Eur. J. Biochem. **268**, 2028–2037 (2001) © FEBS 2001

The structure of the P_{II}–ATP complex

Yibin Xu[1], Paul D. Carr[1], Thomas Huber[2], Subhash G. Vasudevan[3] and David L. Ollis[1]

[1]*Center for Molecular Structure and Function, Research School of Chemistry, Australian National University, Canberra, Australia;* [2]*Supercomputer Facility, Australian National University, Canberra, Australia;* [3]*Department of Biochemistry and Molecular Biology, James Cook University, Townsville, Queensland, Australia*

P_{II} is a signal transduction protein that is part of the cellular machinery used by many bacteria to regulate the activity of glutamine synthetase and the transcription of its gene. The structure of P_{II} was solved using a hexagonal crystal form (form I). The more physiologically relevant form of P_{II} is a complex with small molecule effectors. We describe the structure of P_{II} with ATP obtained by analysis of two different crystal forms (forms II and III) that were obtained by co-crystallization of P_{II} with ATP. Both structures have a disordered recognition (T) loop and show differences at their C termini. Comparison of these structures with the form I protein reveals changes that occur on binding ATP. Surprisingly, the structure of the P_{II}/ATP complex differs with that of GlnK, a functional homologue. The two proteins bind the base and sugar of ATP in a similar manner but show differences in the way that they interact with the phosphates. The differences in structure could account for the differences in their activities, and these have been attributed to a difference in sequence at position 82. It has been demonstrated recently that P_{II} and GlnK form functional heterotrimers *in vivo*. We construct models of the heterotrimers and examine the junction between the subunits.

Keywords: ATP binding protein; GlnK; P_{II}; signal transduction; X-ray structure.

P_{II} is a signal transduction protein found in a variety of organisms. It is best characterized in *Escherichia coli* in which it indicates the nitrogen status of the cell. In many bacteria the level of available nitrogen is closely coupled to the level of glutamine synthetase (GS) activity. P_{II} is involved in the regulation of GS and the transcription of its gene [1–4]. In brief, the nitrogen status of the cell is monitored by the uridylyl transferase/uridylyl removing enzyme (UT/URase). It uridylylates P_{II} at Tyr51 to form P_{II}–UMP [5–7] under conditions of limiting ammonia and catalyses the removal of UMP under conditions of excess ammonia. The uridylylation status of P_{II} reflects the nitrogen status of the cell. P_{II} dictates the direction of the enzymatic regulation of GS through its interaction with adenylyl transferase (ATase). The transcriptional regulation of GS is mediated by P_{II} and the nitrogen regulatory proteins I and II (NRI and NRII) also known as NtrC and NtrB, respectively. Although the regulation of GS has been studied for many years, much remains to be learnt about this system in the light of some recent discoveries. For example, it is now apparent that many organisms possess two P_{II}-like proteins. The second protein in *E. coli* is known as GlnK [8,9]. The physical properties of GlnK are very similar to those of P_{II} and under some circumstances GlnK can substitute for P_{II}. When the cell is grown in nitrogen-poor conditions, P_{II} and GlnK form heterotrimers [10] to fine-tune the regulation of GS [11]. The fine regulation of the nitrogen signal cascade by functional heterotrimers of P_{II} and GlnK introduces a new level of complexity to the regulation of GS. The system is even more complex with both P_{II} and GlnK binding effector molecules α-ketoglutarate (α-KG), and ATP [12]. The former will not bind specifically to P_{II} in the absence of the latter and the affinity of P_{II} for ATP increases in the presence of α-KG. In *E. coli*, under normal conditions, P_{II} exists predominantly as a ternary complex. This is consistent with *in vitro* studies showing that the ternary complex of P_{II} is the substrate of UT/URase and the form of P_{II} that binds NRII.

The structure of P_{II} was determined some time ago [13,14] while the structure of GlnK along with its ATP complex have been described more recently [15]. The structures of the two proteins are similar. Both proteins form trimers in which the individual monomers are related by a threefold axis. In GlnK, ATP binds in a cleft at the interface of two monomers. In this paper we describe the structure of P_{II} with ATP as determined by the structure analysis of two crystal forms (forms II and III) obtained by co-crystallization. Both structures are well determined and show significant differences. We make detailed comparisons of the two structures to show that the differences between the structures are localized to specific regions of the molecule and we comment on the functional significance of these differences. We compare the P_{II} and GlnK structures with an emphasis on how they bind ATP. It is known that P_{II} and GlnK can form heterotrimers *in vivo* [11]. We propose models for the structures of the heterotrimers.

Correspondence to D. Ollis, Research School of Chemistry, Australian National University, GPO 414, Canberra ACT 2061, Australia. Fax: + 61 26125 0750, Tel.: + 61 26125 4377, E-mail: Ollis@rsc.anu.edu.au

Abbreviations: GS, glutamine synthetase; UT/URase, uridylyl transferase/uridylyl removing enzyme; ATase, adenylyl transferase; SA, simulated annealing; NRI and NRII, nitrogen regulatory proteins I and II, also known as NtrC and NtrB; α-KG, α-ketoglutarate; α-OG, oxolylglycine; r.m.s., root mean square.

(Received 14 September 2000, revised 22 January 2001, accepted 6 February 2001)

MATERIALS AND METHODS

Crystallization

The overexpression and purification of P_{II} has been described previously [16]. The protein was stored in buffer A (20 mM Hepes pH 7.5, 1 mM 2-mercaptoethanol, 1 mM EDTA). Two crystal forms, forms II and III, were produced by co-crystallizing P_{II} with ATP and α-KG. Along with the biologically relevant effector, α-KG, an analogue, disodium oxolylglycine (α-OG) was also used in some experiments. This latter material was prepared as described by Rife and Cleland [17].

The form II crystals were obtained at 277 K with the protein concentrated to 18–25 mg·mL^{-1} and stored in buffer A with 20 mM spermidine and a fourfold molar excess of ATP and α-KG. The form II crystals were prepared from a solution that contained sodium acetate, poly(ethylene glycol) 3400 and Hepes pH 7.5 as described previously [18]. The space group was I222. Form III crystals have not been described previously and were obtained at 298 K by using the standard hanging drop technique in which equal volumes of protein and reservoir solution are mixed together. These crystals were produced with both P_{II}/ATP and P_{II}/ATP/α-OG complexes with reagents 1 and 24 from the Hampton Kit1 (Hampton Research, Laguna Niguel, CA, USA). Large crystals of P_{II}/ATP/α-OG (0.4 mm in each dimension) were obtained from 5–10% isopropanol or 2-methyl-2,4-pentanediol and 200 mM $CaCl_2$ buffered in 100 mM sodium acetate pH 4.7. Large crystals of P_{II}/ATP/α-KG were also obtained under these conditions. The crystals were stable in the X-ray beam at room temperature and suitable for high resolution structure analysis. The form III crystals belong to space group $I2_13$. The concentration of Ca^{2+} far exceeded the concentration of ATP and other ligands so that most of the ATP in the solution existed as a complex with Ca^{2+} (form IIIa). It was possible that this complex would not bind to P_{II}; indeed, the first structure obtained with this crystal form did not show any indication of ATP. Although the formation of large crystals required high levels of Ca^{2+}, these crystals could be transferred to solutions with lower concentrations of Ca^{2+} and higher concentrations of ATP (form IIIb).

Data collection

X-ray diffraction data were collected at Rigaku RAXIS-IIC imaging plate detector mounted on a Rigaku RU200 rotating-anode X-ray generator operating at 50 kV and 100 mA. Data for both crystal forms were processed using the HKL program [19] and the results are shown in Table 1. Data for the form II crystals were collected at 100 K as described previously [18]. The form III crystals were mounted in glass capillaries and data were collected at 293 K. The first data set (form IIIa) was collected with a crystal that had been transferred to the reservoir buffer of the crystallization experiment while the second (form IIIb) was collected with a crystal soaked in 10 mM ATP, 10 mM $CaCl_2$, 10% isopropanol and 100 mM sodium acetate pH 4.7.

Structure determination and refinement

The form II and III structures were solved by molecular replacement methods. Rotation and translation functions as well as rigid body refinement were done with programs in the AMORE package [20]. The structure refinement was performed with the program XPLOR 3.1 [21]. The progress of refinement was monitored by using the cross-validation indicator R_{free} values [22] calculated with 5% of reflections with Fo $>$ σ_F. Electron density maps were viewed with FRODO/TOM [23]. Solvent molecules were identified in 2Fo–Fc and Fo–Fc maps. Water molecules were included if they could be built into density that was higher than 3 σ in the Fo–Fc map and were in chemically reasonable positions. Water molecules were removed from the model if the *B* values refined to greater than 60 Å^2. The final models for both structures were produced after several rounds of model building and energy minimization followed by

Table 1. Statistics.

Crystals	Form II (ATP)	Form IIIa (no ATP)	Form IIIb (ATP)
Spacegroup	I222	$I2_13$	$I2_13$
Cell [a, b, c (Å)]	87.6, 87.8, 91.4	90.3	90.5
Subunits in ASU	3	1	1
Data collection			
Temperature (K)	100	293	293
Crystal to detector distance (mm)	100	90	90
Number of oscillations	40	40	40
Oscillation angle (°)	3	1	1
Resolution range (Å)	25–2.05	25–2.2	25–1.8
No. of observations	102195	55269	94739
No. of unique observations	21941	6359	11601
Completeness (%)	97.7	100	99.8
R_{merge} (%)	5.0	6.9	4.4
Outer shell (Å)	2.12–2.05	2.28–2.20	1.86–1.80
Completeness (%)	97.2	100	99.6
R_{merge} (%)	15.6	24.4	29.0

individual B factor refinement. The refinement stopped when R_{free} converged.

Structure determination with the form II crystals

The crystals contained three independent subunits arranged around a noncrystallographic threefold axis. The threefold axis was found near to the body diagonal of the cell so that the crystals were approximately cubic. The search model contained a P_{II} trimer with the T-loops deleted. The initial difference maps showed no density for either the T-loop or the last four residues of the protein. The starting model for refinement contained residues 1–35 and 56–108 for all three subunits. The first round of simulated annealing (SA) from 3000 K to 300 K brought down R_{cryst} from 44.8% to 30.7% and R_{free} from 45.2% to 37.0% against the data with Fo $>$ 2 σ(F) between 6.0 and 2.05 Å resolution. 2Fo–Fc and Fo–Fc maps showed very good density for three ATP molecules. At this stage, three ATP molecules were added to the model. Further SA refinement (2000 K to 300 K) reduced R_{cryst} to 26.8% and R_{free} to 32.7%. There was continuous electron density for residues 109–112 of all subunits in Fo–Fc maps. After rebuilding these C-terminal residues into the model, another round of SA followed by individual B factor refinement resulted in R_{cryst} of 24.9% (R_{free} = 30%). At this time numerous water molecules were added.

Structure determination with the form III crystals

The P_{II} structure obtained with the form I crystals was used as the starting model for molecular replacement calculations. This model consisted of a single subunit with the T-loop removed. After rigid body refinement, a correlation value of 61.8% and *R* value of 38.3% was obtained for data between 8 and 3 Å resolution. The model gave bad contacts between residues 109–112 and symmetry-related molecules. These four residues were discarded from the model and the *R* factor dropped to 37.1%.

Early in the refinement difference Fourier maps indicated that there was a calcium ion located on a crystallographic twofold axis. Problems were experienced in refining the B factor of an atom on a special position. A reasonable value was obtained by changing its value manually and looking at the consequent effects on difference Fourier maps. The final B value for the calcium was 24 $Å^2$. There was no density for an ATP molecule in the difference maps generated during the refinement of form IIIa. The crystal structure showing ATP (form IIIb) was refined with the form IIIa structure (without solvents and calcium) used as the starting model. Before refinement, calculated difference Fourier maps showed very good density for all the atoms in the ATP molecule except for the γ phosphate.

The models of P_{II}/GlnK heterotrimers

Although there are no experimental P_{II}/GlnK heterotrimer structures, the known structures of P_{II} and GlnK are very similar and a reasonable model can be constructed by superimposing the GlnK monomer on the P_{II} subunits in the trimer. This was done with the program LSQKAB in the CCP4 suite of programs [24]. For these calculations, sequence alignments previously published were used [15].

Qualitatively, the stability of protein complexes can be assessed by simple energetic considerations. The energies of the heterotrimer models could be calculated with a classical molecular mechanics force field. We have chosen a different approach for two reasons. First, it is not a trivial task to model the monomer–monomer interfaces in a heterotrimer so that all side chains pack correctly. This is true even in this seemingly easy case where both homotrimer structures are known and the sequence identity between the two subunits is very high. Second, energies from a classical molecular mechanics force field are very sensitive to errors in the structural model. A small displacement in coordinates generally gives rise to a huge energy difference. As a result, the energies obtained with molecular mechanics may not be very meaningful. The use of a low resolution force field alleviates these problems. In this work we use the scoring scheme from the SAUSAGE package [25]. This approach does not rely on side chain conformations and is insensitive to small structural inaccuracies. The force field used in these calculations was developed for the predicting the most likely chain packing for a given amino-acid sequence (the so called threading approach to structure prediction). Although this force field is not based strictly on physics, it does reflect the suitability of an amino acid for a particular environment and is well suited for the present application.

For each of the four distinct trimers ($^3P_{II}$, $^2P_{II}$ ^{1}GlnK, $^1P_{II}$ ^{2}GlnK, ^{3}GlnK) an energy of trimerization was calculated as the difference between the trimer energy and the sum of energies from the constituting monomers in isolation. To check the sensitivity of results with respect to the model, calculations are performed with the amino-acid sequences threaded onto both the P_{II} and GlnK structures (ATP-bound with T-loop missing). The trend in energies of trimerization were the same in both sets of calculations.

RESULTS AND DISCUSSION

The structure of P_{II} with bound ATP has been obtained in two different crystal forms. The form II crystals were fragile and obtained at pH 7.5 with poly(ethylene glycol). Crystal formation required ATP and was successful only at 276 K. The form II crystals were not stable in the X-ray beam at ambient temperatures. The form III crystals were quite sturdy and obtained at 298 K from organic solvents buffered at pH 4.7. The crystals were stable in the X-ray beam for extended periods at 293 K.

Quality of the models

The details of the final model obtained with the form II crystals are given in Table 2. The model includes residues for the three subunits, designated A, B and C. The T-loop is disordered in all three subunits (residues 38–53 in the A and B and residues 40–53 in C). The side chains of Lys85A, Glu18B, Asp54B, Arg103B, Ile112B and Glu109C have poorly defined electron density and have been incorporated as alanine into the model. There was poor density for a number of other amino acids, consequently Glu109A, Ile112A, Glu109B, Arg38C and Gln39C were built as glycine residues. As shown in Table 2, model geometry is good. The program PROCHECK [26] indicates that 94.3% of the ϕ, ϕ angles of all nonglycine residues are in the most

Table 2. Results of structure refinement.

	Form II	Form IIIa	Form IIIb
Diffraction data			
Resolution range (Å)	6.0–2.05	6.0–2.2	6.0–1.8
R_{cryst}/R_{free} ($\|F\| \geq 2\sigma$)	0.206/0.250	0.188/0.248	0.192/0.228
Model			
No. of nonhydrogen atoms	2148	687	699
No. of water molecules	134	25	23
No. of acetates	–	–	1
No. of ATP	3	–	1
No. of Ca^{2+}	–	1	1
Average B factors (Å^2)			
Protein	21	27	24
Solvent	29	35	31
ATP	20	–	35
Acetate	–	–	32
Calcium	–	24	15
R.m.s. deviations from ideal values			
Bond length (Å)	0.011	0.011	0.011
Bond angle (°)	1.844	1.622	1.657

favoured regions of the Ramachandran plot. The only apparent outlier in the Ramachandran plot is Asp108. The conformation of this amino acid is the same in each of the three independent copies of P_{II} and is similar to the conformation of Asp108 observed in two crystal forms of GlnK [15].

Two structures of P_{II} were obtained with the form III crystals. Both structures were well determined as indicated by the statistics given in Table 2. The last three residues at the C terminus and the T-loop (residues 38–53) were disordered in both models. In the absence of ATP, water molecules occupy the cleft. The form IIIb model has an ATP molecule along with an acetate in the cleft. The methyl of the acetate makes contact with the adenine of the ATP and the ring of Phe92. The acetate also forms a hydrogen bond with the side chain carboxylate of Asp5. Normally two anions would repel each other, but at low pH at least one is not ionized and hence a hydrogen bond becomes possible. This kind of hydrogen bond has been found in other proteins at low pH conditions [27]. It should be noted that several data sets were collected with crystals soaked in ATP and an excess of 2-KG or α-OG (data not shown). After refinement, difference Fouriers gave clear signals for ATP but no indication of the T-loop, 2-KG or α-OG.

A description of the P_{II}/ATP complexes

The three subunits in the form II crystals are very similar: the root mean square (r.m.s.) deviations for the least squares superposition of the Cα atoms are 0.26 Å, 0.23 Å and 0.26 Å for A to C, B to C and A to B, respectively. There are a number of significant differences. Specifically, the last few residues of the protein differ by ≈0.6–0.8 Å. The C terminus is close to residue 68 causing differences of ≈0.6 Å in this latter region. The B-loop residues (residues 80–90) differ by ≈0.4 Å. The differences in the subunit structures are consistent with their B factors. B factors for residues near the end of the peptide, around residue 68 and those in the B-loop are significantly higher than the average for the protein. The two structures obtained with the form III crystals exhibit no systematic differences and their B factors do not vary systematically. An r.m.s displacement of 0.22 Å was obtained when the Cα atoms for the form IIIa and form IIIb structure were superimposed.

The structures observed for P_{II} in forms II and III crystals are very similar. The r.m.s deviation for the superposition of form IIIb and the subunit C in form II is 0.68 Å for the Cα atoms of residues 1–36 and 54–108. This deviation goes up to 1.7 Å if residues 109 and 110 are included in the comparison. A plot of the distance between the Cα positions of P_{II} as found in form II and the form III crystals is given in Fig. 1. The two structures differ significantly in their B-loops and C terminus. Specifically, residues 85 and 86 of the B-loop have moved by ≈2.0 Å while residues at the C terminus have shifted by up to 2.5 Å. Some of these differences can be explained by the way the molecules pack in the crystals.

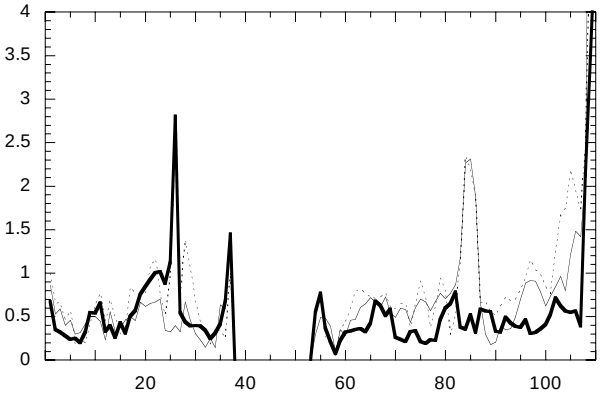

Fig. 1. A plot of distance (Å) between the Cα atoms as a function of residue number. The three lines show form IIIb and form IIC (protein plus ATP; thick line), form IIC and form I (native; broken line) and form IIIb and native P_{II} (thin line).

The crystal packing of the two crystal forms is quite different. Form III crystals have cubic symmetry whereas the form II crystals approximate to cubic symmetry. The cell dimension of the two crystal forms are very similar. However, I222 is not a subgroup of $I2_13$ and the contacts made by the molecules in the two crystal forms are unrelated. For form II, the contacts made by the subunit A are water mediated − there are no direct hydrogen bonds with other molecules. In some cases, charged residues on the surface of the subunit A make water-mediated hydrogen bonds with other subunits. For example, the side chains of Glu22A and Lys58B both form hydrogen bonds with a water molecule, W220. In other cases, side chains of subunit A make nonbonded contact with residues in other subunits. For example, the side chain of Phe11A makes contact with Gly24C. The start and end of the T-loops of all three subunits were found on the border of a large solvent cavity with little to stabilize the T-loop.

Crystal packing in the form III crystals is important for understanding the structure of the molecule. The molecule lies on a crystallographic threefold axis. The T-loop is disordered with the residues at its start and end bordering a solvent channel as was also observed for P_{II} in form II crystals and GlnK in the GlnK/ATP complex. The crystals are stabilized by contacts made between subunits related by the crystallographic twofold axes. There are two important contact areas, one involving a calcium ion located on a crystallographic twofold axis and surrounded by a coordination sphere of the oxygen atoms (Asp14 and Asp15 along with the carbonyl oxygen of Phe11 of helix A). The second site of contact involves residues near to the C terminus. The side chain of Asp108 forms a hydrogen bond with its crystallographic equivalent while Glu107 and Asp09 are linked to Asp109 and Glu107, respectively, in a crystallographically related molecule. Clearly, contacts made in the form III crystals are not likely to be preserved in solution as they depend upon a high concentration of Ca^{2+} and a low pH. In this case packing interactions have had an effect on the ATP conformation. Asp106 makes an interaction with both the N-terminus and Asp67. This shifts the position of Asp68 that interacts with Arg101 which in turn makes a contact with the terminal phosphate of ATP. Thus, the conformation of Arg101 differs between the form II and form III crystals and is probably responsible for the difference in the position of the γ phosphate of ATP that is observed in these crystal forms.

Comparison of P_{II} with and without ATP

The r.m.s. deviation obtained when superimposing the native structure with that of the subunit C in form II is 0.76 Å for the Cα atoms of residues 1–37 and 54–108. This values is clearly higher than those obtained in the comparisons of the form II subunits. As shown in Fig. 1, there are three regions where significant differences are observed between the native and the ATP-bound forms of the protein. The first is around residue 26. Here the native molecule differs from both ligand-bound variants found in forms II and III. The structure of the B-loop in the form II crystals differs from that in the form I (native) and form III crystals which are very similar. The third region where the structures differ is around the C terminus. In this case all three structures differ from one another. Some explanation for these differences can be made.

The differences around Thr26 can be explained by the ordering of the T-loop in the native structure. The side chain of Leu45 in the T-loop of the native structure would collide with Thr26 of form II. Furthermore, Thr26 in the native structure would have to move to accommodate ATP binding. The molecule in the form III crystals has a disordered T-loop and Thr26 is in the same position irrespective of whether ATP is present or not. This suggests that even in the absence of ATP, the peptide around Thr26 can assume a conformation that allows ATP ready access to its binding site.

The structure of the B-loop is coupled with ATP binding and the conformation of amino acids near to the C terminus. The structure of the B-loop of GlnK changes on binding ATP [15] and a similar change is apparent from the comparison of the native (form I) and ATP-bound form (form II) of P_{II}. The structure of the B-loop in the form III crystals is the same in the presence and absence of ATP. The way ATP binds to the protein in form III differs from that seen in form II, due in part to a change in the conformation of the side chain of Arg101.

The differences in the conformation of the C terminus may be due to crystal packing effects. As already noted, the conformation observed in the form III crystals is dependent upon crystal packing. A close inspection of the native (form I) structure reveals that the C terminus makes a number of contacts with symmetry-related molecules. For example, Asp109 makes a hydrogen bond with the side chain of Lys9 in an adjacent molecule. It appears that the conformation of the C terminus in the form I crystals are stabilized, in part, by lattice contacts. Such is not the case for the conformation of the same peptide in the form II crystals. Further, this conformation is very similar to that observed in the two crystal forms of GlnK [15]. All of these observations suggest that the C-terminal residues of P_{II} are capable of adopting a number of conformations and the conformation of the protein in the form II is probably the most stable in solution.

The interaction of P_{II} and ATP in the form II crystals

The model obtained with the form II crystals includes three ATP molecules. The interactions made by the protein and the γ phosphate of the ATP molecule between monomers A and B are not clear as the side chain of Arg103B is disordered. The other two ATP molecules have almost the same interactions with the protein. The present discussion will be restricted to the ATP between the A and C peptides. The adenine base and ribose ring of the ATP molecule adopts the same conformations as the ATP in the GlnK/ATP complex described previously [15]. As shown in Fig. 2, the interactions between P_{II} and the base and sugar of the ATP are very similar in both forms II and III. The P_{II} residues that interact with the phosphate groups come from B-loop and the adjacent C-loop. The α phosphate does not have any direct contacts with the protein but is stabilized by two solvent molecules. The β phosphate in the form II makes two hydrogen bonds with the amide nitrogen atoms of Asp88 and Gly89 and a salt link with Lys90. In this crystal form, Lys90 Nζ also forms a long hydrogen bond (≈ 3.2 Å) with the oxygen atom that links the α and β phosphates. In form II the γ phosphate of ATP is stabilized primarily by

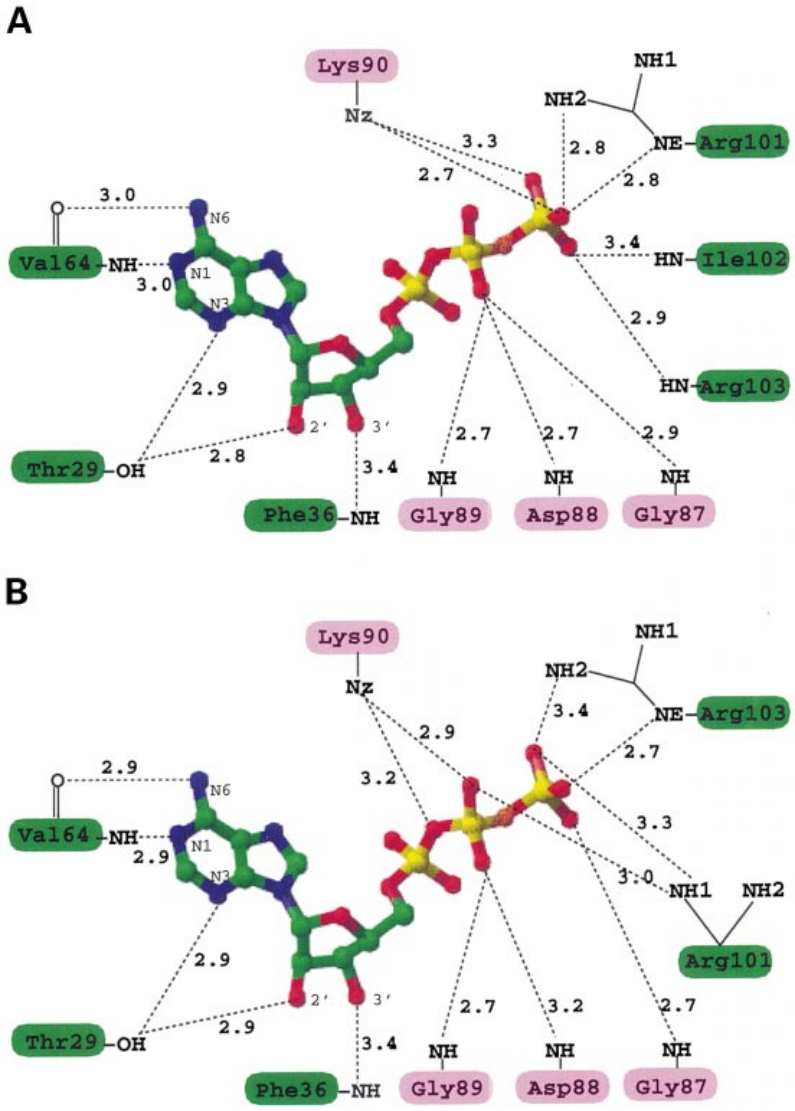

Fig. 2. The schematic diagrams of ATP interactions with the P_{II} protein (A) in the form III crystals and (B) in the form II crystals. ATP is coloured by atom types: carbon, green; nitrogen, blue; oxygen, red; sulphur, yellow. Residues from different monomers are coloured light-green and pink. Hydrogen bonds and lengths are labelled. The interactions between ATP and water molecules in the form II crystals are omitted.

the guanidino groups of Arg101 and Arg103 as well as the amide nitrogen of Gly87. The γ phosphate in form II also forms hydrogen bonds with water molecules. When the P_{II} molecules in forms II and III are superimposed the distance between the γ-P atoms is found to be 3.7 Å. This difference can be attributed to the way the molecules pack in the crystal. It is expected that in solution ATP takes on the conformation observed in the form II crystals.

Comparison GlnK and P_{II}–heterotrimer formation and ATP binding

As shown in Fig. 3A, the overall structures of P_{II} and GlnK are essentially the same. Yet, understanding how these two proteins might form heterotrimers is an extremely difficult task due to the enormous number of possible interactions involved. There are 46 residues in P_{II} and GlnK that experience a change in local environment on trimer formation. The overall environmental change of all residues can be seen in the calculated energy differences given in Table 3 that show a clear trend. The trimer stabilization energies for the P_{II} and GlnK homotrimers differ to the greatest extent while the values obtained for the heterotrimers are intermediate. The calculations show that the amino acids in GlnK are compatible with the local environments found in the P_{II} trimer and the structure of the heterotrimers are probably quite similar to those of P_{II} and GlnK.

P_{II} and GlnK bind the base and sugar of ATP in a very similar manner, but they differ in the way they accommodate the phosphates. The following differences were noted for the base and sugar. In the GlnK/ATP complex there is a hydrogen bond between adenine N7 with Lys90 Nζ that is not present in the P_{II} complex. In addition, the hydrogen bond between the ribose 3′-OH and Phe36 N found in the P_{II} complex is not present in that of GlnK. N7 and 3′-OH shift by ≈ 1.0 Å and ≈ 0.8 Å, respectively, in going from GlnK to P_{II}. It is surprising that the triphosphate conformations (Fig. 3B) differ in the two proteins as the residues forming the ATP binding site in GlnK and P_{II} are very highly conserved in all P_{II}-like proteins [15]. The differences can be explained by one important sequence variation. At position 82 there is a Gln in P_{II} and a Tyr in GlnK. In P_{II} the side chain of Gln82 makes a hydrogen

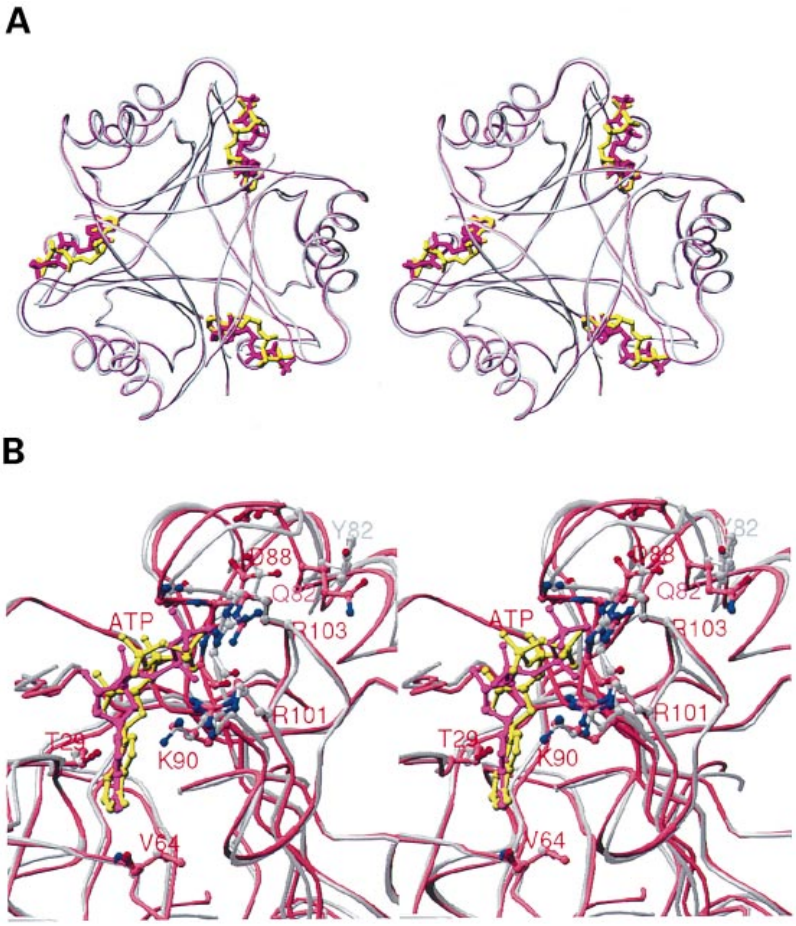

Fig. 3. Structural comparisons of P_{II}/ATP (form II) and GlnK/ATP complex. (A) Stereo diagram of two superimposed trimers with three ATP molecules. (B) Close-up diagram of ATP binding site with side chains of residues that interact with ATP shown (see text). P_{II}/ATP and GlnK/ATP are shown in rose/magenta and silver/yellow, respectively. The figure was produced with RIBBONS [28].

bond with the backbone oxygen of Ile102 from an adjacent monomer (Fig. 4). The conformation of Tyr82 in GlnK is such that it makes nonbonded contact with the backbone atoms of Thr83 and Gly84. The net effect of this amino acid change is a substantial shift in the position of the B-loops so that when GlnK and P_{II} are overlaid, the Cα atoms at position 82 are separated by 1.3 Å. A similar shift is observed in the position of Thr83. In GlnK the side chain of Thr83 forms a hydrogen bond with the side chain of Asp88 [15]. This hydrogen bond is preserved in both proteins and seems to be important in maintaining the structural integrity of the B-loop. In going from P_{II} to GlnK, the B-loop moves as a rigid body; thus the Cα atoms of Asp88 in GlnK and P_{II} are separated by 1.4 Å. On either side of Asp88 there are conserved glycine residues that are important for binding ATP. The effect of the shift in the position of the B-loop is to make the ATP binding cleft much more open in GlnK compared with that found in P_{II}. The shift in position of the B-loop can also explain the differences in the conformation of the α and β phosphates of ATP. By superimposing the two structures, it is evident that an oxygen atom in the α phosphate of the ATP in GlnK is too close (≈ 2.2 Å) to the amide nitrogen atom of Gly89 of P_{II}. As a result the conformation of the triphosphate in GlnK is quite different to those of P_{II} (Fig. 3). The phosphates make different contacts with Arg101 and Arg103 in the two proteins despite the fact that the conformations of these residues are very similar in both P_{II} and GlnK.

In the native GlnK structure it was noted that the ATP binding site was occupied by an anion, probably phosphate or sulfate. The sequence difference at position 82 and the subsequence shift in the B-loop offers an explanation as to why this anion was not observed in the P_{II} structure. This anion would make bad contacts with backbone atoms of both Gly87 and Asp88 in the P_{II} structure.

Table 3. Trimer, monomer and trimerization energies for all PII/GlnK trimers.

Template structure	Type of constituting monomers	Trimer energy [no units]	Sum of monomer energies [no units]	Trimer stabilization energy [no units]
(PII)_3	^{3}PII	−2088	−1738	−349
	$^{(2}$PII)/GlnK	−2212	−1864	−348
	PII/^{2}GlnK	−2304	−1990	−314
	^{3}GlnK	−2420	−2116	−304
(GlnK)_3	^{3}PII	−2027	−1701	−326
	^{2}PII/GlnK	−2094	−1792	−302
	PII/^{2}GlnK	−2168	−1884	−284
	^{3}GlnK	−2234	−1975	−258

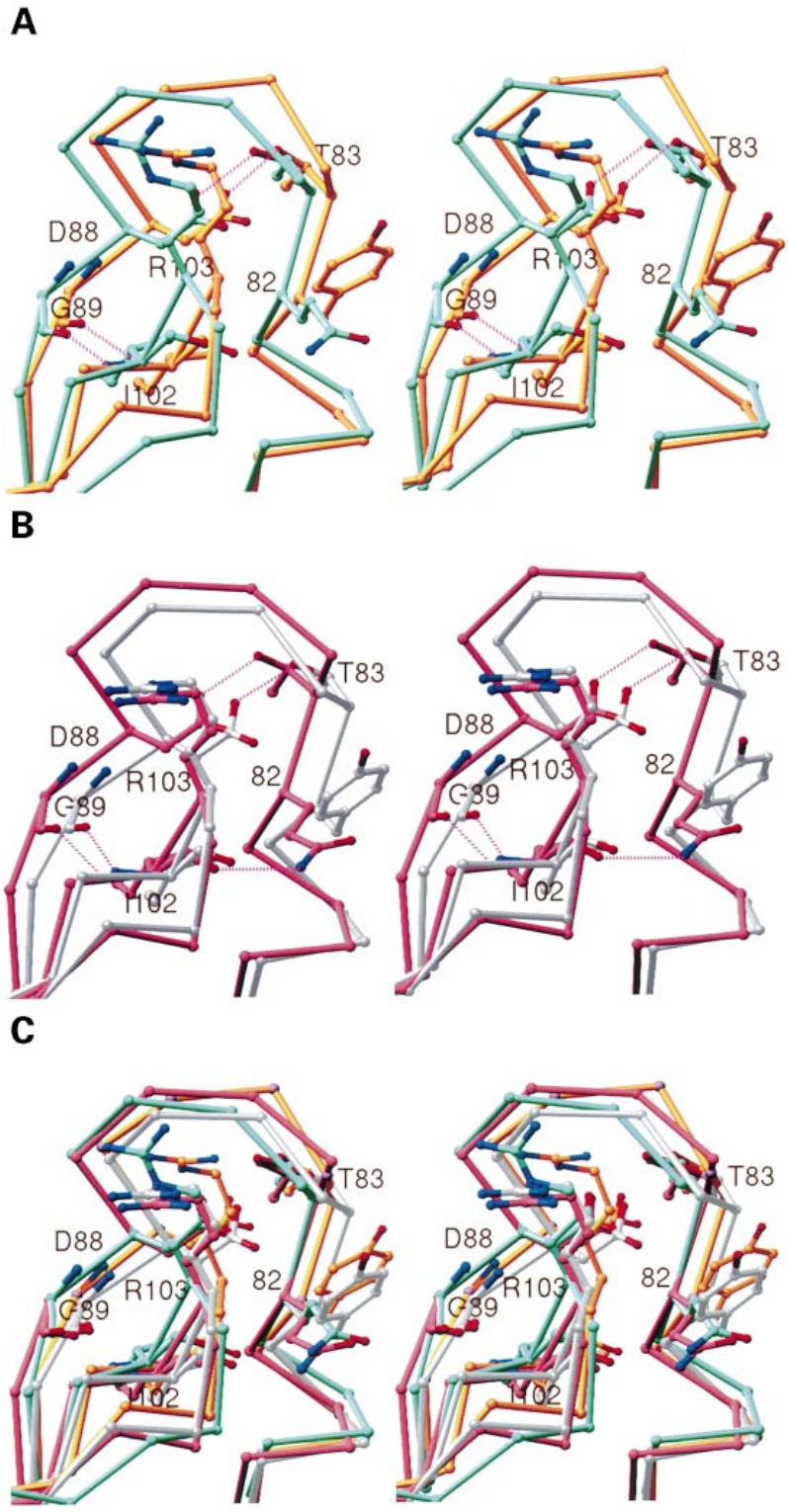

Fig. 4. Comparisons of B- and C-loops. (A) Native (form I) P_{II} (gold) and GlnK molecule (purple). (B) Form II, P_{II}/ATP (rose) and GlnK/ATP complex (silver). (C) The above four structures superimposed together. Residues 82, 83, 88, 89, 102 and 103 are displayed by atom types: nitrogen, blue; oxygen, red; carbon, as the corresponding C^{α} atoms.

CONCLUDING REMARKS

Although the structures and sequences of P_{II} and GlnK are very similar they interact differently with their receptors [8,9]. Both proteins bind and require ATP for activity and the present work shows that the structures of their ATP complexes differ. These differences have been attributed to a sequence change at position 82 that causes a shift in the position of the B-loops of GlnK with respect to that found in P_{II}.

Like other signal transduction proteins, P_{II} and GlnK must interact with a number of receptors, in this case bifunctional enzymes. Specifically, P_{II} and GlnK are the substrate of UT/URase and they modulate the activities of

ATase and NRII. The functionally important areas of P_{II} and GlnK appear to be the T-loop and the ATP binding site. These regions must exhibit a high degree of flexibility so that they can dock with multiple receptors. Although crystallography is not an ideal tool with which to study protein flexibility, comparison of the same protein in multiple crystal forms does give some information about the way protein conformation can vary. Including the present work, there are five crystal structures of P_{II} and GlnK. There are three independent copies of P_{II} in the form II crystals and the native form of GlnK contains two independent copies of that molecule. In all of these structures P_{II} and GlnK appear to have a very rigid core structure with differences restricted to their B and T loops along with their C-termini. Indeed, a comparison of the three copies of the P_{II} molecule in the form II crystals shows that the T-loops are disordered in all subunits and that they differ to the greatest extent in the B-loops and the C termini. These regions also show the largest differences when the different crystal forms of P_{II} are compared. However, the comparisons of the various crystal forms of P_{II} are complicated by crystal packing effects. In most cases, protein crystals are stabilized by numerous weak interactions and the conformation observed in the crystal lattice is one of many that may exist in solution. The conformation of the T-loop observed in the form I crystals is probably one of many possible conformations and we see it because it is compatible with crystal formation. Comparison of the form II structure with that obtained in the absence of ATP (form I) reveals how the T-loop may affect the way the protein interacts with ATP. The form I structure shows an ordered T-loop that interacts with the peptide around position 26. This interaction causes the ATP binding site to partially close and block access to ATP. The comparison of P_{II} in forms I and II show one possible way that the T-loop may influence the structure of the ATP binding site and many others are possible. A comparison of forms I and II also show that the C terminus can take on multiple conformations. Form III shows one conformation of the C terminus that may form at low pH and that is stabilized by the lattice. In summary, the three P_{II} structures show which regions of the molecule are flexible and that are likely to change when binding a receptor. They also provide a structural basis for speculation about other aspects of P_{II} function.

In previous work [15] we have suggested that α-KG binds in the T-loops. Indeed, the structures obtained in the present work are consistent with this suggestion. α-KG was present in the crystallization experiments that produced both the form II and III crystals. Acetate was discovered in the form III crystals indicating that the methodology was capable of detecting a small organic molecule, yet α-KG was not evident in any of the electron density maps examined as part of this work. If α-KG does bind in the T-loops, it would influence the conformation of the loops and hence affect the affinity of the protein for ATP. Form I shows P_{II} in a conformation that is 'closed' to ATP binding whereas forms II and III represent conformations that are 'open' to ATP binding. The structural data support the idea that the equilibrium between the open and closed conformations is influenced by the conformation of the T-loop. This idea provides an explanation for cooperative manner in which P_{II} binds ATP and α-KG. However, a definitive explanation of how P_{II} interacts with both its effectors and receptors awaits the structure determination of a relevant protein–protein complex.

REFERENCES

1. Magasanik, B. (1993) The regulation of nitrogen utilization in enteric bacteria. *J. Cell. Biochem.* **51**, 34–40.
2. Parkinson, J.S. (1993) Signal transduction schemes of bacteria. *Cell* **73**, 857–871.
3. Stock, J.B., Ninfa, A.J. & Stock, A.M. (1989) Protein phosphorylation and regulation adaptive responses in bacteria. *Microbiol. Rev.* **53**, 450–490.
4. Ninfa, A.J. & Atkinson, M.R. (2000) PII signal transduction proteins. *Trends Microbiol.* **8**, 172–179.
5. Son, H.S. & Rhee, S.G. (1987) Cascade control of *Escherichia coli* glutamine synthetase. *J. Biol. Chem.* **262**, 8690–8695.
6. Jaggi, R., Ybarlucea, W., Cheah, E., Carr, P.D., Edwards, K.J., Ollis, D.L. & Vasudevan, S.G. (1996) The role of the T-loop of signal transducing protein P_{II} from *Escherischia coli*. *FEBS Lett.* **391**, 223–228.
7. Jiang, P., Zucker, P., Atkinson, M.R., Kamberov, E.S., Tirasphon, W., Chandran, P., Schefke, B.R. & Ninfa, A.J. (1997) Structure/function analysis of the P_{II} signal transduction protein of *Echerichia coli*: genetic separation of interactions with protein receptors. *J. Bacteriol.* **179**, 4342–4353.
8. van Heeswijk, W.C., Stegeman, B., Hoving, S., Molenaar, D., Kahn, D. & Westerhoff, H.V. (1995) An additional PII in *Escherichia coli:* a new regulatory protein in the glutamine synthetase cascade. *FEMS Microbiol. Lett.* **132**, 153–157.
9. van Heeswijk, W.C., Hoving, S., Molenaar, D., Stegeman, B., Kahn, D. & Westerhoff, H.V. (1996) An alternative PII protein in the regulation of glutamine synthetase in *Escherichia coli*. *Mol. Microbiol.* **21**, 133–146.
10. Forchhammer, K., Hedler, A., Strobel, H. & Weiss, V. (1999) Heterotrimerization of PII-like signalling proteins: implications for PII-mediated signal transduction systems. *Mol. Microbiol.* **33**, 338–349.
11. van Heeswijk, W.C., Wen, D., Clancy, P., Jaggi, R., Ollis, D.L., Westerhoff, H.V. & Vasudevan, S.G. (2000) The *Escherichia coli* signal transducers PII (GlnB) and GlnK form heterotrimers *in vivo*: fine tuning the nitrogen signal cascade. *Proc. Natl Acad. Sci. USA* **97**, 3942–3947.
12. Kamberov, E.S., Atkinson, M.R. & Ninfa, A.J. (1995) The *Escherichia coli* PII signal transduction protein is activated upon binding 2-ketoglutarate and ATP. *J. Biol. Chem.* **270**, 17797–17807.
13. Cheah, E., Carr, P.D., Suffolk, P.M., Vasudevan, S.G., Dixon, N.E. & Ollis, D.L. (1994) Structure of the *Escherichia coli* signal transducing protein PII. *Structure* **2**, 981–990.
14. 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 PII from *Escherichia coli* at 1.9 Å. *Acta Crystallogr.* **D52**, 93–104.
15. Xu, Y., Cheah, E., Carr, P.D., van Heeswijk, W.C., Westerhoff, H.V., Vasudevan, S.G. & Ollis, D.L. (1998) GlnK, a P_I–homologue: structure reveals ATP binding site and indicates how the T-loops may be involve in molecular recognition. *J. Mol. Biol.* **282**, 149–165.
16. Vasudevan, S.G., Gedye, C., Dixon, N.E., Cheah, E., Carr, P.D., Suffolk, P.M., Jeffery, P.D. & Ollis, D.L. (1994) *Escherichia coli* P_{II} protein: purification, crystallization and oligomeric structure. *FEBS Lett.* **337**, 255–258.
17. Rife, J.E. & Cleland, W.W. (1980) Kinetic mechanism of glutamate dehydrogenase. *Biochemistry* **19**, 2321–2328.
18. Edwards, K.J., Suffolk, P.M., Carr, P.D., Wegman, M., Cheah, E. & Ollis, D.L. (1996) Crystallisation and preliminary X-ray

diffraction studies of new crystal forms of *Escherichia coli* PII complexed with various ligands. *Acta Crystallogr.* **D52**, 738–742.
19. Otwinowski, Z. & Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
20. Navaza, J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr.* **A50**, 157–163.
21. Brunger, A.T. (1992) *X-PLOR, Version 3.1 a System for X-Ray Crystallography and NMR.* Yale University Press, New Haven, CT, USA.
22. Brunger, A.T. (1992) The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–474.
23. Jones, T.A. (1985) Interactive computer graphics: FRODO. *Methods Enzymol.* **115**, 157–171.
24. CCP4: Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallog.* **D50**, 760–763.
25. Huber, T., Russell, A.J., Ayers, D. & Torda, A.E. (1999) Sausage: protein threading with flexible force fields. *Bioinformatics* **15**, 1064–1065.
26. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
27. Sawyer, L. & James, M.N.G. (1982) Carboxyl–carboxylate interaction in proteins. *Nature* **295**, 79–80.
28. Carson, M. (1991) Ribbons-2.0. *J. Appl. Crystallog.* **24**, 958–961.