REVIEW ARTICLE

P$_{II}$ signal transduction proteins: nitrogen regulation and beyond

Luciano F. Huergo$^1$, Govind Chandra$^2$ & Mike Merrick$^2$

$^1$Departamento de Bioquímica e Biologia Molecular, Instituto Nacional de Ciência e Tecnologia da Fixação Biológica de Nitrogênio, UFPR Curitiba, PR, Brazil; and $^2$Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Norwich, UK

Correspondence: Mike Merrick, Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, UK. Tel.: +44 1603 450 749; fax: +44 1603 450 018; e-mail: mike.merrick@jic.ac.uk

Received 20 March 2012; revised 26 July 2012; accepted 26 July 2012.

DOI: 10.1111/j.1574-6976.2012.00351.x

Editor: Hauke Hennecke

Keywords
nitrogen metabolism; GlnB; GlnK; protein complexes; allosteric regulation; post-translational regulation.

Abstract
The P$_{II}$ proteins are one of the most widely distributed families of signal transduction proteins in nature. They are pivotal players in the control of nitrogen metabolism in bacteria and archaea, and are also found in the plastids of plants. Quite remarkably, P$_{II}$ proteins control the activities of a diverse range of enzymes, transcription factors and membrane transport proteins, and in recent years the extent of these interactions has been recognized to be much greater than heretofore described. Major advances have been made in structural studies of P$_{II}$ proteins, including the solution of the first structures of P$_{II}$ proteins complexed with their targets. We have also begun to gain insights into how the key effector molecules, 2-oxoglutarate and ATP/ADP, influence the activities of P$_{II}$ proteins. In this review, we have set out to summarize our current understanding of P$_{II}$ biology and to consider where future studies of these extraordinarily adaptable proteins might lead us.

Introduction
In 1968, Bennett Shapiro was studying the process by which the activity of a central enzyme of bacterial nitrogen metabolism, glutamine synthetase (GS), was regulated by adenylylation or deadenylylation of the protein’s 12 subunits. He was attempting to purify the enzyme responsible for deadenylylation of GS and he succeeded in partially purifying two components, both of which were required for the deadenylylating activity. The two components, which were separated by gel filtration chromatography, were labelled by Shapiro as P$_I$ and P$_{II}$ (fig. 1 in Shapiro, 1969), and quite remarkably the protein responsible for the activity in the second peak has retained the simple and unique designation, P$_{II}$, to this day. Little did Shapiro know that he had identified and named one of the most widely distributed signal transduction proteins in nature, the study of which is till today providing remarkable insights into the control of metabolism in bacteria, archaea and plants.

A detailed historical perspective on the elegant biochemical studies that dissected the regulation of GS activity in Escherichia coli is beyond the scope of this review. However, it is summarized in our previous review of P$_{II}$ biology (Arcondéguy et al., 2001) and is described in detail by Earl Stadtman, the key player in the field, in a review he wrote in 1990 (Stadtman, 1990). It wasn’t until 1975 that the P$_{II}$ protein was shown to be the product of a gene called glnB, first described by Magasanik and co-workers 2 years earlier (Prival et al., 1973; Foor et al., 1975). The E. coli glnB gene was finally sequenced in 1987 (Son & Rhee, 1987), thus providing the key information that would subsequently allow recognition of the fact that P$_{II}$ proteins are ubiquitous in bacteria, archaea and plants.

P$_{II}$ proteins function in signal transduction by virtue of their abilities to regulate the activity of a huge variety of target proteins by protein–protein interaction. These targets include membrane proteins, enzymes and transcription factors, the great majority of which are involved in some aspect of cellular nitrogen metabolism. For this reason, P$_{II}$ proteins have become the focus of research in many laboratories studying a variety of biological problems in bacteria, archaea and plants. In the 40 years since their discovery we have learnt a great deal about the role and mode of action of the P$_{II}$ proteins. However, some key facts have only been established within the last few years and still others remain to be elucidated. In recent
years, a number of reviews on specific facets of P II biology have been published (Forchhammer, 2004; Forchhammer et al., 2004; Ninfa & Jiang, 2005; Commichau et al., 2006; Forchhammer, 2007, 2008; Radchenko & Merrick, 2011), but in this review we have attempted to give a broad perspective of the current knowledge of P II biology, and to consider some of the challenges that remain to be tackled.

**P II protein nomenclature and distribution**

It wasn’t until 1995 that the presence of a second P II-like protein in *E. coli* was recognized, for which the structural gene was designated *glnK* (van Heeswijk et al., 1995, 1996). The rapid growth in gene sequencing during the 1990s led to the recognition of many more genes encoding members of the P II family, but the annotation of those P II genes that were linked to *glnA* was rather haphazard with the result that a complex and confusing situation developed. In 2001, Arcondéguy et al. proposed a new nomenclature (Arcondéguy et al., 2001) in an attempt to create a rational framework for describing members of the P II family. They proposed three major subgroups, *glnB*, *glnK* and *nifI*, based on two primary criteria, namely conservation of genetic linkage and similarity at the level of primary amino acid sequence. The proposed *glnB* genes were predominantly linked to *glnA* (the GS structural gene) or *nadE* (encoding NAD synthetase); *glnK* was used for those P II genes that were linked to *amtB* (the structural gene for the ammonia channel protein); and a new designation, *nifI*, was proposed for P II-like genes associated with the structural genes for nitrogenase (*nifH*, *nifD* and *nifK*). The *nifI* genes are normally represented by two adjacent and closely related genes designated *nifI1* and *nifI2* (Leigh & Dodsworth, 2007).

In the last decade, genomic sequencing efforts have identified an ever-increasing number of predicted P II genes that required a re-evaluation of the nomenclature proposed by Arcondéguy et al. (2001) and such a study was carried out by Sant’Anna et al. (2009). Their comprehensive analysis provided new information on the definition and distribution of each parologue, and the conservation of their genetic linkage. It also offered a better comprehension of their probable evolution. Sant’Anna et al. (2009) analysed over 700 sequences from the non-redundant NCBI database. Their survey of the P II superfamily showed that P II homologues group together in accordance with three subfamilies, GlnB/K, NifI, and an uncharacterized group which they called P II-New Group (PII-NG).

This new analysis upheld the basic proposal of Arcondéguy et al. (2001), namely that the majority of *glnK* genes are linked to *amtB*, *glnB* genes are associated with *glnA* or *nadE*, and *nifI* genes are invariably linked to a nitrogenase gene cluster. Furthermore, although there are many P II genes that show other genetic associations, most P II proteins can be phylogenetically related to a particular subfamily. GlnB proteins are distributed predominantly within the *Proteobacteria*, whereas GlnK-like proteins show a broader distribution amongst different taxa. The definition of GlnK proteins was broadened by Sant’Anna et al. (2009) to include some proteins that appear to be phylogenetically linked to GlnK, but whose structural genes are not always genetically linked to *amtB*. A clear example of this is the *Azospirillum brasilense* GlnZ protein that has been shown to have a similar cellular function to *E. coli* GlnK, namely regulation of AmtB, despite the fact that glnZ is not linked to *amtB* (Huergo et al., 2007). Other examples of GlnK proteins that have been given alternative designations occur where an organism has multiple *glnK* *amtB* clusters. These include *Azorarcus* which has two clusters designated *glnK* *amtB1* and *glnK* *amtB2* (Zhang et al., 2001b, 2006b). Within the *Proteobacteria*, the two major P II families are GlnK and GlnB, although some taxonomic groups only encode GlnK. The *Archaea* and the *Firmicutes* encode both GlnK and NifI, although the archaeal GlnK proteins are significantly distinct from the proteobacterial GlnKs (Sant’Anna et al., 2009).

The detailed phylogenetic analysis of Sant’Anna et al. (2009) allowed a re-evaluation of the possible evolutionary history of the P II proteins. Their conclusions support the earlier proposal of Thomas et al. (2000) that the genetic association of *glnK* and *amtB* represents the origin of the P II proteins. They suggest that the *glnK* and *amtB* genes most likely arose in the *Archaea*, and were subsequently transferred between deep-branching lineages of the *Archaea* and *Bacteria*. The subsequent association of the *glnA* gene to give a *glnA*, *glnK*, *amtB* cluster is supported by the occurrence of such a genetic association in *Dehalococcoides ethenogenes* (*Chloroflexi*) and *Aquilox aerolicus* (*Aquificae*) both of which are deep-branching lineages of the *Bacteria*. Subsequent gene duplication in such a bacterial ancestor could then have led to the commonly found *glnK*, *amtB* and *glnB*, *glnA* gene associations.

The NifI proteins divide into two well-supported clades comprising NifI1 and NifI2, and hence the *nifI1* and *nifI2* genes appear to be the products of an ancient duplication event that occurred before the speciation of the last common ancestor. The species within which the *nifI* genes are found are somewhat distantly related and include both archaeal and bacterial species. Furthermore,
their distribution, like that of the nitrogenase genes (Raymond et al., 2004) suggests that it has arisen by horizontal gene transfer.

The phylogenetic studies of both Sant’Anna et al. (2009) and Osanai & Tanaka (2007) support the notion of an origin for the PII proteins in plants as a consequence of a cyanobacterial endosymbiosis leading to the generation of the chloroplast. This hypothesis explains the absence of PII proteins from all Eukaryotes other than the plants. The PII proteins of the Cyanobacteria and the plants cluster in a single group that is closely related to the GlnB/K group. The majority of both cyanobacteria and plants encode a single PII, but PII genes are absent from some of the algae (Uhrig et al., 2009). All the PII proteins so far characterized in the Plantae are found in the chloroplast. In red algae, the PII gene has been mapped to the chloroplast genome (Reith & Munholland, 1995), but in higher plants such as Arabidopsis thaliana the PII gene is in the nuclear genome and the protein is targeted into the plastid by a signal sequence (Hsieh et al., 1998).

In the phylogenetic analysis of Sant’Anna et al. (2009), one clade stands out as being quite distinct: these are the previously mentioned PII-NG proteins. The proteins in this class are devoid of both the recognized PII PROSITE signature patterns (PS00496 and PS00638), but are clearly PII-related on the basis of sequence homology. They are found in all subdivisions of the Proteobacteria (except the Epsilon subdivision) and in some Bacteroidetes. Most remarkably, the PII-NG proteins show highly conserved genetic linkage in that they are adjacent to genes predicted to encode heavy metal efflux pumps. Despite diverse annotation, these genes appear to correspond to the cccCBA genes which encode a proton-cation antiporter for Co$^{2+}$, Zn$^{2+}$ and Cd$^{2+}$. To date, there has been no experimental analysis of the possible role of the PII-NG proteins. The genetic association with genes encoding proteins that could confer resistance to toxic heavy metals suggests that they might be implicated in heavy metal sensing as opposed to the nitrogen-sensing that has here-tofore characterized PII proteins.

The explosion in genome sequencing in recent years allows new information to be derived, both about the distribution of PII proteins and potentially about some of their targets. As already mentioned, this information allowed Sant’Anna et al. (2009) to revise our picture of the PII phylogeny, but as many more genome sequences have been published since that work in 2009 we have carried out new analyses using the completed prokaryotic genomes deposited in Genbank in late 2011. At the time of the analysis, this constituted 2783 Genbank files made up of 120 archaeal genomes and 66 archaeal plasmids, 1493 bacterial genomes and 1104 bacterial plasmids; recognizing of course that the distinction between genome and plasmid can be quite complex.

All these files were searched for PII-encoding genes using BLAST with GlnB, GlnK and NifII as query sequences. GlnB, GlnK, PII-NG and, NifII proteins were used as queries to search, by BLASTP, for PII proteins in all bacterial GenBank files: a low complexity filter was not applied. The expected value threshold was set to 1e−2 and an additional condition of at least 30% amino acid identity with the query protein was imposed. Proteins thus found were assigned to categories (GlnB, GlnK etc.) based on the query protein with which they gave the highest bit score in BLASTP. BLASTP results were parsed into a tabular form and manually checked for false positives and misassignments based on coverage in alignment and genomic context.

One of the first observations from this analysis is that, although PII proteins are undoubtedly found in a huge range of prokaryotes, amongst the sequences analysed, 36 archaeal species and 291 bacterial species do not encode PII proteins (Table S1). In some cases, the absence of PII is apparently a characteristic of a whole genus (19 archaeal genera and 115 bacterial genera) whereas in other cases only some species within the genus lack PII (2 archaeal genera and 24 bacterial genera; Table S2). These lists include many parasitic organisms such as Mycoplasma spp. and Rickettsia spp. which, as observed by Thomas et al. (2000), will rely on their hosts for provision of nitrogen-containing compounds and therefore probably have little need for complex nitrogen control mechanisms. However, it is not apparent that such arguments will explain all the organisms that lack PII and a detailed analysis might reveal some novel physiology and metabolism.

Conserved genetic association in prokaryotes can often indicate related functions. This is very clearly exemplified by the highly conserved linkage between glnK and amtB that led directly to studies on the interaction between GlnK and the ammonia channel protein (Thomas et al., 2000; Couotts et al., 2002). Likewise, the NifI proteins are characterized by linkage to the genes for nitrogenase, an enzyme whose activity they control (Dodsworth & Leigh, 2007). Hence, we have screened the same Genbank dataset for any previously unobserved conserved linkage that might suggest other so far unexplored targets for PII proteins.

The most prominent example is the frequent linkage between glnB and nadE encoding NAD synthetase. This linkage was remarked on by Arcondéguy et al. (2001), but there are still no reports of studies that have explored NAD synthetase as a potential PII target. Within eubacteria, NadE proteins fall into two classes: a single domain ammonium-dependent enzyme, and a two domain
glutamine-dependent enzyme (Pace & Brenner, 2001) in which the N-terminal domain has been shown to function as a glutamine amidotransferase (Bellinzoni et al., 2005). In those cases where glnB and nadE are linked, the NAD synthetase appears to be of the glutamine-dependent type and hence it seems likely that the activity of these enzymes might be regulated by interaction with PII in response to the intracellular glutamine availability.

As observed by Sant’Anna et al. (2009) many PII genes are flanked by genes other than glnA, amtB, nadE, czcA and nifH. When those five classes of linkage were excluded from the dataset no other major linkage group was identified, although in many cases the flanking genes are annotated as hypothetical and this may obscure a yet unidentified target group. Within the Proteobacteria, glnB is often linked to a two-component regulatory pair (YhaA, YfhK), but the role of these regulators is unknown and there is currently no evidence that they have any functional relationship to GlnB. However argB, the structural gene for N-acetyl-L-glutamate kinase (NAGK) is found linked to glnB in some species of Methanococcus, and this might well reflect a functional interaction between PII and NAGK in those organisms. Although we have not identified any new major linkage groups, flanking PII genes may still sometimes identify PII targets. As an example, PII interaction with N-acetyl-L-glutamate kinase (NAGK) was originally identified through yeast two-hybrid studies (see later). Similarly, in a number of species within the Aquifacaceae, including Aquifex aeolicus, Hydrogenobacter thermophilus and Thermocrinis albus, a PII gene is closely linked to genes that are predicted to encode a nitrate transporter and nitrate or nitrite reductases, reinforcing circumstantial evidence from cyanobacteria and plants (Ferrario-Mery et al., 2008; Forchhammer, 2010) that PII may sometimes regulate nitrate utilization.

In summary, our current understanding of the PII superfamily and its evolution undoubtedly supports the thesis that these proteins are indeed one of the most ancient families of signal transduction proteins. It also provides an excellent framework against which to consider the biology of PII proteins and how that might have evolved to result in the incredibly diverse functions that they exhibit today.

**PII protein biochemistry**

**The structure of PII proteins and location of effector binding sites**

Crystallization studies and sedimentation equilibrium analysis of E. coli GlnB established that it is a homotrimer (Cheah et al., 1994; Vasudevan et al., 1994) and it seems likely that most PII proteins will be similar, although Streptococcus mutans GlnK is reported to be hexameric, as is the Methanococcus maripaludis NifI1,2 complex (Dods- worth & Leigh, 2006; Portugal et al., 2011). Given the proposition from phylogenetic analysis that PII proteins evolved to regulate the ammonia channel Amt proteins, as evidenced by the early linkage of glnK with amtB, then this trimeric structure can be related directly to the fact that AmtB is also trimeric (see later). Since the initial solution of the X-ray crystal structure of E. coli GlnB (Cheah et al., 1994) at least 47 PII protein structures have been deposited in the Protein Data Bank (PDB). These represent 19 different proteins from 16 different organisms including bacteria, archaea and plants (Table 1). However, to date there are no crystal structures from representative NifI or PII-NG proteins.

Within the GlnB/K subfamily, the primary amino-acid sequence is strongly conserved so that it is usually precisely 112 amino acids, and the known PII structures present a very unified picture. The PII trimer typically forms a compact barrel around 30 Å high in which each monomer comprises two α-helices and four β-strands arranged so that the α-helices and β-strands form a double βαβ motif connected by a large loop of 19 amino acids (Fig. 1). In many bacteria, this loop contains a site for post-translational modification. This was first characterized in E. coli where residue Tyr51 is subject to uridylylation (Son & Rhee, 1987) and consequently the loop was designated the T-loop (Cheah et al., 1994). A smaller loop (the B-loop) is located between the second α-helix and the fourth β-strand, and a third loop (the C-loop) is located at the C-terminus (Fig. 1).

Within the PII trimer, the T and B loops of one monomer and the C-loop of the adjacent monomer form an inter-subunit cleft that is now known to constitute a ligand binding site (Fig. 1). The trimer is a remarkably stable molecule with a melting temperature that is typically between 60 and 70 °C, and this heat stability can be used to facilitate PII purification (Moure et al., 2012). The T-loops are inherently flexible and are consequently often unresolved in X-ray crystal structures. Furthermore, when they are resolved they are often constrained by the crystal lattice (Cheah et al., 1994; Xu et al., 1998) such that the apparent structure of the loop may not reflect a particular physiologically relevant conformation. The potential flexibility of the T-loop was recognized early on as offering properties that were well suited to interaction with PII target proteins and the T-loop was shown to be essential for interaction of E. coli GlnB with three of its known targets, adenyllyltransferase (GlnE), uridylyltransferase (GlnD) and the histidine protein kinase NtrB (Jaggi et al., 1996; Jiang et al., 1997b, c).

Early studies with E. coli GlnB implicated ATP, Mg²⁺ and 2-oxoglutarate (2-OG) as potential effector molecules...
that could facilitate PII-mediated regulation of its targets (Kamberov et al., 1995). The B-loop region of GlnB was recognized early on as containing a sequence similar to the ‘Walker A’ motif that typifies an ATP-binding site in many proteins (de Mel et al., 1994; Kamberov et al., 1994b; Carr et al., 1996), and the X-ray crystal structure for E. coli GlnK bound to ATP confirmed that the inter-subunit cleft was the site of ATP binding such that each trimer could bind three molecules of ATP (Xu et al., 1998). The B-loop motif TGxxGDGKI constitutes one of the most conserved regions in all PII proteins, emphasizing that ATP-binding is a key property of the whole PII superfamily (Xu et al., 1998) although the biological role of ATP was not apparent from the various PII-ATP

<table>
<thead>
<tr>
<th>Organism</th>
<th>Protein</th>
<th>Ligand</th>
<th>Complex with</th>
<th>PDB code</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. aeolicus</td>
<td>P II</td>
<td></td>
<td></td>
<td>2Z0G</td>
<td>Unpublished</td>
</tr>
<tr>
<td></td>
<td>P II</td>
<td>ATP</td>
<td></td>
<td>2EG1</td>
<td>Unpublished</td>
</tr>
<tr>
<td></td>
<td>P II</td>
<td>ATP</td>
<td></td>
<td>2EG2</td>
<td>Unpublished</td>
</tr>
<tr>
<td>A. brasilense</td>
<td>GlnB</td>
<td>MgATP</td>
<td>2OG</td>
<td>3MNY</td>
<td>Truan et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>GlnZ</td>
<td>ADP</td>
<td>DraG</td>
<td>3OST</td>
<td>Rajendran et al. (2011)</td>
</tr>
<tr>
<td>A. fulgidus</td>
<td>GlnK1</td>
<td></td>
<td></td>
<td>3O8W</td>
<td>Litz et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>GlnK2</td>
<td></td>
<td></td>
<td>3NCP</td>
<td>Heffmann et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>GlnK2</td>
<td>MgATP</td>
<td></td>
<td>3NCQ</td>
<td>Heffmann et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>GlnK2</td>
<td>MgADP</td>
<td></td>
<td>3NCR</td>
<td>Heffmann et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>GlnK3</td>
<td>MgATP</td>
<td></td>
<td>3TZ2</td>
<td>Maier et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>GlnK3</td>
<td>MgADP</td>
<td></td>
<td>3TA0</td>
<td>Maier et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>GlnK3</td>
<td>MgATP</td>
<td>2OG</td>
<td>3TA1</td>
<td>Maier et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>GlnK3</td>
<td>MgATP</td>
<td>2OR</td>
<td>3TA2</td>
<td>Maier et al. (2011)</td>
</tr>
<tr>
<td>A. thaliana</td>
<td>P II</td>
<td>Citrate</td>
<td></td>
<td>2O66</td>
<td>Mizuno et al. (2007a)</td>
</tr>
<tr>
<td></td>
<td>P II</td>
<td>Malonate</td>
<td></td>
<td>2O67</td>
<td>Mizuno et al. (2007a)</td>
</tr>
<tr>
<td></td>
<td>P II</td>
<td>MgATP</td>
<td>NAGK</td>
<td>2RD5</td>
<td>Mizuno et al. (2007a)</td>
</tr>
<tr>
<td>A. variabilis</td>
<td>P II</td>
<td>ADP</td>
<td>PipX</td>
<td>3N5B</td>
<td>Zhao et al. (2010b)</td>
</tr>
<tr>
<td>E. coli</td>
<td>GlnB</td>
<td></td>
<td></td>
<td>2P II</td>
<td>Carr et al. (1996)</td>
</tr>
<tr>
<td></td>
<td>GlnB</td>
<td></td>
<td></td>
<td>1PIL</td>
<td>Cheah et al. (1994)</td>
</tr>
<tr>
<td></td>
<td>GlnK</td>
<td>ADP</td>
<td>AmtB</td>
<td>2NS1</td>
<td>Gruswitz et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>GlnK</td>
<td>ADP</td>
<td>AmtB</td>
<td>2NUU</td>
<td>Conroy et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>GlnK</td>
<td>ATP</td>
<td></td>
<td>1GNK</td>
<td>Xu et al. (1998)</td>
</tr>
<tr>
<td></td>
<td>GlnK</td>
<td>ATP</td>
<td></td>
<td>2GNK</td>
<td>Xu et al. (1998)</td>
</tr>
<tr>
<td>H. seropedicae</td>
<td>P II</td>
<td></td>
<td></td>
<td>1HWU</td>
<td>Benelli et al. (2002)</td>
</tr>
<tr>
<td>M. jannaschii</td>
<td>GlnK1</td>
<td>MgATP</td>
<td></td>
<td>2J9C</td>
<td>Yildiz et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>GlnK1</td>
<td>ADP, AMP</td>
<td></td>
<td>2J9D</td>
<td>Yildiz et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>GlnK1</td>
<td>MgATP</td>
<td></td>
<td>2J9E</td>
<td>Yildiz et al. (2007)</td>
</tr>
<tr>
<td>M. tuberculosis</td>
<td>P II</td>
<td>ATP</td>
<td></td>
<td>3LF0</td>
<td>Shetty et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>P II</td>
<td></td>
<td></td>
<td>3BZQ</td>
<td>Shetty et al. (2010)</td>
</tr>
<tr>
<td>N. meningitidis</td>
<td>P II</td>
<td></td>
<td></td>
<td>2GW8</td>
<td>Nichols et al. (2006)</td>
</tr>
<tr>
<td>S. baltica</td>
<td>P II</td>
<td></td>
<td></td>
<td>3CE8</td>
<td>Unpublished</td>
</tr>
<tr>
<td>S. elongatus</td>
<td>P II</td>
<td>MgATP</td>
<td>2OG</td>
<td>2XZW</td>
<td>Fokina et al. (2010a)</td>
</tr>
<tr>
<td></td>
<td>P II</td>
<td>MgATP</td>
<td>2OG</td>
<td>2XUL</td>
<td>Fokina et al. (2010a)</td>
</tr>
<tr>
<td></td>
<td>P II</td>
<td>MgATP</td>
<td></td>
<td>2XBP</td>
<td>Fokina et al. (2010b)</td>
</tr>
<tr>
<td></td>
<td>P II</td>
<td>PipX</td>
<td></td>
<td>2XG8</td>
<td>Llacer et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>P II</td>
<td>NAGK</td>
<td></td>
<td>21J4</td>
<td>Llacer et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>P II</td>
<td>NAGK</td>
<td></td>
<td>2V5H</td>
<td>Llacer et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>P II</td>
<td></td>
<td></td>
<td>1QY7</td>
<td>Xu et al. (2003)</td>
</tr>
<tr>
<td>S. mutans</td>
<td>P II</td>
<td></td>
<td></td>
<td>3L7P</td>
<td>Unpublished</td>
</tr>
<tr>
<td>Synechocystis</td>
<td>P II</td>
<td>ADP</td>
<td></td>
<td>1UL3</td>
<td>Xu et al. (2003)</td>
</tr>
<tr>
<td>T. maritima</td>
<td>GlnK</td>
<td>ADP</td>
<td></td>
<td>1OS1</td>
<td>Schwarzenbacher et al. (2004)</td>
</tr>
<tr>
<td>T. thermophilus</td>
<td>GlnK</td>
<td>ADP</td>
<td></td>
<td>1V9O</td>
<td>Sakai et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>GlnK</td>
<td></td>
<td></td>
<td>1VFJ</td>
<td>Sakai et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>GlnK</td>
<td></td>
<td></td>
<td>1V3R</td>
<td>Sakai et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>GlnK</td>
<td>ATP</td>
<td></td>
<td>1V3S</td>
<td>Sakai et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>GlnK</td>
<td></td>
<td></td>
<td>1UFL</td>
<td>Sakai et al. (2005)</td>
</tr>
</tbody>
</table>
structures. ATP and 2-OG were known to bind cooperatively (Kamberov et al., 1995), but it was also clear from the GlnK-ATP structure that binding of negatively charged 2-OG close to the phosphates of ATP would require the involvement of a divalent cation such as Mg$^{2+}$ (Xu et al., 1998). A resolution of this situation required the identification of the 2-OG binding site and was to take another 12 years from the initial GlnK-ATP structure!

In the meantime, the first indication that ADP might be a physiologically important effector came from the solution of a structure for a P$_{II}$ protein from _Thermatoga maritima_. In this case, ADP was found in the nucleotide-binding pocket, despite the fact that it had not been specifically added during the crystallization (Schwarzenbacher et al., 2004). The protein had been expressed in _E. coli_ grown in selenomethionine-containing minimal medium, and consequently the reason for the presence of bound ADP was not apparent. The T-loop was also disordered in the _T. maritima_ P$_{II}$ structure (Schwarzenbacher et al., 2004) as it was in a later structure of the ADP-bound of _Thermus thermophilus_ GlnK (Sakai et al., 2005).

Clear evidence for a physiological role for ADP came with the solution of the first structure of a P$_{II}$ protein complexed with one of its targets, namely _E. coli_ GlnK bound to the ammonia channel protein AmtB (Conroy et al., 2007). In the GlnK-AmtB complex, each of the nucleotide-binding sites in GlnK contains a molecule of ADP, and as the complex was purified directly from _E. coli_ cells that had been subjected to an ammonia shock the effector status of the complex was considered to be physiologically informative (Conroy et al., 2007). Within the complex (details of which are discussed later), the GlnK T-loop adopts an extended form with two short antiparallel $\beta$-strands separated by a $\beta$-turn causing it to extend 28 Å above the core of the protein (Conroy et al., 2007). The structure of this complex was also solved by Stroud and colleagues, who co-crystallized the GlnK and AmtB proteins, that had been individually purified, together with addition of ATP (Gruswitz et al., 2007).

Comparison of the ATP- and ADP-bound forms of P$_{II}$ reveals that the two nucleotides bind almost identically. The hydrogen bonding interactions with the base and sugar moieties are the same in GlnK-ATP (Xu et al., 1998), GlnB-ATP (Xu et al., 2001) and GlnK-ADP (Conroy et al., 2007). However, the mobility of the side-chains of two highly conserved arginine residues, Arg101 and Arg103 and the very C-terminal segment of the protein (residues 108–112) together with the mobility of the B-loop on the adjacent sub-unit means that the cleft appears to be able to accommodate a variety of nucleotide-binding modes, differing mainly in the conformation and interactions of the di- or triphosphate moiety.

As earlier studies showed that ATP bound tightly to _E. coli_ GlnB in the presence of 2-OG, this led to the conclusion that ATP probably did not play a regulatory role in vivo (Kamberov et al., 1995; Jiang et al., 1998a). However, subsequent studies of nucleotide binding, in _R. rubrum_ P$_{II}$ (Zhang et al., 2001b, 2006b; Ruppert et al., 2002), _Synechocystis_ P$_{II}$ (Ruppert et al., 2002) and _A. thaliana_ P$_{II}$ (Moorhead & Smith, 2003; Smith et al., 2003), led a number of authors to suggest that P$_{II}$ might be able to sense the cellular energy status. Such a model also provided a rationale for the apparent universal nature of the nucleotide-binding site in P$_{II}$ proteins.

This concept led Ninfa and colleagues to re-evaluate ADP binding, and in a series of experiments they demonstrated that ATP and ADP bind competitively to _E. coli_ P$_{II}$ and that whilst ATP binding is co-operative with 2-OG, ADP acted antagonistically to 2-OG (Jiang et al., 2007).
The data suggested that PII could perhaps adopt two conformations, one when the cellular 2-OG levels are high which would favour co-operative binding of 2-OG and ATP, and one when the 2-OG pool was low which might favour ADP binding, although intermediate states with a mixture of ATP- and ADP-bound sites were feasible (Jiang et al., 2007b). Studies of A. thaliana, Synechocystis and A. brasilense PII came to essentially similar conclusions, namely that binding of 2-OG and ATP are synergistic whilst ADP antagonizes 2-OG binding (Smith et al., 2003; Fokina et al., 2010b, 2011; Gerhardt et al., 2012).

The missing element in the effector binding story was the binding site for 2-OG. This was finally resolved with the solution of a structure for A. brasilense GlnZ complexed with MgATP and 2-OG (Truan et al., 2010) and was confirmed by subsequent structures of S. elongatus PII and Archaeoglobus fulgidus GlnK3 each with bound 2-OG in almost identical locations (Fokina et al., 2010a; Maier et al., 2011). The binding site was in excellent agreement with the earlier predictions of Xu et al. (1998), in that 2-OG was bound within the lateral cleft between adjacent subunits, in such a way that both 2-OG and ATP participated in the coordination of a Mg$^{2+}$ ion located between the two effector molecules (Fig. 2a). The 5-carboxy group of 2-OG formed a salt bridge with A. brasilense GlnZ Lys58, and the sixth ligand to the Mg$^{2+}$ ion was provided by the side chain of GlnZ Gln39: both Gln39 and Lys58 being highly conserved PII residues. An earlier structure published by Kühbrandt and colleagues of Methanococcus jannaschii GlnK1 in which they identified a single molecule of 2-OG bound to one of the T-loops (Yildiz et al., 2007) was not consistent with prior knowledge and is unlikely to be biologically relevant.

As might be expected from their phylogeny, the plant PII proteins are quite similar to the prokaryotic GlnB/K proteins. So far, there are only structures for A. thaliana PII, but this is probably quite typical of other plant PII proteins which, in addition to having an N-terminal chloroplast transit peptide that is cleaved off in the mature protein, possess unique approximately 13- and 15-residue N- and C-terminal extensions (Moorhead & Smith, 2003; Mizuno et al., 2007a). The N-terminal extension in plant PII proteins forms numerous interactions with the R2 helix and projects from the surface of the homotrimer opposite to that occupied by the T-loop. In addition, solvent-exposed residues near the T-loop are highly conserved in plants, but differ in prokaryotes. The additional residues at the C-terminus contribute part of the ATP-binding site (see below) and likely participate in an ATP-induced conformational change (Mizuno et al., 2007a).

NifI proteins are particularly distinguished from the GlnB/K proteins by their predicted T-loops; NifI proteins having shorter T-loops (approximately 14 residues) and NifI2 having longer T-loops (approximately 27 residues) than those in B/K proteins (Leigh & Dodsworth, 2007). Studies of NifI proteins to date suggest that NifI1 and NifI2 probably form a hexamer comprising three subunits of each protein, and that the interaction between NifI1 and NifI2 does not involve their T-loops (Dodsworth & Leigh, 2006). Both NifI1 and NifI2, contain the conserved Gln and Lys residues equivalent to Gln39 and Lys58 in the B/K proteins and might therefore be expected to bind 2-OG in an analogous manner to other PII proteins.

**Allosteric regulation of PII by effectors**

Comparison of the structures of A. brasilense GlnZ complexed with MgATP and 2-OG (PDB:3MHY) to that bound to ADP and no 2-OG (PDB:3O5T) reveals very clearly how fluctuation in the intracellular 2-OG pool might be sensed by PII proteins and translated into a conformational change in the T-loop (Truan et al., 2010;
Rajendran et al., 2011). Given that, in most cases the T-loops serve as the interaction surface with the PII target protein, this then offers an insight into how 2-OG could regulate PII interactions with their targets (Fokina et al., 2010a; Truan et al., 2010; Maier et al., 2011).

The key residue in this response is the highly conserved Gln39 which is located directly at the base of the T-loop (Fig. 2). When the cellular 2-OG pool is high, PII is expected to exist predominantly in the MgATP, 2-OG form where Gln39 contributes to the coordination of the Mg\(^{2+}\) ion and the T-loop projects perpendicular from the GlnZ core axis (Truan et al., 2010). As the 2-OG pool drops, the binding affinity for ATP is reduced and that for ADP increases, such that PII switches to the ADP-bound form (Radchenko et al., 2010). In this state, Gln39 is free to adopt an alternative conformation (Fig. 2b) in which it forms a bond with the side chain of Lys58. This very significant movement of Gln39 leads the T-loop to adopt the distinctive extended structure projecting parallel from the PII core axis as seen in the GlnK-AmtB complex (Conroy et al., 2007). Although this model accommodates very well the current data for GlnK and GlnZ, the interaction between 2-OG binding and nucleotide (ATP or ADP) binding is not universal. The PII proteins of Synecococcus and Arabidopsis 2-OG show synergistic effects with both ATP and ADP (Smith et al., 2003; Fokina et al., 2011) and structural data on Synecococcus PII suggests that 2-OG can lead to a conformational change in the T-loop even when ATP remains bound to the active site (Fokina et al., 2010a).

The graded transition between ‘high 2-OG’ and ‘low 2-OG’ states of PII has been proposed to be facilitated by the strong anti-cooperativity observed for 2-OG binding to the three subunits of the protein. Taking into account the known cellular levels of 2-OG and adenylate nucleotides, and the measured binding affinities for these effectors Ninfa and colleagues concluded that in vivo E. coli GlnB trimers will always be saturated with nucleotides and will be bound by at least one molecule of 2-OG under all physiological conditions (Jiang & Ninfa, 2009c). In fact, the structure of the E. coli GlnK-AmtB complex suggests that PII proteins can exist in a form with three ADP molecules and no 2-OG bound (Conroy et al., 2007). Nevertheless, studies of E. coli GlnB showed that when ATP is saturating the binding of 2-OG displays strong anti-cooperativity, such that the first molecule binds with a $K_d$ in the low $\mu$M range and the binding of the second and third molecules occurs with a $K_d$ approximately 1–2 orders of magnitude higher (Kamberov et al., 1995; Jiang et al., 1998a). It was deduced that this anti-cooperative binding of 2-OG facilitated a shallow response to physiological concentrations of 2-OG; a desirable feature for a biological sensor (Jiang & Ninfa, 2009c).

The crystal forms of S. elongatus PII with ATP and 2-OG bound shed further light on the basis for this anti-cooperative binding of 2-OG. Crystallization of PII in the presence of a low (unspecified) concentration of 2-OG produced a structure for which the asymmetric unit contained three PII trimers; one containing three ATP, one $\text{Mg}^{2+}$ and one 2-OG; the second containing three ATP, two $\text{Mg}^{2+}$ and two 2-OG; and the third containing three each of ATP, $\text{Mg}^{2+}$ and 2-OG (Fokina et al., 2010a). Comparison of the monomer conformations in which the 2-OG site is occupied with those which are not, revealed that binding of the first 2-OG generated unequal ligand-binding sites in the adjacent monomers. The conformational changes were slight, but involved distortion of the phosphate moiety of ATP and these stereochemical properties were proposed to account for the altered 2-OG affinities of the different sites (Fokina et al., 2010a).

Binding of the effectors 2-OG, ATP or ADP to PII proteins appears to constitute a fundamental sensory property that is probably shared by nearly all members of the superfamily. However, there may be exceptions as the A. fulgidus GlnK2 protein has been reported not to bind 2-OG, despite apparently having all the key interaction residues (Helfmann et al., 2010) and we presently have no knowledge of the effector-binding properties of the PII-NG group. Nevertheless, a mechanism by which binding of 2-OG, ATP or ADP influences the T-loop conformation, and thereby the ability to interact with one or more targets is likely to characterize most PII proteins. However, in addition to this mechanism a number of PII proteins have a second sensory response characterized by post-translational modification.

**Covalent post-translational modification of PII proteins**

Early studies of the role of PII in regulating GS activity identified modification of E. coli GlnB by uridylylation of residue Tyr51 in the T-loop as a key factor. The PII-UMP form predominated in nitrogen-deficient cells and was rapidly converted to the deuridylylated form by a uridylyl removing (UR) enzyme (GlnD) when the nitrogen status of the cells improved (Francis & Engleman, 1978). This process is reversible and re-uridylylation of Tyr51 is catalysed by the uridylyltransferase (UTase) activity of the same protein, GlnD. GlnD has been characterized from E. coli (Jiang et al., 1998a; Zhang et al., 2010), R. rubrum (Jonsen & Nordlund, 2007; Zhang et al., 2010), Herbaspirillum seropedicae (Bonatto et al., 2007, 2012) and A. brasiliense (Araujo et al., 2008). It has at least four domains of which the N-terminal domain encodes the UTase activity and the adjacent HD domain encodes the UR activity (Zhang et al., 2010). Hence, the two activities are not thought to...
share an active site as previously proposed (Jiang et al., 1998a). GlnD has a single glutamine binding site such that glutamine binding stimulates UR activity, and in the absence of glutamine UTase activity predominates.

The glnD gene is restricted to the Proteobacteria, the Actinobacteria and a number of other poorly characterized prokaryotes, and to date the PII proteins are the only known substrate for GlnD. However, studies in the alfalfa symbiont Sinorhizobium meliloti have shown that whereas a glnD mutant is ineffective, i.e. it can fix nitrogen but the alfalfa is unable to use that nitrogen for growth, a glnB, glnK double mutant forms an effective symbiosis (Yurgel & Kahn, 2008; Yurgel et al., 2010; Yurgel et al., 2011). This suggests that GlnD may have some other role in S. meliloti, although no other GlnD substrate protein has so far been found (Yurgel et al., 2011).

PII uridylylation is by no means universal, and in the Actinobacteria Streptomyces coelicolor (Hesketh et al., 2002) and Corynebacterium glutamicum (Stroser et al., 2004) GlnD adenylylates, rather than uridylylates, Tyr51 in nitrogen-deficient conditions. Interestingly, in the Actinobacteria, glnD is part of the amtB, glnK, glnD gene cluster, whereas in the Proteobacteria it is usually encoded elsewhere on the genome.

A discreet form of post-translational modification is seen in some Cyanobacteria, including S. elongatus (Forchhammer & Tandeau de Marsac, 1994, 1995) and Synechocystis (Kloft et al., 2005), where the PII T-loop is phosphorylated on residue Ser49. Just as in the Proteobacteria, PII modification occurs in response to nitrogen starvation (Forchhammer & Tandeau de Marsac, 1994), however despite considerable efforts, the kinase responsible is yet to be identified (Irmler et al., 1997). The kinase requires a high 2-OG concentration (Forchhammer & Tandeau de Marsac, 1995) which may well imply a specific conformation of the PII T-loop. Dephosphorylation of S. elongatus PII is driven by a specific phosphatase, PphA, the activity of which is inhibited by 2-OG in concert with Mg-ATP (Irmler & Forchhammer, 2001; Ruppert et al., 2002). Despite conservation of Ser49 in the T-loop, this phosphorylation is not found in Prochlorococcus (Palinska et al., 2002) or in Anabaena (Zhang et al., 2007) nor has it been detected in plants (Smith et al., 2004). In Anabaena another novel modification, namely nitration of Tyr51 has been reported (Zhang et al., 2007).

Despite these various types of post-translational modification, a great many organisms appear not to modify PII. There is, for example, no report of PII modifications in Archaea, in Firmicibacteria such as Bacillus subtilis, in some Cyanobacteria and in plants. Hence, although post-translational modification clearly plays an important part in regulation of some PII targets (see later) it is not necessary in all cases.

**The sensory functions of PII proteins?**

PII proteins have long been recognized as key regulators of the cellular nitrogen status (N-status) and indeed they were discovered through research in this area. Initial work focused on the process of GlnB uridylylation and suggested that the UTase/UR protein (GlnD) responded to the intracellular glutamine/2-OG ratio as the key indicator of N-status (Rhee et al., 1985a). However, subsequent studies showed that GlnD responds specifically to the glutamine pool and PII is the target for 2-OG binding (Kamberov et al., 1995), leading to the revised view that both molecules provided independent inputs into the sensory machinery. Furthermore, 2-OG was considered to have the potential to signal cellular carbon sufficiency as well as nitrogen deficiency, and hence the concept emerged of PII proteins sensing the cellular C/N ratio by integrating a nitrogen signal (glutamine) with a carbon signal (2-OG; Ninfa & Atkinson, 2000; Ninfa & Jiang, 2005; Commichau et al., 2006). However, this model cannot apply in organisms that lack a glutamine sensor, e.g. GlnD, and hence a generic model still requires 2-OG to act as a signal of cellular N-status.

Measurements of cellular 2-OG pools in conditions of nitrogen sufficiency or deficiency support the view that this molecule is well suited to report N-status. In E. coli, 2-OG pools range between ~1.5 and 0.1 mM in response to low and high nitrogen nutrition (Senior, 1975; Reyes-Ramirez et al., 2001; Yuan et al., 2009; Radchenko et al., 2010), and in the archaean M. maripaludis the equivalent range is 0.8–0.08 mM (Dodsworth et al., 2005; Leigh & Dodsworth, 2007). Furthermore, the 2-OG pool responds within minutes to a change in extracellular nitrogen availability (Dodsworth et al., 2005; Yuan et al., 2009; Radchenko et al., 2010) and the 2-OG pool has been estimated to have a half-life of 0.5 s (Yan et al., 2011). Hence, it seems likely that PII alone can and does sense cellular N-status via the 2-OG pool and the refinement of glutamine sensing has evolved later.

So, can PII proteins also act as carbon sensors? The recent recognition of the role (described in detail later) of the chloroplastic PII in A. thaliana in regulating the activity of plastidial acetyl-CoA carboxylase (ACCase) suggests that they may be able to (Feria Bourrellier et al., 2010). PII inhibits ACCase activity and this inhibition is relieved by 2-OG, oxaloacetate (OAA) or pyruvate, in a manner which is PII-dependent (Feria Bourrellier et al., 2010). Previous studies of E. coli GlnB indicated that it might also bind OAA or pyruvate, albeit much less effectively than 2-OG (Kamberov et al., 1995), and so the potential may exist for some element of carbon status sensing by PII proteins through binding of other small molecule effectors as well as 2-OG.
Although a number of studies mentioned earlier have led to a perception that PII proteins may function in vivo to sense and respond to fluctuations in the cellular energy charge, there is presently very little in vivo data to support this proposition. The energy charge is primarily a reflection of the ATP/ADP ratio and, in E. coli at least, this is apparently quite strongly buffered around a value of 0.9 or greater (Chapman et al., 1971; Bennett et al., 2009). Attempts have been made to manipulate the cellular ATP levels in R. rubrum, but even quite dramatic reductions in the ATP pool had little effect on PII-mediated regulation of the activities of the transcription factor NifA or the enzyme nitrogenase (Zhang et al., 2009). In E. coli, we did observe a transient fluctuation in the ATP/ADP ratio following an extracellular ammonium shock, but the cause of this is presently unknown and we cannot estimate what contribution, if any, this made to the response of GlnK in binding to AmtB (Radchenko et al., 2010). So, to date, the role of PII proteins in directly sensing the cellular energy charge remains to be proven.

Protein targets regulated by interaction with PII

As discussed in the Introduction to this review, PII proteins were discovered through studies of the regulation of GS activity, but it has subsequently become apparent that they play a role in almost every facet of nitrogen metabolism in prokaryotes. Although the ancestral GlnK protein appears to have evolved in concert with the ammonia channel protein AmtB, its inherent ability to respond conformationally to the cellular N-status has led to it being co-opted to control other membrane transport systems, a variety of enzymes and a wide range of transcription factors. The list of PII targets has grown over the years and is almost certainly not yet complete. In the sections below, we have summarized current knowledge on the remarkable versatility of PII and its consequently pivotal role in orchestrating cellular nitrogen metabolism.

Regulation of membrane transport proteins

Recognition that the genetic linkage between glnK and the ammonia channel gene amtB was highly conserved strongly suggested that the two proteins were functionally related (Thomas et al., 2000) and indeed this is presently one of the best described of all PII-target interactions. The glnK gene can be located upstream or downstream of amtB and, as mentioned earlier, within the Actinobacteria it is also linked to glnD.

The Amt family is conserved throughout all domains of life. Homologues of E. coli AmtB are found in Bacteria, Archaea, fungi, and plants (von Wirén & Merrick, 2004) and in the animal kingdom, Amt proteins are represented by the closely-related Rhesus (Rh) proteins (Andrade & Einsle, 2007). All Amt and Rh proteins are homotrimers in which each subunit is comprised of 11 or 12 transmembrane helices that surround a single conduction channel. Escherichia coli AmtB is the best characterized Amt protein and it structure was solved in 2004 (Khamedi et al., 2004; Zheng et al., 2004). Although the precise mechanism of action of Amt proteins is still unresolved (Javelle et al., 2007; Lamoureux et al., 2010), they are thought to be unidirectional channels that function to scavenge ammonium from the extracellular medium. Consistent with this hypothesis the glnKamtB operon is transcriptionally regulated in most organisms such that it is only expressed under N-limitation.

The role of GlnK is to physically regulate the flux of ammonia through AmtB such that ammonia uptake is closely linked to the intracellular N-status. Amt proteins are not essential for growth on ammonium as sole nitrogen source, and it is commonly assumed that this is because in conditions of high ammonium supply sufficient NH3 can diffuse across the cell membrane. Initial studies demonstrated that E. coli GlnK was cytoplasmically located under conditions of N-limitation, but was rapidly membrane-associated following an ammonium shock, and that this membrane association was strictly AmtB-dependent (Coutts et al., 2002). Evidence for GlnK-AmtB complex formation has subsequently been observed in Gram-negative, Gram-positive bacteria and Archaea. These organisms include A. brasilienese, C. glutamicum, H. seropedicae, R. rubrum and Rhodobacter capsulatus (Strosser et al., 2004; Huergo et al., 2006b; Wolfe et al., 2007; Teixeira et al., 2008; Huergo et al., 2010) as well as B. subtilis and M. jannaschii (Detsch & Stulke, 2003; Yildiz et al., 2007). It has also been postulated to occur in A. fulgidus (Andrade et al., 2005).

The precise function of complex formation was resolved by the solution of the crystal structure of the complex which was achieved simultaneously in two labs (Conroy et al., 2007; Gruswitz et al., 2007). The interface between the two proteins is a very open one with a relatively small surface area buried between the two proteins (Fig. 3a). The docking interaction is formed almost exclusively by the T-loops of GlnK which, adopt an extended β-turn. These loops act as extended pin, the tips of which insert deeply into the cytoplasmic pore exits. The novel T-loop conformation places a conserved residue (Arg47) at the tip of the loop such that the guanidinium group of Arg47 constricts the exit of the pore, and thereby blocks any possible movement of ammonia through the channel (Fig. 3b). This structure reinforces the concept that PII
proteins co-evolved with an ancestral Amt protein, and that the nature of the T-loops was originally determined by the requirement to control flux through the channel. The structure of the complex also offered significant insights into how the interaction of GlnK and AmtB might be regulated. It was clear that deuridylylation of GlnK would be a pre-requisite for complex formation, as the position of the Tyr51 residue meant that the uridylylated form could not bind to AmtB. This prediction has been experimentally confirmed in vitro (Rodrigues et al., 2011). Likewise, the location of the T-loops suggests that GlnK cannot be modified when bound to AmtB suggesting that dissociation could not be driven by GlnK uridylylation. Indeed, this is consistent with occurrence of GlnK-AmtB complex formation in organisms in which GlnK is apparently not subject to post-translational modification of the T-loops. Furthermore, we have shown that a Tyr51Phe variant of E. coli GlnK, that cannot be uridylylated, undergoes binding to and disengagement from AmtB in response to the cellular N-status in a very similar way to wild-type GlnK (Am Radchenko and M. Merrick, unpublished). An earlier report that contradicts this observation was erroneous because of the presence of a second mutation within the glnK Y51F allele (Javelle et al., 2004; Durand & Merrick, 2006).

The second key observation from the complex structure was the presence of ADP in each of the three effector binding sites of GlnK. This was the first clear indication that ADP binding to PII proteins was physiologically important. Together with the suggestion that uridylylation is not a pivotal process it implicated 2-OG, ATP and ADP as the likely key molecules that could affect the T-loop conformation and control complex association and dissociation. Subsequent in vivo measurements of these key effector pools and their correlation with the interaction of GlnK and AmtB supports this hypothesis (Radchenko et al., 2010).

Hence, the current model for E. coli is that under N-deficient conditions GlnK is cytoplasmically located, it is uridylylated and each trimer is bound by three molecules of MgATP and 2-OG (Radchenko et al., 2010). An increase in the extracellular availability of ammonium results initially in ammonia influx through AmtB which leads to a rapid rise in the glutamine pool and a reciprocal drop in the 2-OG pool. GlnK is deuridylylated owing to the effect of the glutamine pool on GlnD activity. 2-OG is simultaneously lost from GlnK and there is also a transient rise in the ADP pool and a drop in ATP (of unknown origin). This facilitates replacement of the ATP in GlnK by ADP with consequent binding of GlnK to AmtB. As the cellular N-status then drops again, dissociation of the complex is driven by the concomitant rise in the 2-OG pool. Replacement of ADP by MgATP and 2-OG causes a change in the conformation of the T-loops and disengagement from AmtB at which point the T-loops are substrates for uridylylation by GlnD.

There are certainly exceptions to this model. Firstly, in B. subtilis although the association of GlnK with the cell membrane is strictly AmtB-dependent it is apparently not dependent on the cellular N-status (Detsch & Stulke, 2003). Secondly, in some rare cases where glnK is located immediately downstream of amtB, a gene fusion has occurred so as to encode a single AmtB-GlnK peptide. Interogation of the SMART protein domain database (Letunic et al., 2011) identifies such chimeras in three Archaea Natronomonas pharaonis, Halorubrum lacusprofundi and Haloterrigena turkmenica, and in 46 bacterial species predominantly in the genera Clostridium, Eubacterium and Ruminococcus. None of these systems has yet been studied.

Although AmtB is the predominant membrane protein target for PII proteins in prokaryotes, it is not the only one. Studies of DraT-independent nitrogenase inactivation in Azotobacter (see later) have shown that in this organism, GlnK can be sequestered to the cell membrane by the Rnf membrane complex which catalyses electron

Fig. 3. Structure of the Escherichia coli AmtB-GlnK complex (PDB:2NUU). (a) The GlnK trimer (cartoon representation) occupied by ADP (yellow sticks) interacts with the cytoplasmic surface of the AmtB trimer (space filling). (b) The T-loops of GlnK (red cartoon) penetrate deeply into the cytoplasmic end of the AmtB ammonia conduction pores and residue Arg47 (space filling) at the tip of the T-loop effectively blocks substrate flow.
Regulation of enzyme activity by PII proteins

PII proteins play a significant role in regulating the activity of a number of enzymes, either directly or indirectly. The most well-known target, albeit an indirect one, is probably GS, but the influence of PII proteins extends into many other areas of N metabolism, and new enzyme targets continue to be identified.

Glutamine synthetase

The most common form of GS is GSI: a homo-dodecamer formed by two hexameric rings with 12 active sites, and encoded by glnA (Eisenberg et al., 2000). GSI is strictly regulated at the transcriptional and the post-translational level. Expression of glnA is usually enhanced in response to nitrogen starvation through the activity of a global nitrogen regulator (Merrick & Edwards, 1995; Reitzer, 2003; Amon et al., 2010). In most prokaryotes, PII proteins regulate both GS activity and the activity of these global transcription factors (see later).

In many bacteria, GS can be inactivated by reversible adenylation of a conserved tyrosine residue in each subunit (Reitzer, 2003). A bi-functional enzyme, namely ATase or GlnE, is responsible for both adenylation and deadenylation. The expression of glnE in E. coli is constitutive and glnE orthologues are present in most Proteobacteria, but not in the Epsilonproteobacteria, nor in Actinobacteria (Rhee et al., 1985b; van Heeswijk et al., 1993; Amon et al., 2010). When N-limited E. coli cells face a sudden increase in extracellular ammonium, GS is rapidly adenylylated. This GS inactivation is presumed to be necessary to avoid a drastic decrease in the intracellular pools of ATP and glutamate (Kustu et al., 1984; Zhang et al., 2006a; Yan, 2007).

Most data on GlnE have been derived in E. coli. The two activities of GlnE (adenyltransferase – AT, and adenyly-removing – AR) are located in two different, but homologous nucleotidyltransferase domains (NT domains) separated by a central linker. The N-terminal domain has AR activity and the C-terminal domain has AT activity (Jaggi et al., 1997; Clancy et al., 2007; Jiang & Ninfa, 2007). The structures of the separate NT domains have been solved, and a model for the complete ATase and a docking site for GS was proposed (Xu et al., 2004, 2010). Although the isolated AR domain is not active in vitro, AR activity is restored in the presence of a separated AT domain. The reconstituted GlnE is similar to the wild-type enzyme suggesting that the GlnE domains communicate and regulate each others’ activities (Jiang & Ninfa, 2007, 2009b; Jiang et al., 2007a).

In E. coli, both PII proteins GlnB and GlnK participate in the regulation of GlnE, but GlnB seems to be more effective than GlnK in vitro (van Heeswijk et al., 1996, 2000; Atkinson & Ninfa, 1998, 1999). In N-limitation, GlnB-UMP₃, interacts with the AT domain which results in stimulation of AR activity and activation of GS. Conversely, in N-excess, non-modified GlnB interacts with the AR domain thereby stimulating AT activity and GS inactivation (Ninfa & Atkinson, 2000; Reitzer, 2003). Glutamine binds to the AT domain causing two effects: stabilization of the GlnB and GlnE complex (Jiang & Ninfa, 2007) and inhibition of GlnB-UMP₃ binding to GlnE (Jiang et al., 2007a). Interactions of GlnB with one or other GlnE domain are proposed to be communicated to the other antagonistic domain via domain–domain interaction mediated by the central linker region (Jiang & Ninfa, 2007, 2009b; Jiang et al., 2007a) and this balance of the two activities is proposed to optimize steady state growth (Okano et al., 2010).

As expected, GlnB regulation of AT activity is controlled by effectors such that under N-limitation, when glutamine levels are low, GlnB is saturated with MgATP and 2-OG and is fully uridylylated. GlnB-UMP₃ then promotes AR and inhibits AT activity, thereby maximizing GS activity. In N-sufficiency, glutamine levels rise and 2-OG levels drop, GlnB is de-uridylylated and presumably binds little, if any, 2-OG and predominantly ADP. In this form, both GlnB and glutamine stimulate AT and inhibit AR activity resulting in GS inactivation. It has also been proposed that this regulatory system might respond to the cellular energy levels. Low ATP/ADP ratios stimulate GlnB binding and inhibit GlnB-UMP₃ binding to GlnE, therefore favouring GS inactivation by adenylation under low energy conditions (Jiang et al., 1998b, c, 2007a, b; Jiang & Ninfa, 2009c).

The GlnB T-loop region is required for interaction with GlnE (Jiang et al., 1997b; Jiang & Ninfa, 2007) and only one T-loop within the trimer is required, because heterotrimers with one wild-type subunit and two T-loop...
deleted subunits can regulate GlnE (Jiang et al., 1997c). There are however conflicting results on the number and the location of the GlnB-binding sites on GlnE. Some authors suggest that a single site in the AR domain can bind GlnB and GlnB-UMP, competitively (Jaggi et al., 1997); others suggest two different sites for GlnB and GlnB-UMP, one in each domain (Jiang & Ninfa, 2007; Jiang et al., 2007a); whereas another study suggested that GlnB binds near the central linker (Clancy et al., 2007), a hypothesis supported by docking analysis between the GlnB and GlnE structures (Xu et al., 2010).

ATase has been studied in a number of other model systems as well as E. coli. In R. rubrum, the architecture of GlnE is unusual: it has a C-terminal extension with peroxiredoxin activity which protects GS from reactive oxygen species in vitro. One or more PII proteins has been implicated in ATase regulation in many other organisms, including R. rubrum (Zhang et al., 2001b, 2006a; Jonsson & Nordlund, 2007; Jonsson et al., 2008); S. meliloti (Arcondé guy et al., 1999; Jiang et al., 2001; Rudnick et al., 2002; He et al., 2008); Pseudomonas stutzeri (He et al., 2008) and H. seropedicae (Benelli et al., 1997; Persuhn et al., 2000). In A. brasilense, neither GlnB nor GlnZ significantly affects GS adenylylation in vivo (de Zamaroczy, 1998; Huergo et al., 2006b) and hence, as in other organisms where PII is not involved, it is possible that GlnE is solely regulated by glutamine. In the Actinobacteria, S. coelicolor and C. glutamicum GlnK does not control GlnE activity, and it is unclear how GlnE perceives the cellular N-status (Nolden et al., 2001a; Hesketh et al., 2002; Strosser et al., 2004). In M. tuberculosis, glnE is an essential gene and the possible role of GlnK in GlnE regulation has not yet been investigated (Parish & Stoker, 2000; Carroll et al., 2008).

Although archaea does not encode a GlnE protein, PII still plays a role in regulation of GS activity. In the methanogen Methanosarcina mazei, in vitro studies indicate that in low 2-OG levels, e.g. upon ammonium shock, GlnK1 interacts with GS thereby inhibiting its activity. However, high 2-OG levels, as found in N-limited cells promote M. mazei GS activation presumably by direct binding to the enzyme (Ehlers et al., 2005). By complete contrast, in the halophilic Haloferax mediterranei GlnK1 and GlnK2 have been reported to activate GS in vitro (Pedro-Roig et al., 2011).

Post-translational control of nitrogenase activity

In a number of diazotrophs, the activity of nitrogenase is tightly controlled at the post-translational level. This control occurs by a number of different mechanisms, but all of those that have been described in detail to date involves PII proteins (Huergo et al., 2012).

ADP-ribosylation of nitrogenase

In some proteobacterial species, nitrogenase activity is regulated by reversible ADP-ribosylation of the NifH component of nitrogenase, and this is best described in the Alphaproteobacteria A. brasilense and R. rubrum. Addition of ammonium to nitrogen-fixing cultures of these organisms promotes NifH ADP-ribosylation, and hence nitrogenase inactivation (Zhang et al., 1997; Nordlund, 2000). Removal of ADP-ribose occurs after consumption of the ammonium with the consequent restoration of nitrogenase activity.

NifH ADP-ribosylation is catalyzed by Dinitrogenase Reductase ADP-Ribosyl Transferase (DraT), and the process is reversed by Dinitrogenase Reductase ADP-Ribosyl Glycohydrodase (DraG). The two enzymes are subject to opposing regulation by PII proteins (Zhang et al., 1997; Nordlund, 2000) and the first evidence for this was reported in R. capsulatus, where a glnB mutant showed defective ammonium-induced nitrogenase inactivation (Hallenbeck, 1992). Several other studies in A. brasilense (Klassen et al., 2001, 2005), Azorarcus sp. (Martin & Reinhold-Hurek, 2002), R. rubrum (Zhang et al., 2000, 2001a, 2005), R. capsulatus (Drepper et al., 2003; Tremblay et al., 2007) and Rhodopseudomonas palustris (Heiniger et al., 2011) implicated PII proteins and their modification by GlnD in the regulation of NifH-ADP-ribosylation. Furthermore, direct interactions between DraT and GlnB were identified in R. capsulatus (Pawlowski et al., 2003), in R. rubrum (Zhu et al., 2006) and in A. brasilense (Huergo et al., 2006a, 2009). Likewise, interaction between DraG and GlnZ (the glnK gene is named glnZ in A. brasilense) was reported in A. brasilense (Huergo et al., 2007, 2009).

The interaction between DraT and GlnB in A. brasilense is regulated by the cellular nitrogen levels in vitro and by the PII effector molecules ATP, ADP and 2-OG in vitro. Under nitrogen-fixing conditions, GlnB is fully uridylylated and does not interact with DraT at detectable levels. Upon an ammonium shock, GlnB is de-uridylylated and interacts with DraT at the same time that NifH is ADP-ribosylated, suggesting that DraT-GlnB interaction activates DraT (Huergo et al., 2006a). The stoichiometry of the DraT-GlnB complex is 1 DraT monomer: 1 GlnB trimer, and the complex is stabilized by ADP and readily disrupted by MgATP and 2-OG. Hence, the interaction would be favoured after an ammonium shock when the drop in the cellular 2-OG level would facilitate ADP binding to GlnB (Huergo et al., 2009). Activation of DraT by GlnB has been confirmed in vitro (V.R. Moure et al., 2002).
and L.C. Seefeldt, personal communication) and is corroborated by the fact that glnB mutants of A. brasilense and R. capsulatus lack ammonium-induced nitrogenase inactivation in vivo (Drepper et al., 2003; Klassen et al., 2005; Huergo et al., 2006b). Strikingly, in R. rubrum, a glnB mutant had normal nitrogenase regulation, and altered regulation was only observed in a glnB, glnJ double mutant suggesting that both PII homologues can activate DraT in R. rubrum (Zhang et al., 2001b).

In A. brasilense, the DraG-GlnZ complex is stabilized in vitro by ADP and disrupted by MgATP and 2-OG. Uridylylated GlnZ interacts with DraG though the interaction is weaker than with de-uridylylated GlnZ (Huergo et al., 2009). Hence, the DraG-GlnZ complex should only be stable after an ammonium shock because of GlnZ de-uridylylation and the drop in intracellular 2-OG (Huergo et al., 2009).

Nitrogenase ADP-ribosylation did not occur in amtB mutants of R. capsulatus (Yakunin & Hallenbeck, 2002), R. rubrum (Wang et al., 2005) and A. brasilense (Huergo et al., 2006b) despite the mutants being unaffected in ammonium uptake, suggesting that AmtB might regulate DraT and/or DraG activities. Studies in A. brasilense showed that after ammonium addition, GlnZ is de-uridylylated and the DraG-GlnZ complex binds to the membrane in an AmtB-dependent manner (Huergo et al., 2006b). A ternary complex between AmtB-GlnZ-DraG could be reconstituted in vitro in the presence of ADP, suggesting that DraG interaction with the AmtB-GlnZ complex in the membrane results in DraG inactivation (Huergo et al., 2007). The same mechanism is likely to operate in R. rubrum, where DraG associates with the membrane in the presence of ammonium and both DraG membrane-binding and NifH ADP-ribosylation require AmtB1 and GlnJ or GlnB (Wang et al., 2005; Zhang et al., 2006b).

The structure of DraG has been solved for A. brasilense and R. rubrum (Berthold et al., 2009; Li et al., 2009), as has the structure of the A. brasilense GlnZ-DraG complex (Rajendran et al., 2011). In contrast to other PII complexes, the majority of contacts in the GlnZ-DraG complex do not involve the PI T-loop region. Instead, DraG binds in the interface between the ADP-bound monomers of GlnZ with up to three DraG monomers bound to one GlnZ trimer (Fig. 4a). Superposition of the GlnZ MgATP 2-OG structure with the GlnZ ADP structure in the DraG-GlnZ complex readily explains why ADP is required for complex stabilization. With MgATP 2-OG, the GlnZ T-loops are extended at nearly 90° to the GlnZ trimer axis and this conformation would clash with the DraG surface. By contrast, in the GlnZ ADP structure, the T-loops appear to extend nearly parallel to the GlnZ trimer axis, as observed for E. coli GlnK ADP (Conroy et al., 2007). This conformation is spatially compatible with DraG interaction and allows modelling of a putative structure for the AmtB-GlnZ-DraG ternary complex (Fig. 4b; Rajendran et al., 2011).

A docking model between the predicted structure of ADP-ribosylated NifH and the DraG-GlnZ complex suggests that steric hindrance could inactivate DraG upon GlnZ binding (Rajendran et al., 2011). However, in an amtB mutant, the DraG-GlnZ complex is still formed after an ammonium shock, despite the lack of DraG inac-

Fig. 4. Structure of the Azospirillum brasilense DraG-GlnZ complex (PDB:3O5T). (a) The A. brasilense GlnZ-DraG complex (viewed from the top surface of GlnZ) in which three DraG monomers (space filling) are complexed to a single GlnZ trimer (cartoon). (b) Model of the DraG-GlnZ-AmtB ternary complex in which the structure of the A. brasilense GlnZ-DraG complex has been morphed with the Escherichia coli GlnK-AmtB complex.
Nitrogenase regulation by NifI proteins

Nitrogenase post-translational inactivation in response to ammonium has also been studied in detail in *M. maripaludis*. This organism encodes five P\textsubscript{II} proteins, two of which belong to the Nif subfamily and their respective genes, *nifI*\textsubscript{1} and *nifI*\textsubscript{2}, are located between the nitrogenase structural genes *nifH* and *nifD* (Kessler & Leigh, 1999). As *nifI* genes are found in the same genomic context in a number of nitrogen-fixing *Archaea* and in some *Proteobacteria*, *Firmicutes*, *Chlorobi* and *Chloroflexi* species, it seems very likely that they have a similar role in all those organisms.

Deletion of *M. maripaludis* *nifI*\textsubscript{1} or *nifI*\textsubscript{2} abolishes nitrogenase inactivation by ammonium in vivo, and results in higher nitrogenase activities in cell-free extracts (Kessler et al., 2001; Dodsworth et al., 2005). Nitrogenase activity in extracts from wild-type cells, but not from a Δ*nifI\textsubscript{1},nifI\textsubscript{2}* strain, was stimulated by 2-OG. As in the *Proteobacteria*, intracellular 2-OG levels drop significantly after an ammonium shock and it was suggested that NifI could sense this reduction and inactivate nitrogenase accordingly in *M. maripaludis* (Dodsworth et al., 2005).

NifI\textsubscript{1} and NifI\textsubscript{2} bind to each other forming a hetero-oligomer, probably a hexamer, which further oligomerizes to a putative dodecamer in the presence of 2-OG. The NifI\textsubscript{1,2} heteromer physically interacts with the nitrogenase NifDK component causing inhibition of nitrogenase activity in vitro, and both the interaction and nitrogenase inhibition are relieved by 2-OG in the presence of MgATP (Dodsworth et al., 2005). Nitrogenase inhibition is because of competition between NifI\textsubscript{1,2} and NifH to bind NifDK. During the nitrogenase catalytic cycle, NifH reversibly binds and transfers electrons to NifDK, but in the absence of 2-OG NifI\textsubscript{1,2} binds tightly to NifDK occluding the interaction site for NifH and blocking electron transfer (Dodsworth & Leigh, 2007). These data suggest that upon an ammonium shock the intracellular 2-OG level is reduced, NifI\textsubscript{1,2} de-oligomerizes and interacts tightly with NifDK inhibiting its activity by blocking the docking site for NifH. When the ammonium is exhausted by cellular metabolism, the 2-OG levels rise promoting NifI\textsubscript{1,2} oligomerization, and this change in the NifI\textsubscript{1,2} quaternary structure relieves interaction with NifDK.

**Nitrogenase regulation by inactivation of the Rnf1 complex**

In *Azoarcus* sp., nitrogenase inactivation occurs independently of NifH ADP-ribosylation, despite the presence of *draT* genes, as it occurs in *draT* mutants (Oetjen & Reinhold-Hurek, 2009). Furthermore, knockout of *amtB* or *glnK* results in only partial nitrogenase inactivation after ammonium addition (Martin & Reinhold-Hurek, 2002; Noindorf et al., 2006, 2011). Recent studies of this DraT-independent nitrogenase switch-off has shown that ammonium upshift causes GlnK (but not GlnB) to be sequestered to the cell membrane, and that this is dependent on the Rnf1 complex (Sarkar et al., 2012). The Rnf membrane complex catalyses electron transfer coupled to ion transport across the membrane (Biegel et al., 2011) and ammonium upshift leads to decreased Rnf1 activity, potentially as a consequence of GlnK binding (Sarkar et al., 2012). This could lead to a rapid interruption of electron flow to nitrogenase and the observed DraT-independent nitrogenase switch-off.

**Nitrogenase regulation by as-yet unidentified mechanisms**

Reversible nitrogenase inactivation by ammonium, independent of NifH ADP-ribosylation, has also been described...
in *H. seropedicae*, where neither DraT/DraG nor Rnf are present (Fu & Burris, 1989; Pedrosa et al., 2011). In *H. seropedicae*, knockout of *amtB* or *glnK* resulted in only partial nitrogenase inactivation after ammonium addition (Noindorf et al., 2006, 2011) although GlnK is sequestered to the cell membrane by AmtB a few minutes after ammonium addition and so nitrogenase inactivation may involve formation of an AmtB-GlnK complex (Huergo et al., 2010).

**Control of arginine biosynthesis via NAGK**

NAGK catalyses the committed step in arginine biosynthesis, and in organisms having a cyclic arginine pathway, NAGK may be feedback inhibited by arginine (Caldovic & Tuchman, 2003). Arginine is a key nitrogen storage compound in cyanobacteria and plants (Uhrig et al., 2009), and therefore NAGK must be released from inhibition to accumulate arginine when nitrogen is abundant. NAGK has been identified as a PII target in cyanobacteria (Heinrich et al., 2004), in *A. thaliana* (Burillo et al., 2004) and in *Oryza sativa* (Sugiyama et al., 2004), and it co-localizes with PII in the plant chloroplast (Sugiyama et al., 2004; Chen et al., 2006). In rice, two NAGK genes, NAGK1 and NAGK2, were identified only one of which, NAGK1, interacted with PII.

The interactions between NAGK and PII in *S. elongatus* and *A. thaliana* were confirmed in vitro (Heinrich et al., 2004; Chen et al., 2006; Feria Bourrellier et al., 2009). NAGK activity is stimulated by PII in a concentration-dependent manner (Maheswaran et al., 2004; Chen et al., 2006; Ferrario-Mery et al., 2006), and PII both increases the catalytic efficiency of NAGK and relieves arginine feedback inhibition (Maheswaran et al., 2004; Chen et al., 2006; Beez et al., 2009). Accordingly, NAGK activity is reduced in a *S. elongatus* *glnB* mutant (Burillo et al., 2004; Heinrich et al., 2004) and *Arabidopsis* PII knock-outs accumulate less arginine than wild-type plants in response to ammonium supply, despite unaltered expression of genes involved in arginine biosynthesis (Ferrario-Mery et al., 2006).

Arginine-sensitive NAGK is a hexameric ring formed by a trimer of dimers. The dimer structure resembles that of the *E. coli* arginine-insensitive NAGK, and the hexamer is formed by linking three *E. coli*-like dimers connected by a flexible N-terminal helix found only in arginine-sensitive NAGK (Llacer et al., 2008). Each NAGK subunit has two domains; the N-terminal containing the NAGK binding site and the C-terminal containing the ATP binding site. Arginine feedback inhibition is achieved through arginine binding next to the interdimeric junction widening the hexamer ring and promoting the separation of the N-acetyl-l-glutamate and ATP binding sites (Llacer et al., 2008).

**Fig. 5.** Structure of the *Synechococcus elongatus* NAGK–PII complex (PDB:2VSH). One NAGK hexamer (space filling) is sandwiched by two PII trimers (cartoons). The PII T and B loops are indicated by arrows.

Structures of the PII–NAGK complex from *S. elongatus* and *A. thaliana* have been solved (Llacer et al., 2007; Mizuno et al., 2007b), and in both cases one NAGK hexameric ring is sandwiched by two PII trimers, each PII subunit interacting with one NAGK subunit (Fig. 5). The PII B-loop and the β1α1 region contact the NAGK C-terminal domain, whereas the T-loop assumes a compact structure and is inserted in the NAGK interdomain region (Llacer et al., 2007, 2008; Mizuno et al., 2007b). In the *S. elongatus* complex, no PII effector molecule was observed in the crystals, however, in *A. thaliana*, PII was occupied by MgATP (Llacer et al., 2007; Mizuno et al., 2007b). Comparison of *S. elongatus* NAGK–PII with the free NAGK from *T. maritima* revealed that PII binding alters the position of N and C-terminal NAGK domains, the hexamer becomes less wide and the arginine binding site is expanded explaining the increased resistance to arginine inhibition (Llacer et al., 2007, 2008).

Cyanobacteria and plant PII proteins are mutually interchangeable to regulate the corresponding NAGK activity (Beez et al., 2009). The *S. elongatus* PII residues Arg45, Ser49 and Glu85, and the corresponding residues in *A. thaliana* PII (Arg56, Ser60 and Glu97) make important contacts with NAGK and were proposed to be signature residues for PII proteins that regulate NAGK. Mutations in any of these residues prevent interaction with NAGK in vitro (Burillo et al., 2004; Llacer et al., 2007; Beez et al., 2009). In *S. elongatus* PII, Ser49 is phosphorylated when 2-OG levels are high, and maximum NAGK activities in vivo were observed when PII was not phosphorylated (Heinrich et al., 2004) consistent with the structural data that indicates that PII phosphorylation would prevent its binding to NAGK (Llacer et al., 2007).

The role of PII effectors on NAGK–PII complex formation is similar in both *S. elongatus* and *A. thaliana*. In...
S. elongatus, complex formation is promoted by MgATP and is inhibited by ADP or by MgATP together with 2-OG (Heinrich et al., 2004; Maheswaran et al., 2004; Beez et al., 2009). In A. thaliana, although initial studies found no effects of PII effectors (Chen et al., 2006), subsequent work found that, as in S. elongatus, complex formation requires MgATP and is inhibited by MgATP plus 2-OG (Beez et al., 2009; Feria Bourrellier et al., 2009). Unlike S. elongatus, binding of AtP II to AtNAGK was not antagonized by ADP which may reflect the very high affinity of AtP II for ATP (Beez et al., 2009).

Structural analysis supports the view that in vivo, the NAGK–P II complex would only be stable when P II is occupied by MgATP. In the structure of S. elongatus, P II with MgATP and 2-OG, and P II structures with bound ADP, the T-loops exhibit a conformation incompatible with NAGK interaction (Llacer et al., 2007, 2008; Fokina et al., 2010a). Hence, NAGK–P II interaction is predicted to occur only when the cellular energy levels are high (i.e. high ATP/ADP ratios) and nitrogen is abundant (low 2-OG). In this condition, P II would relieve NAGK arginine inhibition, allowing storage of excess nitrogen in the form of arginine. In cyanobacteria, P II phosphorylation provides a second level of control as reduced 2-OG levels favour P II de-phosphorylation and hence interaction with NAGK.

**Regulation of ACCase in the plant chloroplast**

In a study to identify novel P II protein targets in plants, mass spectrometry was used to identify proteins that bind to a His-tagged Arabidopsis P II mobilized on Ni2+ resin. In addition to NAGK, five proteins described as biotin/lipoyl attachment domain proteins were identified. Two of these, namely BCCP1 and BCCP2, are the biotin carboxyl carrier subunits of the plastidial ACCase which catalyses the carboxylation of acetyl-CoA to produce the pre-cursor of lipid biosynthesis, malonyl-CoA (Feria Bourrellier et al., 2010).

ACCase activity in chloroplast extracts was inhibited by P II in the presence of MgATP, and addition of 5 mM 2-OG, oxaloacetate or pyruvate to the assay relieved the inhibition (Feria Bourrellier et al., 2010). This is the first report of a P II protein regulating the activity of an enzyme dedicated to carbon metabolism, thereby expanding P II function beyond nitrogen regulation. The 2-OG effect is probably because of the loss of protein–protein interaction between ACCase and P II, suggesting that 2-OG is the main signal controlling P II inhibition of ACCase (Feria Bourrellier et al., 2010). In plants, ammonium is almost exclusively assimilated in the plastid by the GS-GOGAT pathway which uses 2-OG as the carbon skeleton for amino acid production. The plastid 2-OG pool is presumably fuelled by mitochondrial TCA intermediates which in turn reflect the cellular carbon status (Galvez et al., 1999). Hence, accumulation of 2-OG in the chloroplast indicates a plentiful supply of carbon, or a high C/N ratio and the consequent relief of ACCase inhibition by P II would allow the allocation of available carbon to fuel lipid synthesis.

The effects of oxaloacetate and pyruvate are less clear: although oxaloacetate has been shown to bind to Arabidopsis P II and affect the interaction between NAGK and P II, pyruvate has not been tested in this way (Smith et al., 2003; Feria Bourrellier et al., 2009). It remains to be determined whether and how oxaloacetate and/or pyruvate mediate P II inhibition of ACCase under physiologically relevant conditions.

**Other putative P II enzyme targets**

In R. capsulatus, yeast-two hybrid library screening using GlnB or GlnK as bait identified PcrA and Era as a putative P II targets, respectively (Pawlowski et al., 2003). PcrA belongs to super-family 1 of DNA helicases and Era belongs to the Ras-like GTPase super-family involved in many cellular processes including ribosome biogenesis. These interactions are yet to be confirmed using other approaches.

**Regulation of transcription factors**

Most major classes of prokaryotes studied to date are characterized by the presence of a global transcription factor that coordinates the control of a large regulon comprising many genes involved in cellular N metabolism. A notable common feature of these transcription factors, many of which have been the subjects of specific reviews, is that P II proteins are almost invariably involved in controlling their activity (Merrick & Edwards, 1995; Herrero et al., 2001; Forchhammer, 2007; Leigh & Dodsworth, 2007; Sonenshein, 2007; Amon et al., 2010).

**The two component regulatory system, NtrB/NtrC in Proteobacteria**

The global transcriptional control of nitrogen-regulated genes in Proteobacteria is mediated by a two-component histidine kinase system, NtrB/NtrC, that has been extensively studied in E. coli (Chen et al., 1982). NtrC is a homodimer with the characteristic architecture of σN-dependent transcriptional activators (Reitzer & Magasanik, 1983). The N-terminal receiver domain is phosphorylated by the sensor protein NtrB (Weiss & Magasanik, 1988). The central region of NtrC contains an AAA+ ATPase domain that interacts with σN RNA polymerase and hydrolyses ATP to facilitate transcription initiation, and the C-terminal domain contains a DNA-binding motif.
Upon phosphorylation, NtrC oligomerizes, increasing its affinity for its binding sites on DNA and enhancing ATPase activity (Hunt & Magasanik, 1985; Ninfa & Magasanik, 1986; Weiss et al., 1991; Austin & Dixon, 1992; Rippe et al., 1998). Termination of transcriptional activation is brought about either by the regulated NtrC-P phosphatase activity of the sensor protein NtrB or by a constitutive NtrC auto-phosphatase activity (Keener & Kustu, 1988).

NtrB is a homodimer, each subunit containing a N-terminal sensor domain, a central phosphotransferase/phosphatase/dimerization domain, and a C-terminal kinase domain (Jiang et al., 2000a). One NtrB monomer binds ATP and phosphorylates the opposing monomer, which in turn transfers the phosphate to NtrC (Jiang et al., 2000b). NtrB can also stimulate NtrC-P dephosphorylation, and it is the switch between the opposing activities of NtrB that is coordinated by interaction with GlnB (Ninfa et al., 1993; Kamberov et al., 1994a; Jiang & Ninfa, 1999). Although the predominant regulator of NtrB is GlnB, GlnK can also regulate the NtrC phosphorylation status in vivo, but to a lesser extent than GlnB (Atkinson & Ninfa, 1998). As with many other interactions of PII proteins with their targets, the interaction of GlnB with NtrB is regulated, both by the modification state of GlnB and by PII-bound effectors.

Unmodified GlnB binds to NtrB, enhancing NtrB phosphatase activity and inhibiting NtrB autophosphorylation. This binding involves the GlnB T-loops (Pioszak et al., 2000) and studies with PII heterotrimers suggest that only a single T-loop is necessary to achieve NtrB regulation (Jiang et al., 1997c). Binding of GlnB to NtrB also requires 2-OG. However, within the physiological range of 2-OG, i.e. from 0.03 to 5 mM, GlnB becomes a progressively less effective inhibitor of the autophosphorylation activity of NtrB as the 2-OG level rises (Jiang & Ninfa, 2009a) and consequently NtrC phosphorylation by NtrB will increase as the 2-OG level rises. It has been proposed that the primary effect of 2-OG is mediated through the conformation of the T-loop (Jiang & Ninfa, 2009a). However, the effect in vivo of increasing 2-OG can be difficult to separate from that of a reciprocally falling glutamine pool which will lead to progressive uridylylation of GlnB. As GlnB-UMP₃ cannot interact with NtrB, the NtrB kinase activity then dominates and NtrC becomes phosphorylated (Atkinson et al., 1994; Jiang & Ninfa, 1999; Pioszak et al., 2000). The influence of the adenylate charge (i.e. the ATP/ADP ratio) on GlnB–NtrB interaction has also been studied and the 2-OG regulated binding of GlnB to NtrB was not antagonized by the presence of ADP; a result that contrasts with PII signaling in other systems (Jiang & Ninfa, 2009c).

The model that emerges from these studies is that, under N-sufficient conditions GlnB interacts with NtrB and activates its phosphatase activity such that NtrC is dephosphorylated and genes in the NtrC regulon are not transcribed. As the cell moves to an N-limited condition, the cellular pool of 2-OG rises and this has two effects. It progressively reduces the ability of GlnB to inhibit NtrB autophosphorylation, presumptively through a change in the T-loop conformation. That change in conformation may in turn facilitate the progressive uridylylation of GlnB and its ultimate disengagement from NtrB. NtrB kinase activity then prevails resulting in NtrC phosphorylation and activation of genes in the NtrC regulon.

Despite the lack of in vitro studies of the NtrBC system in other organisms, mutant analyses suggest that NtrB activation by PII is conserved (Arcondéguy et al., 2001). The only apparent exception is in S. meliloti, where GlnB is required for the induction of the NtcA-dependent genes glmA and glnII under nitrogen limitation, suggesting that in this organism GlnB-UMP₃ is required for stimulation of NtrB-kinase activity (Arcondéguy et al., 1997).

The cyanobacterial nitrogen control protein NtcA

In cyanobacteria when ammonium is limiting the NtcA protein, which belongs to the CRP family of transcription factors, induces expression of several genes involved in ammonium assimilation, in the utilization of other nitrogen sources, and in heterocyst formation. NtcA can also act as a repressor if the NtcA site overlaps the promoter region (Jiang et al., 1997a; Herrero et al., 2004; Luque et al., 2004). 2-OG binds to NtcA enhancing its ability to interact with DNA and to activate transcription (Tanigawa et al., 2002; Vazquez-Bermudez et al., 2002).

In S. elongatus, PII is required for NtcA-regulated gene expression during N-limination (Aldehni et al., 2003; Paz-Yepes et al., 2003), but the mechanism involves a co-activator of NtcA, called PipX, that is found exclusively in cyanobacteria. Synechococcus PipX interacts with both PII and NtcA (Espinosa et al., 2006) so that NtcA competes with PipX for PII binding (Llacer et al., 2010). Interaction between PII and PipX is stabilized in vitro by ADP and is dissociated in the presence of MgATP and high 2-OG (Llacer et al., 2010). In vitro NtcA–PipX interaction is stimulated by high 2-OG (Espinosa et al., 2006), a condition which signals N-limitation in cyanobacteria (Muro-Pastor et al., 2001).

The resolution of the structures of NtcA and the PII–PipX complex from Anabaena (Zhao et al., 2010a, b), and of NtcA and the PII–PipX and PipX-NtcA complexes from S. elongatus (Llacer et al., 2010) clarified the mode of NtcA regulation by PII. The PipX–PII complex contains three PipX molecules embraced between the extended T-loops of one PII trimer (Fig. 6; Llacer et al., 2010; Zhao et al., 2010b). In the Anabaena structure PII was occupied...
by ADP, whereas in *S. elongatus* the PII nucleotide binding sites were empty. However, *in vitro* studies support the proposition that the PII-PipX complex would only be stable when PII is occupied by ADP *in vivo* (Fokina et al., 2011). A phosphorylated PII Ser49 could be modelled in the PipX–PII complex of *S. elongatus* without any charge or steric restriction (Llacer et al., 2010). This suggests that PII–PipX interaction is regulated by PII effectors rather than PII phosphorylation, and is consistent with the observation that changes to the PII T-loop residue Ser49 do not affect interaction with PipX (Espinosa et al., 2006).

In the resultant model, PII is bound by MgATP and 2-OG in N-limitation and cannot interact with PipX which is therefore available to bind NtcA. The NtcA–PipX interaction is further stimulated by high 2-OG. When NtcA is complexed to PipX, and also saturated with 2-OG it binds more avidly to DNA and/or makes more efficient contacts with RNA polymerase thereby activating transcription. Conversely, in N-sufficiency low 2-OG facilitates both the binding of ADP to PII and dissociation of 2-OG from NtcA; hence PII interacts avidly with PipX abrogating the PipX activation of NtcA. Regulation of NtcA by PII might also be influenced by cellular energy levels. Low ATP/ADP ratios would facilitate the binding of ADP to PII, augmenting the affinity between PII and PipX, thereby reducing NtcA-dependent gene expression (Fokina et al., 2011).

**The nitrogen regulatory proteins GlnR and TnrA in Bacilli**

Transcriptional regulation of nitrogen related genes in *Bacilli* relies on two similar MerR-type regulatory proteins, GlnR and TnrA. When nitrogen is in excess, GlnR acts as a transcriptional repressor, and when nitrogen is limited TnrA activates transcription of genes involved in the use of alternative nitrogen sources, though TnrA can also act as a transcription repressor (Amon et al., 2010; Fisher, 1999). The DNA-binding activities of both GlnR and TnrA are controlled by their interaction with type I GS.

The activity of TnrA is controlled by a mechanism involving competitive binding of TnrA to GS or GlnK, changes in TnrA cellular localization and nitrogen regulated proteolysis of TnrA. When *B. subtilis* is grown in the poor nitrogen source nitrate, TnrA is active and is also membrane-associated, presumptively by the formation of an AmtB–GlnK–TnrA ternary complex. The role of PII effectors is not completely clear, but the presence of MgATP and 2-OG dissociates both GlnK and TnrA from AmtB *in vitro* (Heinrich et al., 2006). The TnrA–GlnK interaction is inhibited by ATP or ADP, and Mg$^{2+}$ and Mn$^{2+}$ partially relieve the ATP inhibition whereas 2-OG has no effect (Kayumov et al., 2011). In N-sufficiency, GS is feed-back inhibited by glutamine, this form of GS outcompetes GlnK for TnrA binding so that GS then forms a stable complex with TnrA and inhibits its activator function (Wray et al., 2001).

The stability of TnrA also depends on the availability of nitrogen. When *B. subtilis* cultivated on nitrate is shifted to nitrogen-free medium, TnrA moves from the membrane to the cytoplasm and is rapidly degraded (Kayumov et al., 2008) although the physiological role of this proteolysis is unclear, and the mechanism of complex dissociation under this condition is unknown. TnrA is always located in the cytoplasm in *amtB* and *glnK* mutants and is then not degraded upon nitrogen starvation (Kayumov et al., 2008). It was found that TnrA binds constitutively to GlnK or to GS in *amtB* or *glnK* mutant backgrounds, respectively (Kayumov et al., 2011). The C-terminal region of TnrA is necessary for interaction with both GS and GlnK *in vitro*, and the same region is required for proteolysis. Hence, the interactions of TnrA with either GS or GlnK are competitive and apparently both proteins can protect TnrA from proteolysis (Kayumov et al., 2011).

The repressor function of GlnR is also stabilized by complex formation with feed-back inhibited GS and consequently GlnR acts as a repressor in N-sufficiency (Fisher & Wray, 2008; Wray & Fisher, 2008). Given the similarities between TnrA and GlnR, it is likely that GlnR might also interact with GlnK. Indeed, *in vitro* complex formation between GlnR and GlnK was reported in *S. mutans* and the interaction apparently increased the GlnR affinity for DNA (Castellon et al., 2011). However, the role of PII effectors and the physiological relevance of such an interaction remain to be determined.
The GlnR protein in Actinobacteria

Transcriptional regulation of nitrogen genes in Actinobacteria also relies on a transcription factor called GlnR (Fink et al., 2002; Tiffert et al., 2008; Pullan et al., 2011). However, despite sharing the same name with the Bacillales GlnR, the GlnR protein found in Actinobacteria belongs to the OmpR-type family with a C-terminal rather than an N-terminal DNA binding domain. Studies of an S. coelicolor glnK mutant showed no effects on expression of nitrogen metabolism genes suggesting that GlnK was not an important N sensor, at least under the culture conditions tested, and hence it may not regulate GlnR activity (Waldvogel et al., 2011).

The nitrogen regulatory protein AmtR in Corynebacterium

In Corynebacterium, transcription control of nitrogen-related genes is mediated by a repressor protein of the TetR family named AmtR which represses transcription when nitrogen is abundant (Jakoby et al., 2000; Beckers et al., 2005; Buchinger et al., 2009). In glnK or glnD mutants, or in a GlnK Tyr51Phe variant, AmtR-regulated genes were not expressed in N-limitation (Nolden et al., 2001b). This suggested that GlnK might regulate AmtR activity and it was subsequently shown that adenylated GlnK binds to AmtR in vitro and stimulates its dissociation from DNA (Beckers et al., 2005). Furthermore, as in other organisms, GlnK binds to AmtB in N-sufficient conditions, although in C. glutamicum once localized to the membrane most GlnK is rapidly degraded (Strosser et al., 2004). Hence, the model that emerges in Corynebacterium is quite distinct from that in other organisms. Under N-limitation, adenylated GlnK binds to AmtR releasing it from repressor sites and allowing expression of genes in the AmtR regulon. Upon an ammonium shock, GlnK is de-adenylated, binds to AmtB on the membrane and is degraded by proteolysis thereby allowing AmtR to repress transcription of genes involved in the use of alternative nitrogen sources.

It remains to be determined whether PII effectors influence the interaction between AmtR and GlnK. Furthermore, the signals controlling the opposing adenyltransferase/adenyl-removing activities of GlnD remain to be elucidated. Unlike the situation in most Proteobacteria, glutamine seems not to be a major nitrogen signal in C. glutamicum (Nolden et al., 2001b; Rehm et al., 2010), rather the evidence suggests that the primary signal is accumulation of 2-OG which indicates nitrogen deficiency (Muller et al., 2006).

Control of the nitrogen fixation transcriptional activator NifA

In addition to global regulators, one other conserved transcription factor that is subject to regulation by PII is the nitrogen fixation transcriptional activator NifA, which controls nif gene expression in diazotrophic Proteobacteria. Nitrogenase activity can be regulated both at transcriptional and post-translational levels, and the latter has been discussed earlier. The two major environmental signals that regulate nitrogenase in Proteobacteria are the cellular oxygen and nitrogen levels (Dixon & Kahn, 2004). Oxygen regulation is necessary because the metal clusters of nitrogenase are oxygen sensitive, and nitrogen regulation ensures that nitrogenase will operate only when necessary to fuel the anabolic reactions with fixed nitrogen.

The NifA protein is a $\sigma^N$, dependent transcriptional activator which has a similar architecture to NtrC (Drummond et al., 1986, 1990; Morett & Segovia, 1993). The major differences between NifA and NtrC are localized in the N-terminal region. Instead of a typical two-component receiver domain that is the target for phosphorylation, NifA contains an N-terminal GAF domain which controls NifA activity through interaction with other proteins (Drummond et al., 1990). NifA activity can be regulated in response to nitrogen either directly or by an anti-activator protein called NifL. NifA activity can also be influenced by oxygen levels (for reviews on NifA regulation see Dixon & Kahn, 2004; Martinez-Argudo et al., 2004a).

NifL-mediated regulation of NifA activity

In the Gammaproteobacteria Klebsiella pneumoniae (Kennedy, 1977), A. vinelandii (Blanco et al., 1993; Raina et al., 1993) and P. stutzeri (Desnoues et al., 2003), and in the Betaproteobacteria Azooarcus sp. (Egener et al., 2002), nifA is co-transcribed with a regulatory gene named nifL. NifL is a flavoprotein which interacts with NifA to inhibit NifA activity (Dixon & Kahn, 2004). A variety of signals regulate NifL-NifA complex formation such as cellular redox, small metabolite binding and interaction with PII (Martinez-Argudo et al., 2004b). Despite the general similarity, the mechanism of the NifLA system differs between the two model organisms, A. vinelandii and K. pneumoniae.

Constitutive expression of the K. pneumoniae nifLA genes in E. coli resulted in nitrogen-regulated NifA activity, indicating that NifLA can respond to a signalling system also present in E. coli (He et al., 1998). A role for GlnB was excluded (Holter & Merrick, 1989; He et al., 1998), but the subsequent discovery of GlnK again
implicated $P_\Pi$ in NifLA control. NifA is constitutively inactive in a $glnK$ mutant and uridylylation of GlnK does not influence NifA activity indicating that GlnK relieves NifL inhibition of NifA independently of its uridylylation status (Edwards & Merrick, 1995; He et al., 1998; Jack et al., 1999).

The mechanism of NifL regulation by GlnK in *K. pneumoniae* is not completely understood. GlnK interacts with both NifA and NifL in *in vivo* suggesting that these proteins might form a GlnK-NifL-NifA ternary complex (Stips et al., 2004; Gloer et al., 2008). Over-expression of GlnK elevates *nif* gene expression, suggesting that GlnK promotes dissociation of NifL-NifA, thereby enhancing NifA activity (Stips et al., 2004). However, in vitro neither GlnK nor GlnK-UMP₃ affected complex formation between NifL and NifA (Stips et al., 2004). The cellular location of NifL may be significant because it is predominately membrane-bound under nitrogen-fixing conditions, but not when the cells are exposed to oxygen or to ammonium. However, in the absence of GlnK, NifL was located in the cytoplasm even under de-repressing conditions (Klopprogge et al., 2002). By contrast, NifA is always cytoplasmic (Klopprogge et al., 2002; Stips et al., 2004; Milenkov et al., 2011). Hence, membrane-bound NifL could be physically separated from NifA allowing *nif* gene transcription under nitrogen-limiting conditions (Klopprogge et al., 2002).

Despite the function of GlnK being independent of its uridylylation state, the T-loop has been directly implicated in regulating the NifLA complex. Just two amino acid changes (Asp54 to Asn and Thr43 to Ala) in the T-loop of GlnB, converting it to the GlnK T-loop sequence, are sufficient for GlnB to functionally complement a glnK mutant and relieve NifL inhibition of NifA (Arcondégyuy et al., 2000). Changes in the T-loop were also found in a genetic screen for GlnK variants that could not transduce the nitrogen signal to NifLA (Gloer et al., 2008). 2-OG significantly reduces GlnK binding to NifA *in vitro*, but not to NifL, suggesting a role for the T-loop in NifA interaction (Gloer et al., 2008).

Taken together, these data suggest that in *K. pneumoniae* N-limitation GlnK-UMP₃ promotes dissociation of the NifL-NifA complex and NifL binds to the cell membrane allowing NifA to activate *nif* transcription. However, ammonium upshift causes deuridylylation of GlnK which binds to AmtB on the membrane thereby favouring NifL-NifA complex formation and reduction in *nif* gene expression. Studies in *P. stutzeri* support a similar NifLA regulatory mechanism (He et al., 2008; Zhang et al., 2012).

Unlike *K. pneumoniae*, *A. vinelandii* has only a single $P_\Pi$ protein encoded in a *glnK* *amtB* operon (Meletzus et al., 1998) and GlnK uridylylation is essential for nitrogen fixation (Contreras et al., 1991; Rudnick et al., 2002). There is also no evidence for NifL interaction with the cellular membrane in *A. vinelandii*. Unmodified GlnK interacts directly *in vitro* with the C-terminal histidine kinase-like domain of NifL and increases its inhibitory activity on NifA; whereas GlnK-UMP₃ does not interact with NifL (Little et al., 2002).

Direct binding of nucleotides and 2-OG to all three proteins; GlnK, NifL and NifA also regulates complex formation. GlnK interaction with NifL is dependent on both MgATP and 2-OG (Little et al., 2000). NifL binds ATP and ADP, and both nucleotides stimulate NifL-NifA complex formation (Eydmann et al., 1995; Söderbäck et al., 1998; Money et al., 1999), whereas the N-terminal NifA GAF domain binds 2-OG causing a reduction in NifL/ NifA interaction (Little et al., 2000; Little & Dixon, 2003; Martinez-Argudo et al., 2004a). The interaction between GlnK and NifL restores NifA inhibition even when the 2-OG level is saturating, therefore the GlnK signal overrides the 2-OG signal (Little & Dixon, 2003). The data suggest that in N-sufficiency unmodified GlnK binds to NifL promoting NifLA complex formation, and hence *nif* gene inactivation, whereas uridylylation of GlnK in response to N-limitation prevents its interaction with NifL leading to NifA release and *nif* expression.

**Direct regulation of NifA activity by $P_\Pi$**

In diazotrophs belonging to the Alphaproteobacteria and some Betaproteobacteria, *nifL* is not present and in these cases NifA activity is directly regulated by $P_\Pi$ (Dixon & Kahn, 2004). The precise mechanism of this regulation varies: in *A. brasilense*, *H. seropedicae* and *R. rubrum* $P_\Pi$ proteins are required to activate NifA when nitrogen is limiting, whereas in *R. capsulatus*, *A. caulino nodans* and *Gluconacetobacter diazotrophicus* $P_\Pi$ seems to be required to prevent NifA activity when nitrogen is abundant.

In the first example of such regulation, *A. brasilense* GlnB was found to be required for *nif* expression (de Zamaroczy et al., 1993; de Zamaroczy, 1998; Araujo et al., 2004). As *nifA* expression was not affected it was proposed that GlnB regulates NifA activity (de Zamaroczy et al., 1993; Fadel-Picheth et al., 1999). GlnB was subsequently shown to interact with the NifA GAF domain (Chen et al., 2005) and NifA variants carrying deletions in the GAF domain are constitutively active even in a *glnB* mutant (Arsene et al., 1996, 1999). These results suggest that the GAF domain inhibits NifA activity, unless relieved by interaction with GlnB-UMP₃ in N-limiting conditions, whereupon NifA activates *nif* expression. Apparently, GlnB and GlnB-UMP₃ compete for the same binding site on NifA though only the latter is competent to activate NifA *in vivo* (Arsene et al., 1999; van
Dommelen et al., 2002; Huergo et al., 2005). This mechanism allows nif gene expression to respond rapidly to transient fluctuations in the availability of ammonium (Chen et al., 2005; Huergo et al., 2005).

Studies in R. rubrum and H. seropedicae suggest a similar mechanism. In R. rubrum, disruption of glnB or glnD prevents nif expression (Zhang et al., 2000, 2001b). GlnB has been shown to interact with NifA (Zhu et al., 2006) and a GlnB Tyr51Phe variant has reduced nitrogenase activity (Zhang et al., 2000). In H. seropedicae, the NifA GAF domain interacts in vitro with both GlnK and GlnK-UMP3 (Oliveira et al., 2012) and in vivo studies suggest that GlnK-UMP3 activates NifA under N-limiting conditions by relieving the inhibitory function of the NifA GAF domain (Noindorf et al., 2011).

An alternative mechanism for regulating NifA activity is seen in R. capsulatus, A. caulinodans and G. diazotrophicus. Rhodobacter capsulatus has two similar nifA genes with overlapping functions (Masepohl et al., 1988) both of which are activated under nitrogen limitation by NtrC (Foster-Hartnett & Kranz, 1992). Although constitutive expression of nifA does not allow nif gene expression in the presence of ammonium (Hübner et al., 1993), mutations in the NifA1 GAF domain or knockout of both glnB and glnK do allow such expression (Paschen et al., 2001; Drepper et al., 2003). Consistent with this either GlnB or GlnK can apparently interact with NifA to prevent its activity when nitrogen is abundant (Pawlowski et al., 2003) and evidence for a similar mechanism is found in A. caulinodans (Michel-Reydellet et al., 1997; Michel-Reydellet & Kaminski, 1999) and in G. diazotrophicus (Perlova et al., 2002, 2003).

Concluding remarks

Looking back over a decade since our last comprehensive review of PII biology (Arcondéguy et al., 2001), there have undoubtedly been remarkable advances in our understanding of both the role and the mode of action of these proteins. The breadth of their influence in the control of nitrogen metabolism in both the bacteria and the archaea has been clearly confirmed, although in plants we are only now beginning to explore the role that they play.

Structural studies have made one of the most significant contributions to the development of our understanding of PII proteins and major advances have occurred within the last 5 years. These advances fall into two areas; the role of effector molecules, and the mode of interaction of PII proteins with their targets. The recognition of ADP as a physiologically important effector was highlighted by the structure of the GlnK-AmtB complex; whereas the identification of the long sought for binding site for 2-OG resolved many of the earlier data on synergy between ATP and 2-OG. The availability of structures for PII proteins bound to four different targets including enzymes, a membrane protein and a transcription factor regulator, has confirmed the earlier predictions that the flexible T-loops play a key role in many target interactions. However, at the same time it is now apparent that PII proteins can form ternary complexes involving simultaneous interactions with two different targets and that, as evidenced by the structure of the GlnZ-DraG complex, these interactions can involve surfaces of PII other than the T-loops. These structures serve to underline the quite remarkable way in which evolution has capitalized on the inherent signalling attributes of the PII proteins, which derive from their effector-binding properties and their conformational flexibility, so as to recruit PII to control gene expression, substrate uptake and cellular metabolism.

Despite these advances, a number of key questions remain to be addressed. To date, there is a clear threefold symmetry in each of the complexes where structures have been solved, and in an evolutionary sense this can be seen to have originated in the GlnK–AmtB interaction. However, PII proteins are known to target a number of proteins whose native state is dimeric; notably many of the transcription factor targets. So, it will be of particular interest to understand the stoichiometry and form of these interactions. Likewise, the recognition of membrane sequestration and ternary complex formation as modes of action in a variety of situations opens up the possibility that this may be a more common facet of PII biology than we have yet appreciated.

Although we now have knowledge of the binding sites for 2-OG, ATP and ADP and an indication of the structural consequences of the binding of these molecules, especially with regard to their influence on T-loop conformation, we are still some way from a complete understanding of the signal transduction process. There is a wealth of in vitro data on effector binding to a variety of PII proteins, but very little in vivo information on either changes in effector pools or the absolute concentrations of these molecules in the cell. Such data are notoriously difficult to obtain and there are many concerns about how accurately one can assess the intracellular metabolite pools (Schneider & Gourse, 2004; Bennett et al., 2009; Yuan et al., 2009; Yan et al., 2011). Consequently, there is an urgent need to obtain more robust data on these in vivo values. Only then can we be in a position to correlate with some confidence the activity of a particular PII protein with respect to one or more of its targets, and the intracellular metabolite changes that are driving that activity. This topic is of particular relevance to the debate on the role of PII proteins in sensing cellular energy charge. Much has been written about this possibility, and
there is considerable in vitro data on binding of ATP and ADP to PII which supports current hypotheses on the ability of PII to respond to the ATP/ADP ratio. However, we need clear quantitative in vivo data to test these hypotheses and to date that has been difficult to obtain (Zhang et al., 2009).

The recent demonstration that PII can control the activity of ACCase in plant chloroplasts is the first clear demonstration of PII regulating the activity of an enzyme involved in carbon rather than nitrogen metabolism. This links directly to the pivotal position of 2-OG, with the potential to reflect both carbon and nitrogen status, and it remains to be seen whether more examples of PII functioning in roles unrelated to N metabolism will be discovered. Indeed, there is undoubtedly still considerable scope for studies on undescribed PII targets. Gene linkage studies have long suggested a role in regulating NAD synthesis and nitrate uptake in certain organisms (Arcondéguy et al., 2001), and there is certainly no reason to suppose that the list of known PII targets is yet complete.

Other unexplored areas of PII biology include the recently described PII-NG proteins with the intriguing possibility that they might act a metal sensors (Sant'Anna et al., 2009). Both in vivo biological studies and protein structural studies would be very informative here. Likewise, although the basic nature of the role of the NifI proteins has been elegantly described by Leigh & Dods- worth (2007), structural information on the proposed NifI1-NifI2 complex is still lacking.

So, despite tremendous advances in PII biology having been made in the last decade by the concerted efforts of many labs around the world, there are undoubtedly still more important and probably unexpected discoveries to be made in this field in the future. The next decade of research in PII biology could be just as exciting as the last one.

Acknowledgements

L.F.H. acknowledges the financial support from CNPq, INCT and Fundação Araucária. M.M. and G.C. acknowledge support from the Biotechnology and Biological Sciences Research Council (BBSRC). Emanuel M Souza, Fritz Winkler, Jake Malone and Ray Dixon are all thanked for constructive comments on the manuscript.

References


Fokina O, Chellamuthu VR, Zeth K & Forchhammer K (2010b) A novel signal transduction protein P$_1$ variant from *Synechococcus elongatus* PCC 7942 indicates a two-step...


Jiang P & Ninfa AJ (2009a) Alpha-ketoglutarate controls the ability of the *Escherichia coli* PII signal transduction protein to regulate the activities of NRII (NtrB) but does not control the binding of PII to NRII. *Biochemistry* **48**: 11514–11521.


signal transducer enables the Azotobacter vinelandii NifL-NifA regulatory system to respond to conditions replete for nitrogen. *J Biol Chem* **277**: 15472–15481.


Radcenko MV, Thornton J & Merrick M (2010) Control of AmtB-GlnK complex formation by intracellular levels of


**Supporting Information**

Additional Supporting Information may be found in the online version of this article:

**Table S1.** Those fully sequenced prokaryote genomes that do not encode PII genes.

**Table S2.** Those prokaryote genera in which some or all species lack PII genes.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.