

A New P_{II} Protein Structure Identifies the 2-Oxoglutarate Binding Site

Daphne Truan¹, Luciano F. Huergo², Leda S. Chubatsu²,
Mike Merrick³, Xiao-Dan Li⁴ and Fritz K. Winkler^{4*}

¹Macromolecular
Crystallography, Swiss Light
Source, Villigen PSI,
Switzerland

²Instituto Nacional de Ciência e
Tecnologia da Fixação Biológica
de Nitrogênio, Departamento de
Bioquímica e Biologia
Molecular, Universidade Federal
do Paraná, Curitiba, PR, Brazil

³Department of Molecular
Microbiology, John Innes
Center, Norwich, UK

⁴Biomolecular Research, Paul
Scherrer Institut, Villigen PSI,
Switzerland

Received 4 February 2010;
received in revised form
13 May 2010;
accepted 14 May 2010
Available online
21 May 2010

P_{II} proteins of bacteria, archaea, and plants regulate many facets of nitrogen metabolism. They do so by interacting with their target proteins, which can be enzymes, transcription factors, or membrane proteins. A key feature of the ability of P_{II} proteins to sense cellular nitrogen status and to interact accordingly with their targets is their binding of the key metabolic intermediate 2-oxoglutarate (2-OG). However, the binding site of this ligand within P_{II} proteins has been controversial. We have now solved the X-ray structure, at 1.4 Å resolution, of the *Azospirillum brasilense* P_{II} protein GlnZ complexed with MgATP and 2-OG. This structure is in excellent agreement with previous biochemical data on 2-OG binding to a variety of P_{II} proteins and shows that 2-oxoglutarate binds within the cleft formed between neighboring subunits of the homotrimer. The 2-oxo acid moiety of bound 2-OG ligates the bound Mg²⁺ together with three phosphate oxygens of ATP and the side chain of the T-loop residue Gln39. Our structure is in stark contrast to an earlier structure of the *Methanococcus jannaschii* GlnK1 protein in which the authors reported 2-OG binding to the T-loop of that P_{II} protein. In the light of our new structure, three families of T-loop conformations, each associated with a distinct effector binding mode and characterized by a different interaction partner of the ammonium group of the conserved residue Lys58, emerge as a common structural basis for effector signal output by P_{II} proteins.

© 2010 Elsevier Ltd. All rights reserved.

Edited by I. Wilson

Keywords: nitrogen regulation; P_{II} protein; 2-oxoglutarate; GlnZ; *Azospirillum brasilense*

Introduction

P_{II} proteins are involved in the regulation of many aspects of nitrogen metabolism.¹ They can be regarded as signal integration proteins whose output is their signal-dependent interaction with various target proteins that they may activate or inhibit.^{2–4} The universally conserved and probably most ancient signals controlling P_{II} activities are the effector molecules ATP, ADP, and 2-oxoglutarate (2-OG). Additionally, many P_{II} proteins are covalently

modified by enzymes whose activities are regulated by another key nitrogen metabolite, glutamine. P_{II} proteins are compact, cylindrically shaped (homo) trimers composed of 12-kDa to 13-kDa subunits from which three long exposed loops, the so-called T-loops, protrude (Fig. 1), as first reported for *Escherichia coli* GlnB.⁵ The T-loops are significantly conserved in sequence, but as is apparent from the numerous reported structures of P_{II} proteins, they are structurally very flexible.^{3,6} They are vital for P_{II} interactions with many of their targets and are also the sites of reversible covalent modification. In addition to the T-loops, the highly conserved structure of P_{II} proteins is characterized by three lateral intersubunit clefts within which two smaller loops (the B-loops and C-loops) from opposing subunits contribute to an adenyl-nucleotide binding pocket where ADP or ATP can bind competitively (Fig. 1).

*Corresponding author. E-mail address:
fritz.winkler@psi.ch.

Abbreviations used: 2-OG, 2-oxoglutarate; PDB, Protein Data Bank; NAGK, N-acetylglutamate kinase; ACCase, acetyl-CoA carboxylase; Mes, 4-morpholineethanesulfonic acid.

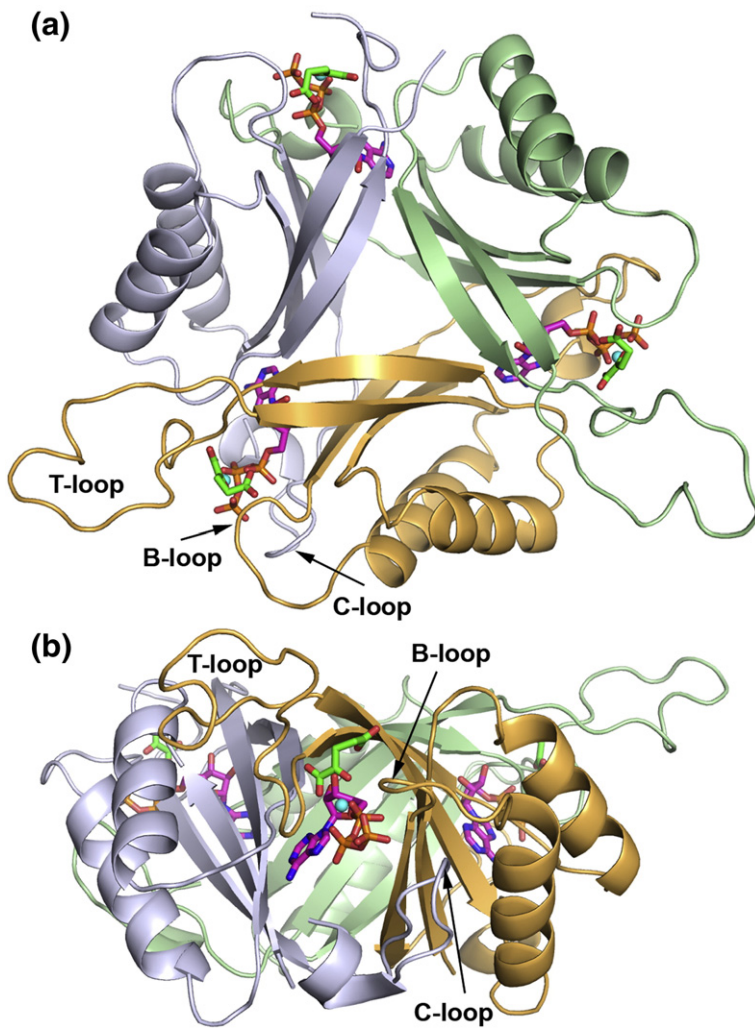


Fig. 1. Cartoon representation of the GlnZ trimer. The trimer with bound ATP (magenta), 2-OG (green), and Mg²⁺ (blue) is shown in top view (a) and side view (b). The T-loop of one protomer (chain A) is partly disordered. The T-loop (residues 37–55), B-loop (residues 82–88), and C-loop (residues 102–105) at one binding site are indicated by arrows.

The binding mode of the other key effector 2-OG, which is assumed to be conserved among all P_{II} proteins, has remained enigmatic and controversial.³ Here we report a high-resolution structure of a complex of the P_{II} protein GlnZ from *Azospirillum brasilense* complexed with the effectors ATP and 2-OG. GlnZ and its orthologues are specifically involved in the regulation of nitrogenase activity in some nitrogen-fixing bacteria.^{7,8} The observed 2-OG binding site is in excellent agreement with biochemical data, in contrast to a previously reported binding mode.⁹ This new structure will greatly facilitate an understanding of the link between changes in cellular effector pools and P_{II} signaling.

Results and Discussion

Crystal structure

Purified GlnZ was crystallized in the presence of ATP, 2-OG, and Mg²⁺, and the structure was solved by molecular replacement using the *E. coli* GlnK trimer as search model [Protein Data Bank (PDB) ID 2GNK].¹⁰ First electron density maps revealed clear

density for bound MgATP and 2-OG in all three independent binding sites of the GlnZ trimer (chains A, B, and C) present in the asymmetric unit. Refinement using data up to 1.4 Å resolution, as described in Materials and Methods, led to a final model with an *R*-factor of 0.160 (*R*_{free} = 0.198) and excellent stereochemistry (Table 1). The final electron density at the effector binding site is shown in Fig. 2. Two of the three T-loops (residues 37–55) of each trimer are completely ordered (Fig. 1) and are observed in the same conformations with similar interactions with neighboring trimers in the crystal lattice. The third, belonging to chain A, appears disordered from residues 43 to 52.

The 2-OG binding site

In our structure, ATP and 2-OG both bind to GlnZ by participating in the coordination of an Mg²⁺ ion (Figs. 2 and 3). Three of the oxygen ligands are provided by the α-phosphate, β-phosphate, and γ-phosphate of ATP, and two of the oxygen ligands are provided by the 2-oxo acid moiety of 2-OG whose 5-carboxy group is involved in a salt bridge with Lys58. The sixth ligand, completing the nearly perfect octahedral Mg²⁺ coordination, is provided

Table 1. Data collection and refinement statistics

<i>Data collection</i>	
Space group	C222 ₁
Cell dimensions <i>a</i> , <i>b</i> , <i>c</i> (Å)	64.98, 87.95, 116.88
Resolution (Å)	47.7–1.30 (1.40–1.30) ^a
<i>R</i> _{merge} ^b	0.078 (0.576) ^a
<i>I</i> / σ <i>I</i>	15.3 (2.8) ^a
Completeness (%)	98.0 (96.5) ^a
Number of unique/measured reflections	80,692/585,432
Redundancy	7.3 (7.0) ^a
<i>Refinement</i>	
Resolution (Å)	20.0–1.4
Number of reflections	65,019
<i>R</i> _{work} / <i>R</i> _{free} ^c	0.160/0.198
Number of atoms	2991
Protein	2487
Ligand/ion	178
Water	321
Average <i>B</i> -factors (Å ²)	
All	21.3
Protein (chains ABC)	19.6
Ligand/ion (chains ABC)	20.6
Water	35.0
r.m.s.d. ^d	
Bond lengths (Å)	0.006
Bond angles (°)	0.90
Ramachandran statistics	
Outliers (%)	0.0
Favored (%)	99.1

^a The values in parentheses refer to statistics in the highest-resolution bin.

^b $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl) \rangle$ is the mean value for its unique reflection; summations are over all reflections.

^c $R_{\text{work}}/R_{\text{free}} = \sum_h (F_o(h) - F_c(h)) / \sum_h F_o(h)$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. R_{free} was calculated with 5% of the data excluded from the refinement.

^d r.m.s.d. from ideal values.

by the carboxamide oxygen of the Gln39 side chain. The carboxylate oxygen coordinating the Mg²⁺ also makes a hydrogen bond to the main-chain NH of Gln39, while the other oxygen of this carboxylate

makes two further hydrogen bonds to the main-chain NH groups of T-loop residues 37 and 41 (Fig. 3). These 2-OG main-chain interactions help to stabilize the unique conformation of T-loop residues 37–41. In addition to these ionic and hydrogen-bonding interactions, there are hydrophobic interactions of the two methylene groups of 2-OG with the side chains of Thr43, Leu56, and Ile86, as well as with the α -carbon of Gly87. The conformation of bound MgATP and its interactions with the protein, except for those with T-loop residues, are very similar to those observed in *Methanococcus jannaschii* GlnK1 (MjGlnK1)⁹ and in the complex of *Arabidopsis thaliana* P_{II} with *N*-acetylglutamate kinase (NAGK).¹¹ However, the observed binding mode of 2-OG with its participation in Mg²⁺ coordination is completely different from the 2-OG binding reported for MjGlnK1,⁹ which we will discuss later.

We are confident that the 2-OG binding interactions observed in GlnZ will essentially be the same in other P_{II} proteins, as the arrangement observed here provides straightforward structural explanations for a number of biochemical data with different P_{II} proteins. First, direct effector molecule binding studies have shown a strong synergistic binding of ATP and 2-OG for several P_{II} proteins.^{3,4,12} Second, studies on P_{II} complexes with different Amt targets,^{9,13–15} with NtrB,¹⁶ and with NifL¹⁷ have shown a requirement for ATP, 2-OG, and Mg²⁺ (or Mn²⁺) in combination for effective complex dissociation or association. Third, the stereochemically precise recognition of the 2-oxo acid moiety makes biological sense. Fourth, the direct involvement of two highly conserved residues, Gln39 and Lys58,¹⁸ in 2-OG binding explains why *E. coli* GlnB-Q39E is strongly impaired in 2-OG binding¹⁹ and why Q39E and K58A of *Rhodospirillum rubrum* GlnJ, an *A. brasilense* GlnZ orthologue, are insensitive to the presence of (MnATP+2-OG) with regard to their target interactions.¹⁵ The additional negative charge of Q39E may disfavor its ligation to the Mg²⁺ site as

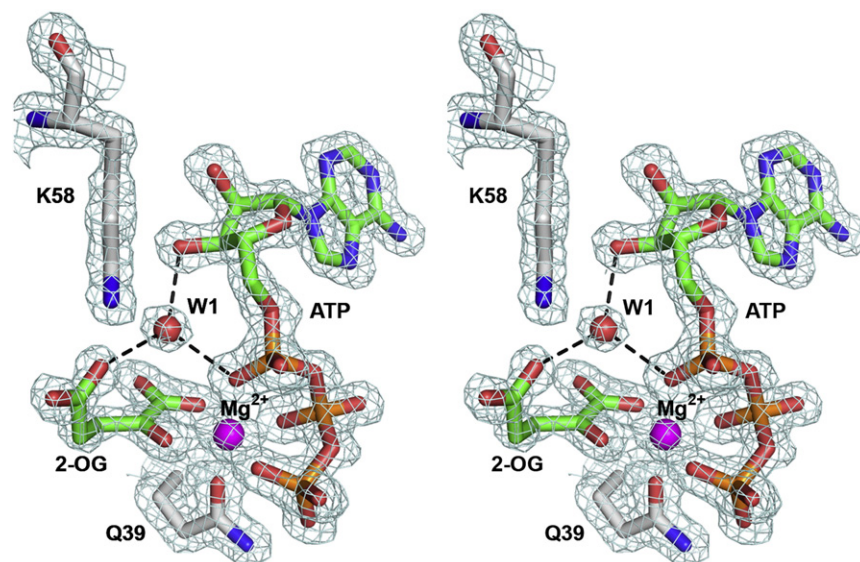


Fig. 2. Electron density at the 2-OG binding site. Stereo view of the final electron density ($2mF_o - DF_c$ map, contoured at 1.6σ) at the MgATP–2-OG binding site illustrating near-atomic resolution. The side chains of residues Q39 and K58 and a tightly bound water molecule W1 forming hydrogen bonds (broken lines) complement the binding interactions. For clarity, the rest of the protein that forms many more contacts with the shown effector molecules is not shown. The same arrangement is observed in all three crystallographically independent binding sites.

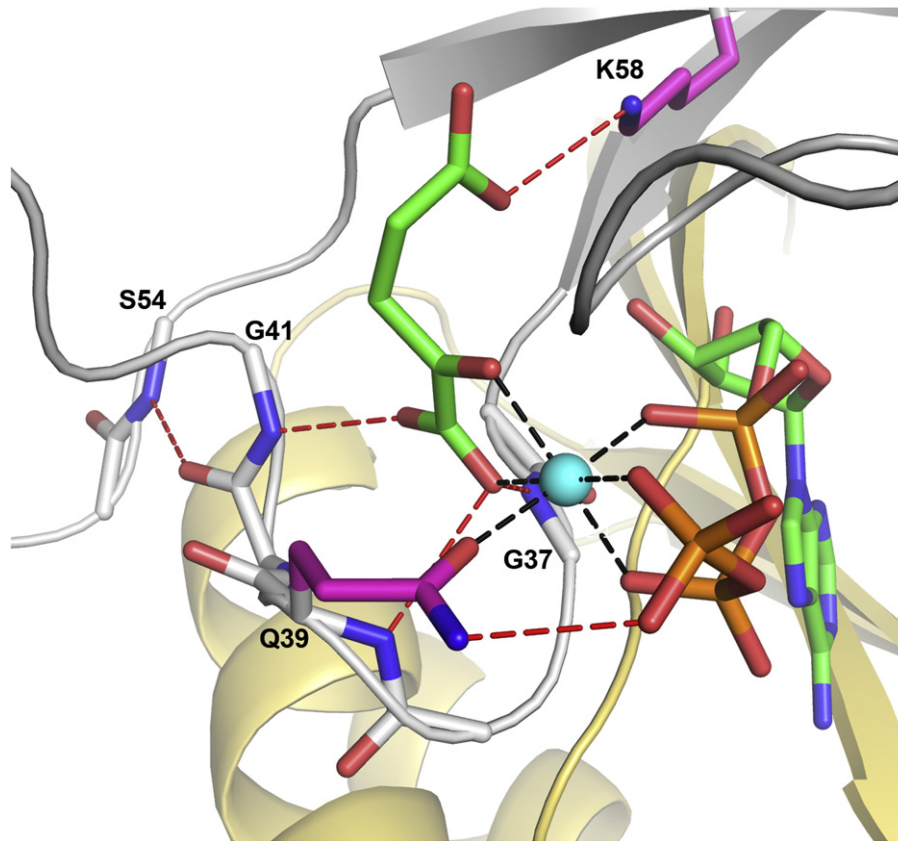


Fig. 3. The 2-OG binding site. 2-OG and ATP are shown in green stick representation, Mg^{2+} is shown as a cyan sphere, and its coordination is shown with black broken lines. Selected GlnZ residues and main-chain atoms are shown in magenta and gray sticks, respectively, and selected hydrogen bonds are indicated with red broken lines. GlnZ protomers are shown in gray and pale yellow cartoon representation.

ATP and 2-OG already provide excess negative charge. Furthermore, Q39E may also stabilize particular T-loop conformations, thereby influencing particular target interactions. Very recently, it has been shown that mutants I86T and I86N of a *Synechococcus elongatus* P_{II} protein are insensitive to 2-OG but still bind ATP and, based on these results, a direct interaction of the side chain of I86 with 2-OG, as observed here, has been postulated.²⁰

The 2-OG binding mode reported for MjGlnK1⁹ at the apex of a compactly folded T-loop is characterized by hydrogen-bonding interactions with main-chain and side-chain donors/acceptors of residues 52–54 and by stacking interactions with the conserved Tyr51 side chain (PDB ID 2J9E). The arguments as to why this structure is unlikely to represent a physiologically relevant interaction between 2-OG and MjGlnK1, or any other P_{II} proteins, are as follows. First, the 2-OG binding site might be expected to involve highly conserved P_{II} residues. In our structure, Gln39, Lys58, and Gly87 all satisfy this criterion (Fig. 4), but this is not so for the Asp54 contact in MjGlnK1.¹⁸ Indeed, the low crystallization pH of 4.6 used in those studies⁹ may have been important in permitting the formation of a hydrogen bond between the MjGlnK1-Asp54 side chain and the 5-carboxy group of 2-OG. Second, a T-loop deletion mutant $\Delta 47$ –53 of *E. coli* GlnB shows no impaired 2-

OG binding.¹⁹ Third, in PDB ID 2J9E, the 2-oxo group points to solvent and is not recognized by the protein, which is unexpected for specificity reasons. Moreover, glutamate and glutamine can be modeled very well into this binding site with similarly good or even better interactions. This is in conflict with data for *E. coli* GlnB, where glutamate is shown to be 10³-fold to 10⁴-fold less effective than 2-OG both in target activation and in competition with 2-OG binding, while glutamine is essentially ineffective.¹⁶ Fourth, nearly all eubacterial P_{II} proteins align with their C-terminus at residue 112, which could suggest a structurally defined role of the C-terminal carboxylate group. In our structure, the C-terminal carboxylate is hydrogen bonded to the adenine base of ATP and fits into the extended interaction network around the Mg^{2+} site. However, in the engineered MjGlnK1 protein, which carries a C-terminal tag (Leu-Glu-His₆), this structure is precluded, and three to five of the extra C-terminal residues are seen to form multiple interactions with the folded T-loop.

As recently reported,²¹ 2-OG also regulates the interaction of *A. thaliana* P_{II} with the biotin carboxyl carrier subunit of chloroplast, acetyl-CoA carboxylase (ACCase). This heteromeric enzyme is involved in the first steps of lipid biosynthesis, and the authors propose that plant P_{II} contributes to the fine-tuning of chloroplastic carbon flow to several

	10	20	30	40	50	60	
AbGlnZ	MKLVMAIIKP	FKLDEVREAL	TSLGIQGLTV	SEVKGFGGRQK	GQTEIYRGAE	YSVS-FLPKVK	
RrGlnJ	MKFVIAITKP	FKLDEVREAL	GALGIQGMTV	TEVKGFGGRQK	GQTEVYRGAE	YVFN-FLPKVK	
EcGlnB	MKKIDAIIKP	FKLDDVRELR	AEVGITGMTV	TEVKGFGGRQK	GHTELYRGAE	YMVD-FLPKVK	
EcGlnK	MKLVTVIIKP	FKLEDVREAL	SSIGIQGLTV	TEVKGFGGRQK	GHAELYRGAE	YSVN-FLPKVK	
SePII	MKKIEAIIKP	FKLDEVKIAL	VNAGIVGMTV	SEVRGFGGRQK	GQTERYRGSE	YTVE-FLQKVK	
MjGlnK1	MKKVEAIIKP	EKLEIVKVAL	SDAGYVGMTV	SEVKGGRVQGG	GIVERYRGRE	YIVD-LIPKVK	
AtPII	..FYKVEAIVRP	WRIQQVSSAL	LKIGIRGVTV	SDVRGFGAQQG	GSTERHGGSE	FSEDKFVAKVK	72
				* * * *		* *	

	70	80	90	100	110	
AbGlnZ	VEVAVSDDQY	EQVVEAIQKA	ANTGRIGDGK	IFVLDIAQAV	RIRTGETNTE	AL 112
RrGlnJ	IELAVPDNLL	DQVVEIQST	AQTGKIGDGK	IFVFDLGEAV	RIRTGERGEE	AL 112
EcGlnB	IEIVVPDDIV	DTCVDTIIRT	AQTGKIGDGK	IFVFDVARVI	RIRTGEBEDA	AI 112
EcGlnK	IDVAIADDQL	DEVIDIVSKA	AYTGKIGDGK	IFVAELQRVI	RIRTGEADEA	AL 112
SePII	LEIVVEDAQV	DTVIDKIVAA	ARTGEIGDGK	IFVSPVDQTI	RIRTGEKNAD	AI 112
MjGlnK1	IELVVKEDV	DNVIDIICEN	ARTGNPGDGK	IFVIPVERVV	RVRTKEEGKE	AL 112
AtPII	MEIVVKKQV	ESVINTIIEG	ARTGEIGDGK	IFVLPVSDVI	RVRTGERGEK	AE... 123
			**			

Fig. 4. Alignment of P_{II} amino acid sequences for *A. brasilense* GlnZ, *R. rubrum* GlnJ, *E. coli* GlnB and GlnK, *S. elongatus* P_{II}, *M. jannaschii* GlnK1, and *A. thaliana* P_{II}. The numbering above the alignment refers to the six bacterial sequences. N-terminal and C-terminal extensions in *A. thaliana* P_{II} have been omitted. The positions of the B-loops, C-loops, and T-loops are indicated above the alignment, and the residues identified in the GlnZ structure as contributing to the 2-OG binding site are identified by asterisks below the alignment.

metabolic pathways. ACCase activity is inhibited by P_{II} in the presence of MgATP. The inhibition results from P_{II} binding to the biotin carboxyl carrier subunit and can be released by 2-OG, with 50% recovery being achieved at ~0.1 mM 2-OG, which is within the physiological range. Interestingly, oxaloacetate and pyruvate were also shown to relieve P_{II} inhibition of ACCase, but this effect was only reported at a concentration of 5 mM, and titration of these metabolites was not reported.²¹ Oxaloacetate and pyruvate are both involved in acetyl-CoA synthesis and thus might conceivably act as effector molecules for plant P_{II} proteins.

These *in vitro* experiments and others with *E. coli* GlnB, which indicated that oxaloacetate and glutamate could potentially bind GlnB, albeit with much lower affinities than 2-OG,¹⁶ raise the possibility that 2-OG might not be the only physiologically relevant effector of P_{II} proteins. The binding of other effectors to the 2-OG site, as defined here, appears structurally plausible for oxaloacetate or pyruvate, which both share the 2-oxo acid moiety with 2-OG. However, lower binding affinities would be expected in agreement with the relatively high concentrations required for their *in vitro* effects. As stated previously, with *E. coli* GlnB, glutamate was 10³-fold to 10⁴-fold less effective than 2-OG.¹⁶ The inferred strongly reduced binding appears plausible as (i) the α -amino group of glutamate, substituting for the 2-oxo group of 2-OG, would almost certainly have to be deprotonated, and (ii) as indicated by modeling, the octahedral Mg²⁺ coordination would become significantly distorted.

T-loop conformations and target interactions

The mechanism by which effector molecule binding to P_{II} proteins regulates their interaction with target proteins is not understood in detail. However, a plausible hypothesis has been that their

binding changes the conformational energy landscape of T-loops, which are known to be the key determinants of target interactions.³ The 2-OG binding mode observed in this structure not only extends these observations but also allows the definition of key determinants of functional T-loop conformations. Even with MgATP and 2-OG bound, the tip of the T-loop appears to retain conformational flexibility. However, all three crystallographically independent copies show well-defined density and the same conformation up to residue 41 and after residue 54, characterized by a conserved main-chain hydrogen bond between residues 40 and 54 (Fig. 3). The conformation of the N-terminal and C-terminal T-loop residues enabling this intraloop main-chain interaction is seen in no other known P_{II} structure and appears uniquely related to the MgATP+2-OG form.

Superposition of the GlnZ structure with P_{II} structures involved in complexes with target proteins, namely AmtB and NAGK, suggests that interactions of the conserved Lys58 side chain characterize distinct families of functionally important T-loop conformations. In our new structure, its ammonium group forms a salt bridge with the 5-carboxy group of 2-OG (Figs. 3 and 5) together with a hydrogen bond to the main-chain carbonyl oxygen of the B-loop residue Gly87, with this latter interaction being observed in nearly all P_{II} structures. In the *E. coli* GlnK–AmtB complex, where ADP rather than ATP is bound to GlnK,²² the place of the 2-OG 5-carboxy group is taken by the carboxamide group of Gln39 (Fig. 5). However, in the P_{II}–NAGK complexes of *S. elongatus*²³ and *A. thaliana*,¹¹ the place of 2-OG is taken by the carboxylate group of Glu44 or the homologous Glu55, respectively (Fig. 5). Moreover, in the *A. thaliana* complex, which has MgATP bound to P_{II}, the main-chain carbonyl oxygen of Gly48 (Gly37 in GlnZ) is coordinated to Mg²⁺ and occupies the

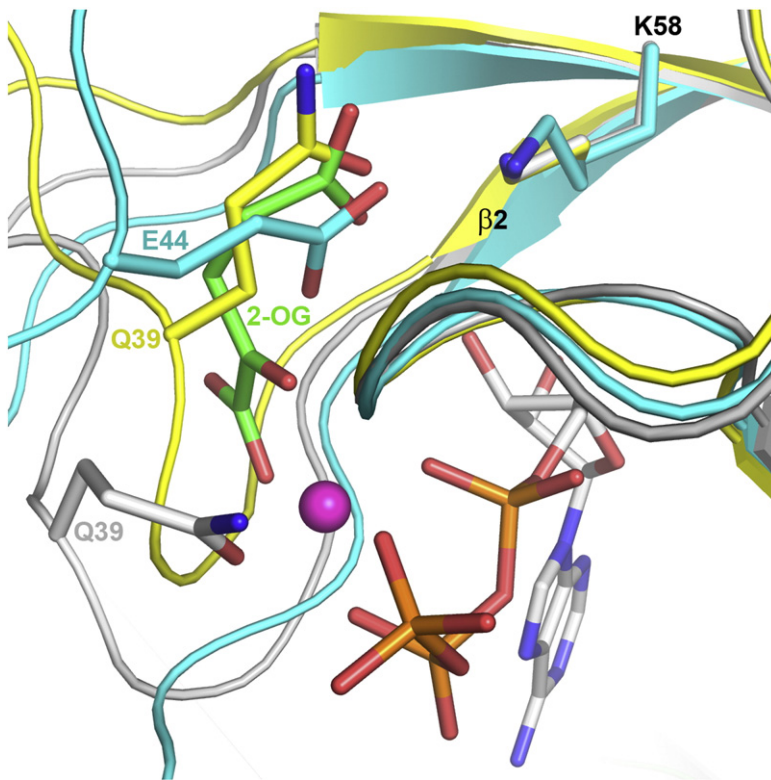


Fig. 5. Comparison of three functionally important T-loop conformations. K58 interacts with 2-OG, Q39, or E44 depending on effector molecule binding. (i) GlnZ (gray) with bound 2-OG (green) and MgATP (magenta/gray); (ii) *E. coli* GlnK (yellow) bound to ADP (not shown) in its complex with AmtB (PDB ID 2NUU); and (iii) the *S. elongatus* P_{II} protein (cyan) in its complex with NAGK (PDB ID 2V5H), which is likely to have MgATP bound *in vivo*. In the *A. thaliana* P_{II}-NAGK complex (PDB ID 2RD5), which has MgATP bound to P_{II}, the same intramolecular salt bridge is observed between Glu55 and Lys70. Note the diverging T-loops after strand β2 (Q39 of PDB ID 2V5H lies outside the image).

position of the coordinating carboxyl oxygen of 2-OG. Although bound MgATP appears not to be mandatory for *S. elongatus* P_{II} binding to NAGK *in vitro*,²³ its Gly37 is very similarly positioned in the complex. Comparison of the various interactions of Lys58 observed in each of these structures suggests that 2-OG binding is incompatible with T-loop conformations that engage Lys58 in interactions with T-loop residues.

In addition to ATP and 2-OG, it has recently been recognized that ADP also significantly affects P_{II} target interactions and their associated activities.^{15,22,24} These observations have led to the suggestion that P_{II} proteins may sense cellular adenylate charge through the ATP/ADP ratio.^{12,14,15} From a structural perspective, it is conceivable that each of the three binding sites of a trimer could be occupied by ADP, MgATP, or MgATP+2-OG. Based on existing structural information, we believe that these three functionally important effector binding modes are characterized by mutually exclusive interactions of Lys58 with Gln39, Glu44, or the 2-OG 5-carboxy group, respectively. Thus, for each of these three binding modes, the conformational space of the engaged T-loop is constrained in a very specific way, thereby significantly influencing the interaction of a P_{II} protein with its targets according to the particular effectors that are bound. For each such effector binding mode, the residual conformational flexibility of the T-loop, particularly at its tip, would still permit a range of conformations to occur in complexes with different targets. We cannot exclude, however, that in some cases target interactions may override the interactions of Lys58

with Gln39 or Glu44 in the ADP-bound and MgATP-bound states, respectively. It is perhaps important to note that both the GlnK–AmtB complex and the P_{II}–NAGK complex have a 3:3 P_{II}/target stoichiometry. Structures have yet to be determined for a P_{II} complex that has a different stoichiometry. Nevertheless, the discussed conformational constraints should also apply to such complexes, which *in vivo* should all have effector molecules bound to P_{II}.

Negative cooperativity between effectors

The synergistic binding of ATP and 2-OG is a common feature of P_{II} proteins. However, in *E. coli* GlnB, it appears limited to the first site of the trimer, and increasing negative cooperativity is observed for occupation of additional sites.¹² This has been proposed to be important for the regulation of its targets ATase and NRII (NtrB),^{4,12,25} which discriminate 2-OG/P_{II} stoichiometries of 1:3 and 3:3, respectively. Whether such negative cooperativity is a characteristic property of P_{II} proteins in general is not known. In the case of *Azotobacter vinelandii*, it is at best weak;¹⁷ for the *S. elongatus* P_{II} protein,²⁰ it is considerably smaller than for *E. coli* GlnB. At present, we have no information on whether anti-cooperative behavior is relevant to the biological function of GlnZ. The fully ligated trimer structure observed here shows no significant differences between the three crystallographically independent binding sites and gives no information on how the structure could be perturbed in a putative 1:3 complex. Resolving the structural basis of negative cooperativity will need additional biochemical and

structural data for P_{II} members that display this property very strongly.

In conclusion, the resolution of the 2-OG binding sites, within the lateral clefts of the P_{II} homotrimer and in close proximity to the site of ATP/ADP binding, brings considerable clarity to how these three effector molecules can interact to influence the conformation of P_{II} and, most specifically, of the T-loops. This information should now provide an excellent platform for future studies both of the structures of novel P_{II} complexes and of the links between cellular physiology and the signal transduction activities of these remarkably versatile proteins.

Materials and Methods

Protein and DNA

GlnZ from *A. brasilense* was expressed in *E. coli* BL21 (DE3) cells carrying pMSA4²⁶ at 37 °C overnight. The protein was purified using a heparin column, as described previously,¹⁴ followed by an additional gel-filtration step (Superdex 200) in a buffer containing 50 mM Hepes (pH 7.5) and 0.1 M NaCl (buffer A). GlnZ peak fractions were concentrated to 8 mg/ml for crystallization using ultrafiltration. Protein purity was assessed by SDS-PAGE, and molecular mass was confirmed by electrospray ionization mass spectrometry analysis.

Crystallization and data collection

Initial screening for crystals was performed using a Phoenix nanoliter-drop-dispensing robot with three commercial screens from Nextal (the Classics, JCSQ+, and PEG suites) and 200-nl drop sizes for protein and precipitant solutions. Protein solutions containing 10 mM MgCl₂ and different effector molecule concentrations (0.1/0.5, 0.5/0.1, 0.2/0.2, and 5/5 mM ATP and 2-OG, respectively) were prepared for crystallization. Crystals were observed under a number of similar conditions with different ATP/2-OG concentrations, typically in the pH range 6.5–7.5, and using different polyethylene glycols for precipitation. The crystal used for structure determination was grown by mixing a protein solution containing 8 mg/ml GlnZ, 10 mM MgCl₂, 0.1 mM ATP, and 0.5 mM 2-OG in buffer A with a precipitant solution containing 0.1 M 4-morpholineethanesulfonic acid (Mes; pH 6.5) and 40% (wt/vol) polyethylene glycol 200.

Diffraction screening and data collection for crystals mounted on cryoloops (directly from the crystallization drops) and frozen in a cold nitrogen gas stream at 100 K were carried out at beamline X10SA (PXII) of the Swiss Light Source at the Paul Scherrer Institut (Villigen, Switzerland). All crystals examined belonged to the same crystal form with diffraction up to 1.3 Å resolution in the best cases. Diffraction data ($\lambda=1.0000$ Å) were collected using a rotation angle of 0.5° per image on a MarCCD225 detector. Several data sets were collected and processed using XDS†.²⁷ XDS indexing indicated space group C222₁ in all cases, and processing in this space group yielded very good symmetry *R*-factors. The data

were further evaluated using phenix.xtriage,²⁸ and possible twinning was indicated by a value of 5.3 for the multivariate *Z*-score *L*-test, as well by the intensity statistics for acentric reflections (with Wilson ratio and moments about halfway between the ideal value for twinned data and the ideal value for untwinned data). Processing the data in lower-symmetry monoclinic space groups, however, did not significantly lower the symmetry *R*-factors.

Structure solution and refinement

We solved the structure in space group C222₁ by molecular replacement using the program Phaser²⁹ with the *E. coli* GlnK trimer as search model (PDB ID 2GNK). The initial model was corrected and improved in several rounds using automated restrained refinement with the program PHENIX³⁰ and interactive modeling with Coot.³¹ ATP and 2-OG, coordinated to a Mg²⁺, were identified in the first electron density map based on molecular replacement phases and were included in the model. The long T-loop (residues 37–55, found to be disordered in many P_{II} protein crystal structures) was found to be well defined in chains B and C, but partly disordered in chain A (residues 44–52). In subsequent rounds of refinement and interactive modeling, we identified one surface-bound Mes buffer molecule, a pentacoordinated Mg²⁺ hydrate, (indicated by about 2.1-Å distances between the central atom and its nearest-neighbor atoms and located on the intersection between a 2-fold axis along *b* and the noncrystallographic 3-fold), three short segments of polyethylene glycol chains and bound water molecules. Where clearly indicated by difference electron density, double conformations were chosen for a small number of residues. Anisotropic atomic displacement parameters were introduced for the protein and the bound MgATP, 2-OG, and Mes ligands. At the final stage, riding hydrogens (for protein residues, 2-OG, and Mes), as implemented in PHENIX, contributed to the calculated structure factors. The final *R*-factors were 0.162/0.198 for *R*_{work} and *R*_{free}, respectively, and refinement statistics are listed in Table 1.

The final model was analyzed using the program MolProbity.³² Ramachandran statistics show no outliers (Table 1). Figures 1, 2, 3, and 5 were prepared using PyMOL.³³

In view of the high resolution and the very good quality of the data, the final *R*-factors appear somewhat high. As the intensity statistics (phenix.xtriage) had indicated possible twinning, we explored pseudo-merohedral twinning in three possible monoclinic space groups (nonisomorphic subgroups C211, C121, and C112₁ of C222₁). In all cases, the twin fraction refined to 0.50, *R*_{work} and *R*_{free} improved quite substantially, and the two trimers retained essentially the same structure apart from some side-chain conformations of surface residues. The uncertainty about the twinning axis strongly indicated that the model of a twinned monoclinic crystal was unlikely to be correct. We suspect that further improvement of the orthorhombic refinement statistics is prevented by some unknown pathology of this crystal form. It is, however, a minor and largely technical issue and does not affect the structural model in a significant way.

PDB accession number

Coordinates and structure factors of the *A. brasilense* GlnZ complex with MgATP and 2-OG have been deposited under PDB ID 3MHY.

† http://xds.mpimf-heidelberg.mpg.de/html_doc/XDS.html

Acknowledgements

We thank Antonietta Gasperina for supervising part of the laboratory work of D.T. L.F.H. and L.S.C. received support from CNPq/INCT (Brazil). M.M. acknowledges support from the Biotechnology and Biological Sciences Research Council, UK.

References

- Arcondeguy, T., Jack, R. & Merrick, M. (2001). P_{II} signal transduction proteins, pivotal players in microbial nitrogen control. *Microbiol. Mol. Biol. Rev.* **65**, 80–105.
- Ninfa, A. J. & Atkinson, M. R. (2000). P_{II} signal transduction proteins. *Trends Microbiol.* **8**, 172–179.
- Forchhammer, K. (2008). P(II) signal transducers: novel functional and structural insights. *Trends Microbiol.* **16**, 65–72.
- Ninfa, A. J. & Jiang, P. (2005). P_{II} signal transduction proteins: sensors of alpha-ketoglutarate that regulate nitrogen metabolism. *Curr. Opin. Microbiol.* **8**, 168–173.
- Carr, P. D., Cheah, E., Suffolk, P. M., Vasudevan, S. G., Dixon, N. E. & Ollis, D. L. (1996). X-ray structure of the signal transduction protein P_{II} from *Escherichia coli* at 1.9 Å. *Acta Crystallogr. Sect. D*, **52**, 93–104.
- Nichols, C. E., Sainsbury, S., Berrow, N. S., Alderton, D., Saunders, N. J., Stammers, D. K. & Owens, R. J. (2006). Structure of the P_{II} signal transduction protein of *Neisseria meningitidis* at 1.85 Å resolution. *Acta Crystallogr. Sect. F*, **62**, 494–497.
- Huergo, L. F., Souza, E. M., Araujo, M. S., Pedrosa, F. O., Chubatsu, L. S., Steffens, M. B. & Merrick, M. (2006). ADP-ribosylation of dinitrogenase reductase in *Azospirillum brasilense* is regulated by AmtB-dependent membrane sequestration of DraG. *Mol. Microbiol.* **59**, 326–337.
- Wang, H., Franke, C. C., Nordlund, S. & Noren, A. (2005). Reversible membrane association of dinitrogenase reductase activating glycohydrolase in the regulation of nitrogenase activity in *Rhodospirillum rubrum*; dependence on GlnJ and AmtB1. *FEMS Microbiol. Lett.* **253**, 273–279.
- Yildiz, O., Kalthoff, C., Raunser, S. & Kuhlbrandt, W. (2007). Structure of GlnK1 with bound effectors indicates regulatory mechanism for ammonia uptake. *EMBO J.* **26**, 589–599.
- Xu, Y., Cheah, E., Carr, P. D., van Heeswijk, W. C., Westerhoff, H. V., Vasudevan, S. G. & Ollis, D. L. (1998). GlnK, a P_{II}-homologue: structure reveals ATP binding site and indicates how the T-loops may be involved in molecular recognition. *J. Mol. Biol.* **282**, 149–165.
- Mizuno, Y., Moorhead, G. B. & Ng, K. K. (2007). Structural basis for the regulation of N-acetylglutamate kinase by P_{II} in *Arabidopsis thaliana*. *J. Biol. Chem.* **282**, 35733–35740.
- Jiang, P. & Ninfa, A. J. (2007). *Escherichia coli* P_{II} signal transduction protein controlling nitrogen assimilation acts as a sensor of adenylate energy charge *in vitro*. *Biochemistry*, **46**, 12979–12996.
- Durand, A. & Merrick, M. (2006). *In vitro* analysis of the *Escherichia coli* AmtB–GlnK complex reveals a stoichiometric interaction and sensitivity to ATP and 2-oxoglutarate. *J. Biol. Chem.* **281**, 29558–29567.
- Huergo, L. F., Merrick, M., Pedrosa, F. O., Chubatsu, L. S., Araujo, L. M. & Souza, E. M. (2007). Ternary complex formation between AmtB, GlnZ and the nitrogenase regulatory enzyme DraG reveals a novel facet of nitrogen regulation in bacteria. *Mol. Microbiol.* **66**, 1523–1535.
- Teixeira, P. F., Jonsson, A., Frank, M., Wang, H. & Nordlund, S. (2008). Interaction of the signal transduction protein GlnJ with the cellular targets AmtB1, GlnE and GlnD in *Rhodospirillum rubrum*: dependence on manganese, 2-oxoglutarate and the ADP/ATP ratio. *Microbiology*, **154**, 2336–2347.
- Kamberov, E. S., Atkinson, M. R. & Ninfa, A. J. (1995). The *Escherichia coli* P_{II} signal transduction protein is activated upon binding 2-ketoglutarate and ATP. *J. Biol. Chem.* **270**, 17797–17807.
- Little, R., Colombo, V., Leech, A. & Dixon, R. (2002). Direct interaction of the NifL regulatory protein with the GlnK signal transducer enables the *Azotobacter vinelandii* NifL–NifA regulatory system to respond to conditions replete for nitrogen. *J. Biol. Chem.* **277**, 15472–15481.
- Sant'Anna, F. H., Trentini, D. B., de Souto, W. S., Cecagno, R., da Silva, S. C. & Schrank, I. S. (2009). The P_{II} superfamily revised: a novel group and evolutionary insights. *J. Mol. Evol.* **68**, 322–336.
- Jiang, P., Zucker, P., Atkinson, M. R., Kamberov, E. S., Tirasophon, W., Chandran, P. *et al.* (1997). Structure/function analysis of the P_{II} signal transduction protein of *Escherichia coli*: genetic separation of interactions with protein receptors. *J. Bacteriol.* **179**, 4342–4353.
- Fokina, O., Chellamuthu, V. R., Zeth, K. & Forchhammer, K. (2010). A novel signal transduction protein P_{II} variant from *Synechococcus elongatus* PCC 7942 indicates a two-step process for NAGK–P_{II} complex formation. *J. Mol. Biol.* 2010 Apr 15. [Epub ahead of print].
- Feria Bourrellier, A. B., Valot, B., Guillot, A., Ambard-Bretteville, F., Vidal, J. & Hodges, M. (2010). Chloroplast acetyl-CoA carboxylase activity is 2-oxoglutarate-regulated by interaction of P_{II} with the biotin carboxyl carrier subunit. *Proc. Natl Acad. Sci. USA*, **107**, 502–507.
- Conroy, M. J., Durand, A., Lupo, D., Li, X. D., Bullough, P. A., Winkler, F. K. & Merrick, M. (2007). The crystal structure of the *Escherichia coli* AmtB–GlnK complex reveals how GlnK regulates the ammonia channel. *Proc. Natl Acad. Sci. USA*, **104**, 1213–1218.
- Llacer, J. L., Contreras, A., Forchhammer, K., Marco-Marin, C., Gil-Ortiz, F., Maldonado, R. *et al.* (2007). The crystal structure of the complex of P_{II} and acetylglutamate kinase reveals how P_{II} controls the storage of nitrogen as arginine. *Proc. Natl Acad. Sci. USA*, **104**, 17644–17649.
- Jiang, P. & Ninfa, A. J. (2009). Sensation and signaling of alpha-ketoglutarate and adenylate energy charge by the *Escherichia coli* P_{II} signal transduction protein require cooperation of the three ligand-binding sites within the P_{II} trimer. *Biochemistry*, **48**, 11522–11531.
- Jiang, P. & Ninfa, A. J. (2009). Alpha-ketoglutarate controls the ability of the *Escherichia coli* P_{II} signal transduction protein to regulate the activities of NRII (NtrB) but does not control the binding of P_{II} to NRII. *Biochemistry*, **48**, 11514–11521.
- Araujo, L. M., Huergo, L. F., Invitti, A. L., Gimenes, C. I., Bonatto, A. C., Monteiro, R. A. *et al.* (2008). Different responses of the GlnB and GlnZ proteins upon *in vitro* uridylylation by the *Azospirillum brasilense* GlnD protein. *Braz. J. Med. Biol. Res.* **41**, 289–294.

27. Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**, 795–800.
28. Zwart, P. H., Grosse-Kunstleve, R. W. & Adams, P. D. (2005). Characterization of X-ray data sets. *CCP4 Newsl.* **42**.
29. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.
30. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W. *et al.* (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. Sect. D*, **58**, 1948–1954.
31. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D*, **60**, 2126–2132.
32. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X. *et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383.
33. DeLano, W. L. (2002). *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA. <http://www.pymol.org>.