a = b = 109.70$, $c = 109.68$ Å, and contained a trimer in the asymmetric unit.

The structure of the SsPII protein was determined by molecular replacement using a model of a trimer of EcPII (Carr *et al.*, 1996) from which the T-loop (residues 37–55) and the four C-terminal residues had been deleted. Calculations were undertaken using the program *AMoRe* (Navaza, 1994) as implemented in the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). The best rotation-function solution had a correlation coefficient of 0.424 compared with the highest incorrect peak of 0.170. The best translation-function solution had a correlation coefficient of 0.509 and an R factor of 0.465 compared with 0.237 and 0.566, respectively, for the highest incorrect solution. After rigid-body refinement, these values improved to 0.587 and 0.428, respectively. The model generated from this solution gave good crystal packing when inspected on a graphics terminal. Further refinements were undertaken using *CNS* (Brünger *et al.*, 1998) and phases were improved using *DM* (Cowtan, 1994). During subsequent model building, it became apparent that there was density for another molecule of PII in the asymmetric unit. A further monomer of PII was added to the model. A well packed molecule was obtained from four subunits of PII. The original three molecules formed a trimer around a non-crystallographic triad, whilst the fourth produced trimers from crystallographically related molecules. A Matthews coefficient (Matthews, 1968) V_M of 2.8 Å^3 Da^{-1} is obtained for four molecules per asymmetric unit, with a corresponding solvent content of 56%. This compares with a V_M of 3.6 Å^3 Da^{-1} for a single trimer per asymmetric unit or 1.8 Å^3 Da^{-1} for two trimers per asymmetric unit.

Further rounds of refinement using simulated-annealing, positional and individual B-factor refinement using standard *CNS* protocols (Brünger *et al.*, 1998) were interspersed with manual rebuilding of the model. The final model had an R_{work} and R_{free} of 0.219 and 0.251, respectively. Refinement statistics are listed in Table 1.

The structure of the SnPIIA protein was solved by molecular replacement with *MOLREP* (Vagin & Teplyakov, 2000), incorporated in the *CCP*4 suite. The high-resolution structure of the *E. coli* PII trimer without the T-loop was used as a search model (Carr *et al.*, 1996). *MOLREP* gave rise to a

Table 1
Summary of X-ray diffraction data and refinement statistics.

Values in parentheses are for the outer data shell.

	SnPII	SsPII
X-ray diffraction data		
Resolution (Å)	25–2.0 (2.1–2.0)	25–2.0 (2.1–2.0)
Observed reflections	68888	156138
Unique reflections	19247	31543
Completeness (%)	84.0 (80.2)	99.5 (98.2)
R_{merge} (%)	5.9 (24.1)	5.3 (16.7)
Average $I/\sigma(I)$	8.5 (3.1)	16.7 (4.93)
Refinement statistics		
Resolution (Å)	25–2.0	25–2.0
No. of reflections (working)	17030	29381
No. of reflections (test)	1869 (8.2%)	1577 (5%)
R_{cryst}/R_{free}† (%)	20.0/22.7	21.9/25.3
No. protein atoms	2610	3499
No. solvent atoms	132	217
No. glycerol molecules	—	2
No. ions	3 Ni^{2+}, 6 SO_4^{2-}	2 Ca^{2+}
R.m.s. deviation from ideal geometry		
Bond length (Å)	0.005	0.006
Bond angle (°)	1.28	1.26
Dihedrals (°)	23.64	26.64
Improper (°)	0.67	1.17
Average B factors (Å^2)		
Protein	18.4	29.8
Water	30.0	42.3
Sulfate	43.4	—
Glycerol	—	51.2
Ions	23.7	34.2
Ramachandran plot (%)		
Most favoured regions	96.9	98.1
Additional allowed regions	3.1	1.9
Generously allowed regions	0.0	0.0
Disallowed regions	0.0	0.0

† $R_{cryst} = (|F_{obs}| − |F_{calc}|)/|F_{obs}|$; R_{free} was calculated in the same manner as R_{cryst} using a test set of data not used in the refinement process.

dominant solution with $R = 0.481$ and a correlation coefficient of 0.476 to 3.0 Å resolution. Rigid-body refinement and further structure refinement were implemented using *CNS* v.1.1 and model building was carried out with *O* (Jones *et al.*, 1991). Non-crystallographic symmetry restraints were applied throughout the refinement. The occupancies of disordered side chains were set to zero during the refinement. A number of non-peptide features appeared in the electron-density map during the course of refinement. Within each ATP-binding site, there were two features that appeared to be sulfates from the crystallization buffer. In addition, there was a spherical feature in the maps that appeared to be Ni^{2+} ions, also from the crystallization buffer. Statistics of data collection and structure refinement are given in Table 1.

3. Results and discussion

3.1. The models

The models confirm that the cyanobacterial PII proteins have the same overall topology as other known PII structures. The molecules form homotrimers around threefold axes, with β-sheet interactions involving both of the neighbouring protomers. These hydrogen bonds hold the central core structure of the trimers together. Each monomer consists of a

Table 2
R.m.s. displacement values of superposition of main-chain atoms of various PII structures.

Species	Space group	Chain	R.m.s. (Å)
Synechocystis	$R3$	*A*	0.60
		B	0.59
		C	0.58
		D	0.00
Synechococcus	$I4_122$		1.24
E. coli	$P6_3$		0.95
	$I2_13$		1.19
E. coli GlnK	$P2_13$		1.08
H. seroedicae	$P2_12_12_1$	*A*	0.92

central four-stranded antiparallel β-sheet with a pair of antiparallel helices on one side. These form an interlocking double $\beta\alpha\beta$ motif. There are two major loops emanating from the central sheet. The first and largest, between strands 2 and 3, is known as the T-loop and contains the site of phosphorylation. In structures solved to date, this loop has only been visible where crystal-packing forces have stabilized one particular conformation. This has led to the conclusion that it is flexible in solution. The T-loop is not visible in the SsPII structure reported here, but is visible in the electron density of SnPII (Fig. 5). Lattice contacts appear to stabilize this conformation. The second loop is found between strand 4 and the carboxyl-terminus. In both of the structures reported here and in several other PII structures this loop contains two short antiparallel β-strands. In other PII-like structures, for example, EcGlnK and HsPII, this loop contains a 3_{10}-helix rather than β-strands.

The model of SsPII consists of four protein chains. Chains *A*, *B* and *C* form a trimer around a non-crystallographic triad. Chain *D* forms a trimer by symmetry operations on itself. Like several other forms of PII (Xu *et al.*, 2001), the T-loop (residues 37–53) is not visible in the electron density, presumably owing to mobility. The model consists of residues *A*1–36, *A*54–112, *B*1–36, *B*55–112, *C*1–36, *C*55–112, *D*1–38 and *D*54–110. In addition, two Ca^{2+} ions, two glycerol molecules and 217 water molecules were included. The following residues did not exhibit good side-chain density and were truncated to alanines in the refinement: *A*76, *A*101, *A*106, *B*76, *C*76, *D*38 and *D*54. The following glutamate residues exhibited alternate conformations of their side chains: *A*15, *A*62, *B*5, *B*15, *B*62, *C*5, *C*15 and *D*15. Alternate conformers were modeled at half-occupancy. The final model exhibited good stereochemistry when checked by *WHAT_CHECK* (Hooft *et al.*, 1996) and *PROCHECK* (Laskowski *et al.*, 1993). The Ramachandran plot showed 98.1% of residues falling into the most favoured region, whilst the remaining 1.9% were in additional allowed regions. No residues fell into the disallowed or generously allowed regions.

The overall fold of the SsPII protein is essentially the same as that of EcPII protein. The individual protein chains were overlaid using *LSQKAB* (Kabsch, 1976). The r.m.s. values for main-chain atom displacement were less than or equal to 0.6 Å. For comparison, the *D* chain was also superimposed on
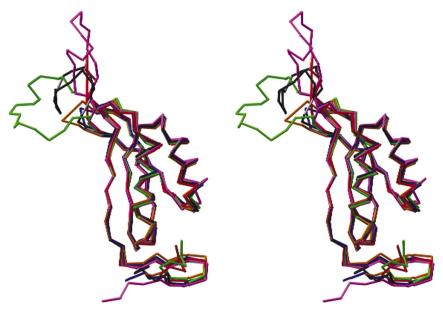
Figure 2
Superposition of various PII monomers. Colour scheme: SnPIIA, magenta; SsPIIA, blue; SsPIID, orange; EcPII, black; GlnKB, green; HsPII, red. Note that T-loops (upper left), where present, show variable conformations. The C-loop structure (lower right) undergoes rigid-body movements relative to the stable double $\beta\alpha\beta$ core between structures.

Inspection of the temperature factors reveals that the central rigid core of interlocking β-strands has the lowest B factors ($\sim$18 Å^2). The B-loop and helices have values that are typically double those of the core. The C-loop, however, is more stable ($\sim$25 Å^2) owing to inter-trimer strand formation. There is variation between the individual protomers that cluster around the non-crystallographic threefold axis, although it difficult to see why this should be as they have similar packing environments.

The crystal packing appears to have influenced the current structure by interactions at two sites. The first is at the Ca^{2+} ion-binding sites, which occur between symmetry-related molecules. The metal ion is liganded to the acidic side chains of residues Asp14 and Glu15 and the carbonyl O atom of Phe11. It is also bound to the symmetry mates of these three atoms in a neighbouring molecule, giving the Ca^{2+} ion full octahedral coordination (Fig. 3). The
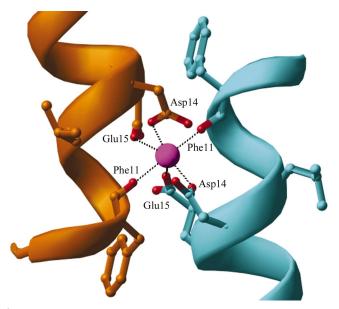

Figure 3
Close-up of crystal contact in SsPII at the Ca^{2+}-binding site. Colour coded by (symmetry-related) protein chains.

a number of other PII structures (Table 2). In all of the superpositions the basic structure of the protein chains align well, with just three regions showing variations. Firstly, in structures where crystal packing stabilizes the T-loop (EcPII, EcGlnK and SnPIIA) it adopts different conformations. Secondly, the C-loop, which exhibits β-sheet topology, undergoes small rigid-body movements with respect to the central strands. Thirdly, the last three C-terminal residues differ between structures. Fig. 2 shows an overlay of the various PII protein chains.
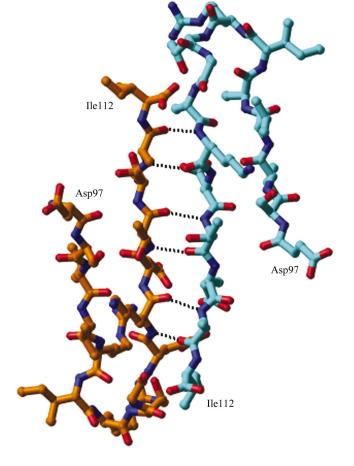

Figure 4
Hydrogen bonds stabilize the C-loops of symmetry-related molecules of SsPII.

second site of packing-influenced interactions occurs at the final β-strand, which forms part of the C-loop. Here, the backbone amide N and carbonyl O atoms of residues 107 and 109 form hydrogen bonds in a classic antiparallel β-sheet arrangement with the same residues in a neighbouring symmetry-related molecule (Fig. 4). This has the effect of producing a large extended β-sheet of 12 strands spanning two trimers. Similar packing interactions were observed in the EcPII $I2_13$ structure (Xu *et al.*, 2001).

The model of the SnPIIA protein consists of a trimer of almost identical subunits. The T-loops are clearly ordered as shown in Fig. 5. The main interaction stabilizing the T-loops appears to be main-chain hydrogen bonds forming β-sheets involving residues 49–53 with their symmetry-equivalent residues from another trimer. Although Ser49 has been mutated to alanine in this structure, it does not appear that side-chain interactions involving the serine would affect the

local conformation of the loop. The side chains of residues Arg38, Gln39, Lys40, Gln42 Arg47, Glu50, Gln57, Gln69 and Glu85 in all three subunits have poorly defined electron density. The C-terminal loop has β-sheet topology. There are three Ni^{2+} ions that link the N- and C-termini of each subunit, as shown in Fig. 6. In addition, there are 132 solvent molecules and six sulfates that bind in the three ATP-binding clefts. *PROCHECK* (Laskowski *et al.*, 1993) indicates that 96.7% of the φ, ψ angles of all non-glycyl residues are in the most favoured regions of the Ramachandran plot. The additional 3.3% were in the additionally allowed regions.

Inspection of the B factors again shows the central core of interlocking β-strands to have the lowest mobility ($\sim$12 Å^2). All loops, including the T-, B- and C-loops, have values that are approximately doubled. The external helices are intermediate between these two values. The nickel ions have B factors similar to the liganding residues. The SO_4^{2-} ions have higher B factors than the nickel ions, closer to typical solvent values.

3.2. Molecular structures

The tertiary structures of the cyanobacterial SsPII and SnPIIA proteins are essentially the same as the EcPII and GlnK structures. However, there are significant differences between the T- and C-loop regions of SsPII and SnPII and the same loops in EcPII and GlnK.

The T-loop of the SsPII protein is disordered as has been observed in some EcPII and GlnK structures. The T-loop in the SnPIIA protein extends from $\beta2$ and $\beta3$ without a disruption of secondary structure, so that these strands are longer than those found in EcPII and GlnK. $\beta2$ of SnPIIA consists of residues 29–41, this being five more than the number of residues in $\beta2$ (37–41) of the EcPII molecule. Similarly, $\beta3$ of SnPIIA consists of residues 50–65, while $\beta3$ of the other PII molecules consists of residues 56–65. There is little contact between the extension of $\beta2$ and $\beta3$ and the core of the *Synechococcus* PII molecule. The T-loop region of SnPIIA takes part in crystal packing. Residues 49, 51 and 53 in a given monomer of the trimer form hydrogen bonds with their symmetry counterparts of one monomer within a neighbouring trimer. The specific interactions involve backbone hydrogen bonds between Tyr51 and its crystallographic equivalent, while Ala49 and Val53 are linked to Val53 and Ala49, respectively, in a crystallographically related molecule. These interactions appear to stabilize the T-loop.
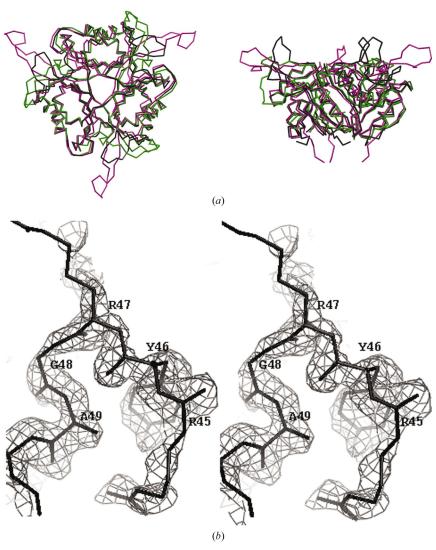

Figure 5
Trimers of PII with visible T-loops. All are stabilized by crystal contacts. Colour scheme: SnPII, magenta; GlnK, green; EcPII, black. (*a*) Left, view looking down the threefold axis. Right, perpendicular to the threefold axis looking into the ATP-binding cleft. (*b*) An omit map of the T-loop region in SnPII.

The C-terminus of the SsPII protein is very similar to that observed in a low-pH structure of EcPII (Xu *et al.*, 2001). In both, the C-terminus is involved in crystal contacts. The conformation of the C-terminus of SsPII differs from that observed in EcPII and GlnK structures obtained at higher pH values. The conformation of the C-terminus of SnPIIA differs from those of other PII proteins. It points towards the N-terminus of the same chain and is held in place by coordination with a metal ion, presumably a Ni^{2+} that was added for crystallization. The coordination site involves Asp110, a residue that is not conserved in either the SsPII or the EcPII proteins. The metal-ion coordination ties down the C-terminus and probably facilitates crystallization, but is not likely to have a physiological function.

SnPIIA binds two anions in the ATP-binding cleft. In the absence of ATP, GlnK was also observed to bind an anion in this cleft. The GlnK anion-binding site also appears to be occupied in SnPIIA. This anion forms interactions with the backbone N atom of Gly87 and the side-chain N atoms of Arg101 and Arg103 fr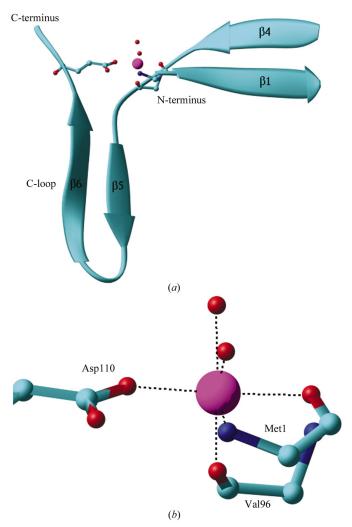om a neighbouring monomer within the trimer. The second anion forms links to the side-chain N atoms of Lys90 and the neighbouring Arg101 as well as the symmetry-related Arg103. The latter anion is exposed to solvent and has a higher B factor.

4. Concluding remarks

The structures observed for the cyanobacterial PII proteins are very similar to those observed for the *E. coli* proteins. The core of the molecule is remarkably well preserved; this similarity is particularly remarkable when it is recalled that the cyanobacterial PII proteins interact with a set of receptors that are quite different to those of *E. coli*. This is consistent with the idea that recognition occurs through flexible loops and that the flexibility allows the recognition of multiple receptors (Xu *et al.*, 1998). The T-loop in the present structures is disordered in one case and in the other is dependent upon crystal packing. Clearly, this piece of peptide is highly flexible. The C-terminal peptide is also capable of assuming multiple conformations although it tends to favour a 3_{10}-helical or a strand structure. As noted earlier (Xu *et al.*, 2001), it may favour a helical structure in solution, but this structure cannot be particularly stable. Indeed, the flexibility of this peptide may also be important for receptor recognition.

The function of ATP in the PII-like proteins is not exactly clear, yet its binding pocket is highly conserved in terms of sequence (Xu *et al.*, 1998). The cyanobacterial PII structures again show a highly conserved binding cleft.

There are features of the cyanobacteria proteins that appear unique. A BLAST search to find other proteins that were similar to *Synechococcus* PII gave a total of 167 hits. Amongst these sequences, cyanobacteria (*Synechocystis*, *Nostoc punctiforme*, *Fremyella diplosiphon*, *Thermosynechococcus elongatus* BP-1 and *Prochlorococcus marinus*) have a conserved Gln at the position 57. All proteobacteria and actinobacteria have a proline at position 57 in their GlnB and GlnK proteins. Other organisms have a variable non-proline at the corresponding position. These organisms include the archaebacteria, aquificaceae, thermatogales, deinococci, red algae and dicotyledenous plants. Like the cyanobacteria proteins, these proteins are not uridylylated as evidenced by the fact that *glnD* (coding for UTase) does not exist in their genomes. Residue 57 is located in a section of the molecule that is likely to effect receptor recognition. We suggest that residue 57 may be involved in recognition events that are specific for cyanobacteria.

Figure 6
Ni^{2+} ion linking the N- and C-termini in SnPII. (*a*) Location of the metal-binding site. (*b*) Close-up showing the metal coordination. Unbonded red spheres represent solvent molecules.

References

Aldehni, M. F., Sauer, J., Spielhaupter, C., Schmid, R. & Forchhammer, K. (2003). *J. Bacteriol.* **185**, 2582–2591.

research papers

Arcondeguy, T., Jack, R. & Merrick, M. (2001). *Microbiol. Mol. Biol. Rev.* **65**, 80–105.

Atkinson, M. R. & Ninfa, A. J. (1999). *Mol. Microbiol.* **32**, 301–313.

Benelli, E. M., Buck, M., Polikarpov, I., de Souza, E. M., Cruz, L. M. & Pedrosa, F. O. (2002). *Eur. J. Biochem.* **269**, 3296–3303.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Carr, P. D., Cheah, E., Suffolk, P. M., Vasudevan, S. G., Dixon, N. E. & Ollis, D. L. (1996). *Acta Cryst.* D**52**, 93–104.

Cheah, E, Carr, P. D., Suffolk, P. M., Vasudevan, S. G., Dixon, N. E. & Ollis, D. L (1994). *Structure*, **2**, 981–990.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cowtan, K (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.

Forchhammer, K. (2003). *Symbiosis*, **35**, 101–115.

Forchhammer, K. & de Marsac, N. T. (1995a). *J. Bacteriol.* **177**, 5812–5817.

Forchhammer, K. & de Marsac, N. T. (1995b). *J. Bacteriol.* **177**, 2033–2040.

Garcia-Dominguez, M. & Florencio, F. J. (1997). *Plant Mol. Biol.* **35**, 723–734.

Heeswijk, W. C. van, Stegeman, B., Hoving, S., Molenaar, D., Kahn, D. & Westerhoff, H. V. (1995). *FEMS Microbiol. Lett.* **132**, 153–157.

Hisbergues, M., Jeanjean, R., Joset, F., de Marsac, N. T. & Bédu, S. (1999). *FEBS Lett.* **463**, 216–220.

Hooft, R. W. W., Vriend, G., Sander, C. & Abola, E. E. (1996). *Nature (London)*, **381**, 272.

Irmler, A. & Forchhammer, K. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 12978–12983.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.

Laskowski, R. A., MacArthur, M. W., Moss, D. W. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Lee, H. M., Flores, E., Forchhammer, K., Herrero, A. & de Marsac, N. T. (2000). *Eur. J. Biochem.* **267**, 591–600.

Lee, H. M., Flores, E., Herrero, A., Houmard, J. & de Marsac, N. T. (1998). *FEBS Lett.* **427**, 291–295.

Love, C. A., Lilley, P. E. & Dixon, N. E. (1996). *Gene*, **176**, 49–53.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Merrick, M. J. & Edwards, R. A. (1995). *Microbiol. Rev.* **59**, 604–622.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Ninfa, A. J. & Atkinson, M. R. (2000). *Trends Microbiol.* **8**, 172–179.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Ruppert, U., Irmler, A., Kloft, N. & Forchhammer, K. (2002). *Mol. Microbiol.* **44**, 855–864.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Smith, C. S., Aalim, M. W. & Moorhead, G. B. G. (2003). *Plant J.* **33**, 353–360.

Vagin, A. & Teplyakov, A. (2000). *Acta Cryst.* D**56**, 1622–1624.

Xu, Y., Carr, P. D., Huber, T., Vasudevan, S. G. & Ollis, D. L. (2001). *Eur. J. Biochem.* **268**, 2028–2037.

Xu, Y., Cheah, E., Carr, P. D., van Heeswijk, W. C., Westerhoff, H. V., Vasudevan, S. G. & Ollis, D. L. (1998). *J. Mol. Biol.* **282**, 149–165.