The structure of a PII signaling protein from a halophilic archaea reveals novel traits and high-salt adaptations

Carles Palanca,[1,] Laia Pedro-Roig,[2,*] José Luis Llácer,[1,**] Mónica Camacho,[2] María José Bonete,[2] and Vicente Rubio[1,3]

[1]*Instituto de Biomedicina de Valencia of the CSIC (IBV-CSIC), Valencia, Spain.*
[2]*División de Bioquímica y Biología Molecular, Departamento de Agroquímica y Bioquímica, Facultad de Ciencias, Universidad de Alicante, Alicante, Spain.*
[3]*Group 739 of the Centro de Investigación Biomédica en Red sobre Enfermedades Raras (CIBERER) del Instituto de Salud Carlos III, Spain*

Correspondence to:
Carles Palanca (cpalanca@ibv.csic.es) and Vicente Rubio (rubio@ibv.csic.es)
Instituto de Biomedicina de Valencia, C/ Jaime Roig 11, Valencia-46010, Spain.
Phone: +34 963 391 772   Fax: +34 963690800

*Present address:
MRC Protein Phosphorylation and Ubiquitylation Unit, College of Life Sciences, University of Dundee, Dundee, DD1 5EH, U.K.

**Present address:
MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK.

Article type   : Original Article

ABSTRACT

To get insight on archaeal nitrogen signaling and on haloadaptation of the nitrogen/carbon/energy-signaling protein PII, we determined crystal structures of recombinantly produced GlnK2 from the extreme halophilic archaea *Haloferax mediterranei,* complexed with AMP or with the PII effectors ADP or ATP, at respective resolutions of 1.49, 1.45 and 2.60 Å. A unique

trait of these structures was a three-tonged crown protruding from the trimer body convex side, formed by an 11-residue, N-terminal, highly acidic extension that is absent from structurally studied PII proteins. This extension substantially contributed to the very low pI value which is an haloadaptive trait of *H.mediterranei* GlnK2, and participated in hexamer-forming contacts in one crystal. Similar acidic N-extensions are shown here to be common among PII proteins from halophilic organisms. Additional haloadaptive traits prominently represented in *H. mediterranei* GlnK2 are a very high ratio of small residues versus large hydrophobic aliphatic residues; and the largest ratio of polar to non-polar exposed surface for any structurally characterized PII protein. The presence of a dense hydration layer in the region between the three T-loops might also be a haloadaptation. Other unique findings revealed by the GlnK2 structure that might have functional relevance are the adoption by its T-loop of a 3-turn alpha helical conformation, perhaps in relation with the GlnK2 ability to directly interact with glutamine synthetase; and the firm binding of AMP, confirmed by biochemical binding studies with ATP, ADP and AMP, raising the possibility that AMP could be an important PII effector at least in Archaea.

INTRODUCTION

One of the strategies of halophilic archaea for coping with the high osmolarity/salinity of their environment is to internalize and accumulate salts (salt-in strategy) [1]. This strategy requires that the intracellular machinery, including cellular proteins, adapt to a highly saline internal medium [2]. Although our understanding of this adaptation process has advanced considerably thanks in part to structural studies [2-8], more structural data on halophilic proteins is necessary particularly for proteins involved in signaling processes mediated by transient interactions with other proteins. In these cases adaptation might pose particular challenges, since salinity decreases the strength of ionic bonds and increases that of hydrophobic interactions [9,10]. Furthermore, an increased number of fixed water molecules was found at the surface of some halophilic proteins [3-5], and thus, if this layer were found generally in halophilic proteins, it might result in the shielding by water of protein groups that could otherwise be involved in direct protein-protein interaction.

In this context, an interesting signaling protein to study is PII, a very ancient, highly conserved (Fig. 1) and widely distributed homotrimer which plays pivotal roles in nitrogen/carbon/energy regulation [11,12]. PII senses the energy status of the cell and also the nitrogen/carbon abundance balance by binding the allosteric effectors ATP/ADP and 2-oxoglutarate (2OG) [11-15]. This last compound is a carbon and nitrogen status reporter, since it is abundant if there is carbon fullness and the ammonia load is low [16, 17]. These PII effectors modulate the binding ability of PII to its protein targets, which include key nitrogen-assimilating enzymes, ammonia channels, and adaptor proteins involved in gene expression regulation [11-14]. The PII trimer, composed of a hemispheric body formed by an inner closed β-sheet layer covered by α helices and nine prominent loops, three long-flexible ones (the T-loops) and six shorter ones (three B-loops and three C-terminal loops) [18], uses the T-loops in many of its interactions with target proteins [14, 19-23].

We have chosen for structural study a PII protein from the extreme halophilic and thermophilic archaea *Haloferax mediterranei*, an organism first isolated from hot salterns in the Spanish Mediterranean coast [24]. This organism grows optimally at 2.5 M NaCl and 51ºC [25], and is believed to accumulate, as other haloarchaea [1], very high salt concentrations, making it an excellent model to study the impact of saline adaptation on the PII protein. The study of PII in *H. mediterranei* is endowed with additional interest since the two PII proteins of this organism, GlnK1 and GlnK2 (called here hmGlnK1 and hmGlnK2) were reported to interact directly with the classical

PII target glutamine synthetase (GS) [26,27] instead of indirectly controlling this enzyme via an intermediate GS-modifying enzyme, as in enterobacteria [28]. Indeed, the direct interaction of PII with GS might be a characteristic of archaeal PII signaling, since it was reported also in *Methanosarcina mazei* [29]. In fact, although much is known about PII from many bacteria, our knowledge of PII and of its signaling network is more limited for archaea, with structures having been determined only for four different archaeal PII proteins, one from *Methanococcus jannaschii* [30] and three from *Archaeoglobus fulgidus* [14,31,32]. Nevertheless, analyses were made with archaeal PII proteins for structure, ligand binding, cooperativity, T-loop structure and protein-protein interactions [14,29-32], which provides already a reference within the same life domain for comparison of the present data with our halophilic archaeal protein.

We determine here crystal structures of recombinantly produced hmGlnK2 complexed with AMP or with the PII effectors ADP or ATP, at respective resolutions of 1.49, 1.45 and 2.60 Å. The studies reported here reveal as unique structural features the presence of a crown of acidic N-terminal extensions (one per subunit in the hmGlnK2 trimer), the presence of a localized hydration layer, the fold of the T-loop as a 3-turn α-helix in the ATP crystal, and the firm binding of AMP to the nucleotide site of PII, a binding that is confirmed also here by biochemical binding studies. These features, some of which appear related to mechanisms of haloadaptation exemplified by PII as a paradigm for a signaling protein, may represent structural correlates of functional traits of PII regulation that could be exclusive to Archaea.

RESULTS

Crystallization and overall structure of *H. mediterranei* GlnK2.

Although purified recombinant hmGlnK1 and hmGlnK2 were used in extensive crystallization trials, only hmGlnK2 yielded crystals. The crystals prepared without added ligands or with AMP or ADP added, were of cubic habit and 0.3-0.5 mm dimension and they diffracted synchrotron X-rays at 1.81, 1.49 and 1.45 Å resolution, respectively (Table 1). The crystals without added ligands nevertheless contained bound ADP and, being of lower resolution and not having any differential traits with respect of those with ADP added, will not be discussed further. After phasing by molecular replacement (see Materials and Methods), the crystals were found to host one typical PII subunit in the asymmetric unit, with the entire polypeptide visible, excepting most of the T-loop (residues 49/50-63/64 not visible) and residues 1-4/5. Application of the crystal symmetry yielded typical PII trimers having the characteristic hemispheric shape and fold for the trimer body, nucleated by three 6-stranded β sheets formed by 4 strands of one subunit and a hairpin from the C-loop of the adjacent subunit, and covered outwards, in each subunit, by two α helices running parallel to the longest β-strands (Fig 2A,B). Nevertheless, these structures presented, as a novel unique feature (Figs. 2B,C), a three-tongued protruding crown sitting over the convex side of the hemispheric PII body, formed by the N-terminal extensions of the three subunits, which are not found in classical PII proteins (Fig. 1 and see below).

We also grew crystals in the presence of ATP. These crystals, which diffracted to 2.6 Å-resolution (Table 1), differed importantly from the other three crystals, being tetragonal bipyramids of 0.2 mm maximum dimension, belonging to a different space group and having another unit cell, which contained in the asymmetric unit a complete PII trimer. Interestingly, the majority or a large part of the T-loop was patent in two subunits (missing residues, 51-53 in one subunit and 51-57 in another subunit). Otherwise, the overall subunit fold and the trimeric architecture were virtually identical in the four crystals, as reflected in good root mean square deviation (rmsd) values for superimposition of C^α atoms of individual subunits (rmsd ≤ 0.24 Å) and of entire trimers (rmsd ≤ 0.30

Å). These trimers, devoid of the T-loops and of the N-terminal extensions, were also highly similar to the PII trimers from all archaeal PII proteins for which structures have been determined (rmsd values ≤ 1.1 Å for 276 C^α atoms, data not shown), as illustrated by the superimposition with the structure of GlnK2 from *A. fulgidus* (Fig. 2D). These PII trimers were also highly similar to those of *Escherichia coli* GlnK (rmsd value, 1.38 Å for 282 C^α atoms), considered the paradigm for eubacterial PII proteins. These results indicate that the adaptation to high salinity does not involve distorting the characteristic structure of the PII body, although, as discussed below, it was associated with the addition of the N-extension.

The N-terminal extension

The observation of clear-cut putative Shine-Dalgarno sequences preceding the initial ATG codons of the *hmglnK1* and *hmglnK2* genes (respectively, GGAGG at bases -10 to -14 and AGGAGG at bases -12 to -17 [33]), and the results of fingerprinting proteomics studies of hmGlnK1 and hmGlnK2 isolated from *H. mediterranei* [34] indicate that the N-terminal extension is indeed present in the naturally produced hmGlnK proteins. In fact, we found more PII proteins with putative N-terminal extensions in sequence databases (Table S1), including PII proteins from halophilic organisms. In these organisms the extension tends to be richer in acidic residues that in non-halophiles (Fig. 3A, Table 1).

In the structures of the hmGlnK2 protein, residues 5/6-11 of the N-terminal extension are visible (Fig. 2E), revealing a 3_{10} helix turn (residues 8-10) at the end of the extension, which is connected to the protein body by Gly10 and Gly11. Despite the presence of these glycines, the visible part of the extension is essentially fixed in the same position and conformation, as shown by superimposition of the C^α atoms in our hmGlnK2 structures (rmsd ≤0.77 Å, for C^α of residues 6-11). This is largely due to the firm anchoring on the protein body of the 3_{10} helix turn, via the side chains of Asn8 and of the two residues preceding it (Leu6 and Pro7) (Fig. 2E). The firmly anchored and rigid part of the N-extension is the stand from where the first 5 residues of the extension must be projected outwards as a highly mobile or flexible element that is not visible in the crystals. This mobile part is polar and negatively charged, since it hosts Ser2, Asp3 and Asp5 (Fig 1), which, together with Asp9, belonging to the fixed part of the extension, decreases the pI value of the protein by one pH unit, favoring protein solubility in high salt medium, since a low pI value is a well-known trait of haloadaptation [35]. In fact, hmGlnK2 and GlnK1 appear to be among the GlnK proteins with the lowest sequence-estimated pI values (Fig. 3B), judged by the sampling represented by the proteins annotated as GlnK in the Comprehensive Microbial Resource [closed after this manuscript submission; GlnK proteins in this database are listed, together with their pI values and (Ala+Gly+Ser+Thr)/(Ile+Leu+Met+Val) ratios, in Supplementary Table S2].

In the crystal with ATP, the N-terminal extensions of two trimers mediate hexamer formation by cogwheel-type trimer-trimer contacts (mainly with exposed residues from α2 and β5 from the other trimer; not shown) across the twofold crystallographic axis that crosses perpendicularly the aligned threefold axes of both trimers (Fig. 2F). Despite the small extension of the contact surface (560 $Å^2$ per trimer) and a poor predicted stability for this hexamer in solution (estimated with the PISA server http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html), a role of the N-terminal extension in making interactions with another PII molecule or with other proteins cannot be excluded in the water-restricted, highly saline environment of *H. mediterranei*.

Other haloadaptive traits

An haloadaptive trait which appears to be particularly related to the strengthening of hydrophobic interactions in high salt medium [36] is the high propensity for small residues in halophilic proteins [9,10,35]. This trait is presented by the hmGlnK proteins, as reflected in their high (Ala + Gly + Ser + Thr)/(Ile + Leu + Val + Met) ratio (Fig. 3C), among the three largest ones for the proteins that were annotated as GlnK in the Comprehensive Microbial Resource (closed after manuscript submission and listed in Supplementary Table S2). In addition, the fraction of the solvent-accessible surface of hmGlnK2 that is polar, estimated from our structures, is the highest for any known PII protein structure (Figs. 3D), as graphically exemplified by the higher surface potential of hmGlnK2 than of *A. fulgidus* GlnK2 protein (Fig. 4A). A large polar exposed surface, combined with a low pI, should favor protein solubility [37], preventing hmGlnK2 from being salted out in the highly saline cytoplasm of *H. mediterranei*.

A feature that was proposed to be a potentially important element of halophilic protein adaptation, as highlighted in studies with glucose dehydrogenase from *H. mediterranei* [5], was the existence of a dense hydration layer around the protein (Fig. 4B). However, only 72 fixed water molecules were observed to be attached per PII subunit in the structure of hmGlnK2 exhibiting the highest resolution (1.45 Å) (Fig. 4C), corresponding to ~0.6 water molecules per protein residue, much less than the 1.9 fixed water molecules per protein residue reported in the structure at 1.6 Å resolution of glucose dehydrogenase from the same organism [5]. Furthermore, we could not find any of the pentagonal rings of surface water molecules observed in abundance in the latter structure [5]. Therefore, our findings suggest that these hydration shell traits cannot be considered constant adaptive features of soluble halophilic proteins even when these belong to the same organism. Nevertheless, 81 water molecules formed a cushion between each two PII trimers generated by crystallographic symmetry in the crystals of the AMP and ADP complexes (Fig. 4D). The two trimers interact mutually via crystal contacts mediated by the bases of the T-loops, in fact forming relatively stable hexamers (ΔG^{diss} -1,2 Kcal/mol, estimated with the PISA server), and the water molecules are sandwiched between the two trimers, suggesting the existence of strong solvation forces that might be related to the halophilic character of the protein.

The allosteric site and the complex with AMP

Although no nucleotides were added in our initial crystallization trials of hmGlnK2, we found a mass of electron density fitting an ADP molecule binding at the expected nucleotide site in the structure obtained for this crystal at 1.81 Å resolution (not shown). Since the electron density was better for the AMP moiety than for the β-phosphate, we did set up crystallization trials adding either ADP or AMP to the crystallization droplets. The crystals obtained, which diffracted X-rays at respective resolutions of 1.45 and 1.49 Å, respectively, contained ADP and AMP, respectively, in the nucleotide site. As already indicated, we also obtained ATP-containing crystals, although the resolution for these crystals was less good (2.6 Å).

Our success in obtaining AMP-containing crystals was unexpected, since ADP and ATP are the well established nucleotidic allosteric effectors of PII [13,15], and also because, among the more than fifty structures of PII proteins deposited in the Protein Data Bank (PDB), there is AMP in only one case, in one single subunit among 12 subunits in the asymmetric unit of the crystal of PII from the archaea *M. jannaschii* (PDB file 2J9D [30]). Since both *M. jannaschii* and *H. mediterranei* are Archaea, it is conceivable that AMP might be a genuine allosteric PII ligand in the archaeal domain of life. Nevertheless, AMP has been reported with the PII of the nitrogen-fixing bacterium *Azospirillum brasilense* to prevent and to stimulate, respectively, complex formation with the PII target DraG and

with the targets Amt and DraT [38,39], raising the possibility that AMP could be a poorly studied allosteric effector of any PII protein.

We corroborated biochemically that AMP was indeed bound by hmGlnK2 with high affinity, by centrifugal ultrafiltration binding studies. We used as a control *Synechococcus elongatus* PII because the binding of ATP and ADP to this highly studied PII protein has been well characterized (see for example [15]). Our studies with this PII protein revealed (Fig. 5A) the binding of one molecule of ATP or of ADP per PII subunit agreeing with the three-site model and approximating the affinities recently reported for this protein [15]. *S.elongatus* PII was also able to bind AMP, but its affinity for this nucleotide was much lower than for ADP or ATP (Fig. 5A). When we assayed the binding of these nucleotides to hmGlnK2 (Fig. 5B), ATP and ADP again were bound with high affinities resembling those for ATP of *S. elongauts* PII, but, in addition, AMP was bound with comparable affinity, an affinity similar to that for ADP of *S. elongatus* PII, confirming biochemically the strong tendency of hmGlnK2 to form a complex with AMP.

In the crystal structure of the hmGlnK2-AMP complex, the nucleotide was bound to its site (Fig. 6A) essentially as the AMP moiety of ADP (Fig. 6C) or ATP (Fig. 7), at the boundary between two subunits (Fig. 6B,D,E; residues of one subunit are distinguished from those of the other subunit by an asterisk). The adenine sits in a cavity, sandwiched between the edge of the invariant Phe103, on one side, and the side-chain of Ser38* and the main chain of residues 38*-40* on the other side (Fig. 6A). The phosphate of the AMP (as well as the di- and tri-phosphate chains of ADP and ATP in the corresponding crystals) is sandwiched between the beginning of the T-loop and the P-loop (Figs. 6 and 7). The phosphate of the AMP is firmly anchored on the protein structure, with one of its non-bridging O atoms making two hydrogen bonds with main-chain N atoms of the contouring P-loop (Fig. 6A,B) and with another non-bridging O atom hydrogen-bonded to the Lys101 side-chain, whereas the third O atom is connected through two fixed water molecules with His67 and Lys69 and with the ribose moiety of the nucleotide (Fig. 6B). The main differences with the ADP complex concern hydrogen bonding of the phosphates (Fig. 6C,D). In the ADP complex one of the P-loop hydrogen bonds has migrated to the β phosphate. This phosphate makes another bond to the N atom of Ser49, at the root of the T-loop, and a water molecule interconnects both phosphates, anchoring them on the protein via another fixed water molecule that makes bonds with main chain O and N atoms of Gly100 and Arg114*, respectively (Fig. 6D, and inset therein). Despite these differences, the similar number and distances of the bonds made by AMP and ADP with PII supports the binding of both nucleotides with analogous strengths, as observed, and, even more important, that they could favor the same PII conformation. The main difference with our complex with ATP (Figs. 6D and 7) is that the β phosphate is not hydrogen-bonded with the P-loop, which, as in the complex with AMP, makes two bonds with the α phosphate. The γ phosphate makes novel bonds with the side chain of Arg114* and, in two subunits, with the side-chain of Gln112* (Fig. 6E).

In any case, despite the changes in protein sequence associated with haloadaptation, all the residues that are centrally involved in nucleotide binding in other PII proteins (highlighted in Fig. 1; see for example [14,31]) are conserved, including the two residues shown with *E. coli* GlnK (Gln39 and Lys58, corresponding to Gln50 and Lys69 in hmGlnK2) to be crucial for the recently reported PII ability to hydrolyze ATP [40] that may be a crucial signaling feature of PII proteins. The conservation of the residues that are hydrogen-bonded to the 2OG carboxylate groups in the three PII structures that showed 2OG bound [14,41,42], Lys69, Gly48 and Gly52 (hmGlnK2 residue numbering, Fig. 1), the preservation of the 2OG-contacting residue Gly98 (Fig. 1), and the lack of apparent steric clash (determined by in silico mutation in the *A. brasilense* PII-MgATP-2OG structure, PDB file 3MHY;

not shown) with the side-chains of non-conserved residues within the 2OG site suggest that 2OG can bind to hmGlnK2.

An alpha helical T-loop

Although in the structures with AMP or ADP the T-loop (residues 48-65) could not be traced, as it is frequent for PII when not complexed to a target, in the trimer found in the asymmetric unit of the crystal of hmGlnK2-ATP a nearly complete T-loop (only 3 residues missing) was observed in one subunit (Fig. 7) and a half-complete T-loop (missing residues, 51-57) was observed in another subunit (not shown). Both visible T-loops were involved in crystal contacts with other PII trimers that could help fix them in the crystal. Interestingly, the loop that was more complete presented an imperfect α-helical conformation spanning nearly three helix turns (residues 56-64) (Fig. 7). The other partially visible T-loop also exhibited a shorter near-helical structure (residues 60-65, not shown). T-loops rarely present canonical secondary structure elements when not in complexes with target proteins, although in some of these complexes they have been found to form a β-hairpin [19-22], which, in the case of the complex with acetylglutamate kinase, made a hybrid β-sheet with β-strands of the target [21]. Only three PII structures in the whole protein databank (GlnK from *E. coli*, 1GNK [18], GlnZ from *A. brasilense,* 3MHY [41] and PII from *Thermus thermophilus*, 1VFJ [43]) showed three to five T-loop residues in helical conformation, but in no case a helix of the size of the present structure was reported for a T-loop. In the T-loop helix of hmGlnK2 (Fig. 7), Tyr62 has its phenolic group buried, clamped between two residues at the beginning of the loop (Arg47 and Gln50) and making a hydrogen bond via its phenolic OH with the main-chain O atom of Gly48. If this tyrosine is uridylylated in hmPII, as concluded from proteomics results [34], the position of the helical T-loop observed in our structure, and even the helical folding, might not be possible for the covalently modified T-loop.

DISCUSSION

This work represents the first structural study of a PII protein from a halophilic archaea, and one of the few carried out with archaea [14,30-32]. It confirms the high structural conservation of PII proteins in all domains of life in which they are present (only absent from the animal kingdom), including extremophilic archaea, a conservation that encompasses important functional elements such as the high flexibility and the polysterism of the T-loops [43], the constancy of the allosteric sites, particularly for the nucleotide [15], and the possibility that the recently discovered trait of a slow regulatory ATPase [40] could also occur with the PII proteins of extremophilic archaea.

However, our work reveals peculiarities that help understand how this widely conserved regulatory protein adapts to high salinity while still keeping the main characteristics of the PII family. Both hmGlnK1 and hmGlnK2 exhibit a highly acidic pI and a strong predominance of short residues, in line with current knowledge about these traits as key elements for haloadaptation [2,7,8,35]. The high fraction of exposed surface that is polar, which also favors haloadaptation, is largely due to surface-exposed O atoms (not shown), many of them from main-chain peptide bonds that are unshielded because of the predominance of small residues. Small residues also predominate over large hydrophobic residues in the protein core, which may prevent the protein from becoming too rigid in its normal high salt medium, given the increase in the strength of hydrophobic interactions that occurs in high salinity media [9,10,37]. Another purported haloadaptation [3-5], strong surface hydration, is also identified here, although restricted to a given region of the protein, instead of being general for the entire exposed surface as reported for glucose dehydrogenase of the same organism [5]. An extensive fixed water layer shielding the regions of interaction of PII with its signaling partners or its

allosteric binding sites would not favor PII signaling and its regulatory roles. This certainly differs from the less strict surface requirements for a soluble enzyme such as glucose dehydrogenase, in which only the active site must be available for specific chemical interactions with the ligand, and thus other parts of the protein can be heavily hydrated. In any case, despite the doubts raised recently about the importance of the hydration layer for haloadaptation, mainly because of the similar water dynamics of mesophilic and halophic-like versions of the same protein [44], the presence of the dense hydration layer at the surface between the bases of the three T-loops is suggestive of some role of surface water in haloadaptation.

An interesting peculiarity revealed by the structure of hmGlnK2 that characterizes halophilic PII proteins and which has not been reported in any PII protein of known structure, is the presence of an 11-residue N-terminal extension. PII proteins exhibit a remarkable degree of sequence conservation [45,46] (Fig. 1), which may represent a challenge for adapting to their widely different environments in diverse organisms, particularly for extremophiles. The addition of the short N-terminal extension of hmGlnK2 and of other halophilic PII proteins may be a way to reconcile the high conservation with the requirement for lowering the pI value for adapting to the high salt world. Indeed, other types of acidifying insertions were already reported for *Halobacterium salinarum* ferredoxin [47] and *Halobacterium species* NRC-1 cysteinyl-tRNA synthetases [48]. In addition to help the protein adapt to a high salt concentration, these extensions might be involved in protein-protein interactions similar to those observed here between PII trimers, since they are exposed, they are charged, and they represent a substantial fraction of the total surface of the protein.

Besides its halophilic character, *H. mediterranei* is also a moderate thermophile and its proteins should stand the relatively high living temperature (optimal, 51 ºC) of this organism. However, this should not have been a problem even for the PII proteins of mesophilic organisms, since these proteins are highly stable and have been purified after a 70 ºC heating step [49]. Correspondingly, no special characteristics associated with the thermophilic character are observed in hmGlnK2 other than the more charged protein surface [50] than in other PII proteins (see comparison of surface of hmGlnK2 and *A. fulgidus* GlnK2, Fig. 4A). Even for this trait, it is remarkable that only one ion pair network of a mere three elements (Asp78-Arg120-Asp121; not shown) is observed in each subunit, not extending across intersubunit boundaries. It is uncertain whether the crystal packing of our PII protein into hexameric dimers of trimers (Figs. 2F and 4D) may reflect an aggregation trend that has been associated in some proteins with increased thermal stability [50,51].

Another peculiarity revealed by the present hmGlnK2 structure is the tendency of hmGlnK2 to bind AMP, as reflected in the crystal structure with this nucleotide and in the affinity observed for it in the binding studies, much higher than the affinity for AMP of a paradigmatic bacterial PII protein such as *S.elongatus* PII. As already indicated, the binding of AMP is exceptional among the reported structures of PII proteins. Comparison of the nucleotide sites of *S.elongatus* PII and of hmGlnK2 (Fig. 8) fails to reveal any other gross difference between them than the replacement in hmGlnK2 of a glycine residue that is present in the vast majority of PII sequences (Gly27 of *S. elongatus* PII, Fig. 1) by a serine residue (Ser38 of hmGlnK; marked with a vertical arrow in Fig. 1). As already indicated, the adenine ring is sandwiched between a constant phenylalanine (Phe103) and either the glycine (*S. elongatus* PII) or the serine (hmGlnK2), which belong to the adjacent subunit. Therefore, the side-chain of the serine may provide a better sandwiching of the adenine ring, resulting in better hold of the adenosine moiety and thus in better affinity for AMP. In any case, the similar number and strength of the bonds with AMP and with ADP (Figs. 6B and 6D) fits quite nicely the closeness of the binding affinity for AMP and for ADP and ATP revealed here. Therefore, with hmGlnK2 the choice of PII

between ADP and AMP may depend much more on the relative concentrations of these nucleotides in the solution than with other organisms such as *S. elongatus*, in which the affinity for AMP is much lower. The change in the ratio in ATP/AMP may be more relevant to describe the energy status in the cells than ATP/ADP alone, and therefore hmGlnK2 could be a better sensor of the energy charge than other PII proteins having poorer affinity for AMP.

Given the high affinity of hmGlnK2 for ATP, ADP and AMP, it would have been expected that the nucleotide bound in the *E. coli* cells in which PII was expressed were the most abundant one in the cellular cytoplasm, ATP. However, the final model of the hmGlnK2 crystal without added nucleotides clearly had ADP bound (data not shown), as expected for ATP cleavage by the slow ATPase activity recently reported for PII proteins [40]. Our bound ADP molecule shares with a recently reported high-affinity-bound (to PII) form of ADP [15] the formation of a charged hydrogen bond between the α-phosphate and a lysine side chain (Lys101 of hmGlnK2, Fig. 6C,D). However, the reported bonds between the β-phosphate and the side chains of the C-loop residues Gln112* and Arg114* [15] are not preserved in our crystal structure with ADP (Fig. 6C,D), suggesting that the interaction with the lysine is enough to ensure the high-affinity binding of ADP observed in our present binding studies. In any case, the ability of hmGlnK to bind AMP raises the question of whether AMP and ADP promote the formation of the same complexes with its partners, as for example Amt. In principle, the interactions made by both nucleotides with the PII protein appear similar enough to suggest that this is indeed the case.

The observation made here in the hmGlnK2-ATP complex of a T-loop folding as a relatively large helix (Fig. 7), extends the repertoire of known T-loop conformations, in line with the concept of polysterism recently coined for this loop [43]. At present, in the absence of structures in which hmGlnK2 is complexed with 2OG or with its protein partners, it would be speculative to attribute specific T-loop conformations to given functional states of this halophilic protein. Nevertheless, by analogy with the *E. coli* PII-AmtB complex [19,20], an extended β-hairpin conformation appears likely for the complex demonstrated in *H. mediterranei* between GlnK and Amt [52]. Given the observations with *E. coli, S. elongatus* and *A. fulgidus* [14,19,20,22], ADP may promote this extended conformation. Since the interactions of AMP and ADP with hmGlnK2 are quite similar, this extended conformation could also be promoted by AMP. Structural studies on the hmGlnK-glutamine synthetase complex will be required to clarify whether the helical T-loop conformation observed here is involved in the direct interaction with glutamine synthetase reported for hmGlnK2 [26,27].

MATERIALS AND METHODS
HmGlnK2 preparation, crystallization and X-ray diffraction
Recombinant hmGlnK2 was produced in *E. coli* from a pET-14b vector containing the cloned *H. mediterranei glnK2* gene and was purified as previously reported [27]. For placing the protein in the final conservation solution (20 mM Tris-HCl pH 8.0, NaCl 0.5 M and 10 mM $MgCl_2$) and for attaining the high concentrations of protein required for crystallization (~5 mg/ml; Bradford assay [53] using bovine serum albumin as standard) the protein was subjected to several cycles of ultrafiltration at 4ºC using a centrifugal device (Amicon Ultra, 10K, from Millipore).

Crystals were obtained by the vapor diffusion technique at 21ºC in hanging drops prepared manually by mixing 1 μl of protein solution with the indicated nucleotide and 1 μl of crystallization solution, using 24-well plates and commercial crystal screen solutions (plates and screen from Hampton Research). Initial trials in the absence of nucleotides yielded the best crystals, having cubic

habit, with solution 14 of Crystal Screen 2, consisting of 2 M $(NH_4)_2SO_4$, 0.1 M Na citrate pH 5.6 and 0.2 M Na/K tartrate. These same conditions were used for growing crystals with 10 mM AMP or 2 mM ADP. However, in the presence of ATP (20 mM) good-diffracting crystals having a tetragonal bipyramidal habit only grew with solution 50 of the classical Crystal Screen (Hampton Research), consisting of 0.5 M Li_2SO_4 and 15% PEG 8000. Crystals were flash frozen in liquid nitrogen after passage through crystallization solution supplemented with 28% PEG 400 (crystals without ligands or with AMP), or with 25% glycerol (ATP complex), or, in the case of the ADP-containing crystal, without using a cryoprotectant.

The crystals were diffracted at 100K using synchrotron radiation from the indicated beamlines from either the European Radiation Synchrotron Facility (ESRF), Grenoble, or the Diamond Light Sources (Didcot, Oxfordshire) (Table 1). Data were processed with XDS [54] or iMosflm [55] and scaled with Aimless [56].

Phasing, model building and refinement

Crystals grown either without added ligands or in the presence of AMP or ADP belonged to the H32 space group and had identical cells dimensions. In contrast, the crystal grown with ATP belonged to space group $P3_121$ having a larger unit cell. Phases for the form without ligands were determined by molecular replacement with Phaser [57], using a polyalanine model of subunit A of PII from *S. elongatus* PCC 7942 devoid of its T-loop (PDB code, 2XUL [42]). The solution, consisting of one subunit in the asymmetric unit, was refined by iterative cycles of rigid body and restrained refinement using REFMAC [58], followed by manual model building using COOT [59]. We used PHENIX Refine [60] to perform a final refinement cycle with automatic TLS group finding and refinement. The model of the subunit produced in this way at 1.8 Å resolution, without the bound nucleotide, was used for phasing with the diffraction data from the AMP, ADP and ATP crystals. In the case of the ADP and AMP-containing crystals the asymmetric unit contained a single PII subunit. With these crystals we performed in the final stages anisotropic refinement (with REFMAC) instead of using TLS, obtaining models at 1.45 and 1.5 Å resolution, respectively. In the case of the crystal grown in the presence of ATP, the solution obtained upon phasing by molecular modeling consisted in three PII subunits organized as a classical PII trimer. Refinement and model building were carried out as for the ligand-free form, yielding a final model at 2.6 Å resolution.

Geometry analysis of the protein main chain torsion angles with RAMPAGE [61] revealed good values for all the models.

Surface analysis was carried out with GETAREA [62], using a probe of 1.4 Å-radius. All of the structural figures were prepared with Pymol (http://www.pymol.org).

Centrifugal ultrafiltration assays of ATP, ADP and AMP binding to PII proteins

Eppendorf tube-size centrifugal ultrafiltration devices (Nanosep 3K Omega, Pall Life Sciences, Ann Arbor, MI, USA) were prewashed by flushing through them (10-min centrifugation, $14,000 \times g$, Eppendorf microcentrifuge) 0.2 mL of the appropriate incubation mixtures (see below) lacking protein and nucleotides, followed by 2-min centrifugation of the empty ultrafiltration devices. Mixtures of 0.1 mM (subunit concentration) of either hmGlnK2 (protein estimated from sequence-deduced $A_{280}^{1\%}$ of 4.63 [63]) or *S. elongatus* PII (prepared and assayed for protein concentration as reported [21]) and variable concentrations of the appropriate nucleotide (range up to 0.3 mM) were incubated 10 min at 24 ºC in 50 ml Hepes 50 mM pH 7.5 and NaCl 0.5 M (hmGlnK2) or 0.1 M (*S. elongatus* PII), and were placed in the washed ultrafiltration devices and immediately centrifuged 15

secs*, discarding the ~2 μl ultrafiltrate. After 30-seconds re-centrifugation, the effluent (~5 μL) was collected and the concentration of the nucleotide was determined by UV absorption (ε_{259} of 15.4×10^3 M^{-1} cm^{-1} [64]) using a microspectrophotometer (Nanodrop ND-1000; from Thermo Scientific-NanoDrop, Wilmington, Delaware, USA). This concentration was compared with that in an ultrafiltrate from an identical incubation lacking PII protein. In this way, the fractions corresponding to free and PII-bound nucleotide were estimated. Data were adjusted to three-site binding using GraphPad Prism (GraphPad Software, San Diego, California).

Author contributions. C.P., L.P.R., M.C., M.J.B., and V.R. planned the experiments, C.P., L.P.R., J.L.L. and M.C., performed the experiments and C.P. and V.R. analyzed the data and wrote the paper.

REFERENCES

1. Roberts MF (2005) Organic compatible solutes of halotolerant and halophilic microorganisms. *Saline Systems* **1**, 5.
2. Eisenberg H, Mevarech M & Zaccai G (1992) Biochemical, structural, and molecular genetic aspects of halophilism. *Adv Protein Chem* **43**,1–62.
3. Dym O, Mevarech M & Sussman JL. (1995) Structural features that stabilize halophilic malate dehydrogenase from an archaebacterium. *Science* **267**, 1344-1346.
4. Frolow F, Harel M, Sussman JL, Mevarech M & Shoham M. (1996) Insights into protein adaptation to a saturated salt environment from the crystal structure of a halophilic 2Fe-2S ferredoxin. *Nat Struct Biol* **3**, 452-458.
5. Britton KL, Baker PJ, Fisher M, Ruzheinikov S, Gilmour DJ, Bonete MJ, Ferrer J, Pire C, Esclapez J & Rice DW (2006) Analysis of protein solvent interactions in glucose dehydrogenase from the extreme halophile *Haloferax mediterranei. Proc Natl. Acad Sci USA* **103**, 4846–4851.
6. Coquelle N, Talon R, Juers DH, Girard É, Kahn R & Madern D (2010) Gradual adaptive changes of a protein facing high salt concentrations. *J Mol Biol* **404**, 493–505.
7. Siglioccolo A, Paiardini A, Piscitelli M & Pascarella S (2011) Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface. *BMC Struct Biol* **11**, 50.
8. Tadeo X, López-Méndez B, Trigueros T, Laín A, Castaño D & Millet O (2009) Structural basis for the amino acid composition of proteins from halophilic archea. *PLoS Biol* **7**, e1000257.
9. Timasheff SN (1992) Solvent effects on protein stability. *Curr Opin Struct Biol* **2**, 35–39.
10. Ghosh T, Kalra A & Garde S (2005) On the salt-induced stabilization of pair and many-body hydrophobic interactions. *J Phys Chem B* **109**, 642–651.
11. Leigh JA & Dodsworth JA (2007) Nitrogen regulation in bacteria and archaea. *Annu Rev Microbiol* **61**, 349–377.
12. Forchhammer K (2008) P_{II} signal transducers: novel functional and structural insights. *Trends Microbiol* **16,** 65–72.

13.   Radchenko M & Merrick M (2011) The role of effector molecules in signal transduction by PII proteins. *Biochem Soc Trans* **39**, 189–194.

14.   Maier S, Schleberger P, Lü W, Wacker T, Pflüger T, Litz C & Andrade SLA (2011) Mechanism of disruption of the Amt-GlnK complex by PII-mediated sensing of 2-oxoglutarate. *PLoS ONE* **6**, e26327.

15.   Zeth K, Fokina O & Forchhammer K (2014) Structural basis and target-specific modulation of ADP sensing by the *Synechococcus elongatus* PII signaling protein. *J Biol Chem* **289**, 8960-8972.

16.   Muro-Pastor MI, Reyes JC & Florencio FJ (2001) Cyanobacteria perceive nitrogen status by sensing intracellular 2-oxoglutarate levels. *J Biol Chem* **276**, 38320–38328.

17.   Laurent S, Chen H, Bédu S, Ziarelli F, Peng L & Zhang CC. (2005) Non metabolizable analogue of 2-oxoglutarate elicits heterocyst differentiation under repressive conditions in *Anabaena sp.* PCC 7120. *Proc Natl Acad Sci USA* **102**, 9907-9912.

18.   Xu Y, Cheah E, Carr PD, van Heeswijk WC, Westerhoff HV, Vasudevan SG & Ollis DL (1998) GlnK, a PII-homologue: structure reveals ATP binding site and indicates how the T-loops may be involved in molecular recognition. *J Mol Biol* **282**, 149–165.

19.   Gruswitz F, O'Connell J 3rd & Stroud RM (2007) Inhibitory complex of the transmembrane ammonia channel, AmtB, and the cytosolic regulatory protein, GlnK, at 1.96 Å. *Proc Natl Acad Sci USA* **104**, 42–47.

20.   Conroy MJ, Durand A, Lupo D, Li X-D, Bullough PA, Winkler FK & 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.

21.   Llácer JL, Contreras A, Forchhammer K, Marco-Marín C, Gil-Ortiz F, Maldonado R, Fita I & Rubio V (2007) The crystal structure of the complex of PII and acetylglutamate kinase reveals how PII controls the storage of nitrogen as arginine. *Proc Natl Acad Sci USA* **104**, 17644–17649.

22.   Llácer JL, Espinosa J, Castells MA, Contreras A, Forchhammer K & Rubio V (2010) Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. *Proc Natl Acad Sci USA* **107**, 15397–15402.

23.   Rajendran C, Gerhardt EC, Bjelic S, Gasperina A, Scarduelli M, Pedrosa FO, Chubatsu LS, Merrick M, Souza EM, Winkler FK, Huergo LF & Li XD (2011) Crystal structure of the GlnZ-DraG complex reveals a different form of PII-target interaction. *Proc Natl Acad Sci USA* **108**, 18972–18976.

24.   Rodríguez-Valera F, Juez G & Kushner D (1983) *Halobacterium mediterranei* spec.nov., a new carbohydrate-utilizing extreme halophile. *Syst Appl Microbiol* **4**, 369–381.

25.   Torreblanca M, Rodríguez -Valera F, Juez G, Ventosa A, Kamekura M & Kates M (1986) Classification of non-alkaliphilic halobacteria based on numerical taxonomy and polar lipid composition, and description of *Haloarcula* gen. nov. and *Haloferax* gen. nov. *Syst Appl Microbiol* **8**, 89–99.

26.   Pedro-Roig L, Camacho M & Bonete MJ (2011) In vitro proof of direct regulation of glutamine synthetase by GlnK proteins in the extreme halophilic archaeon *Haloferax mediterranei*. *Biochem Soc Trans* **39**, 259–262.

27.   Pedro-Roig L, Camacho M & Bonete MJ (2013) Regulation of ammonium assimilation in *Haloferax mediterranei*: interaction between glutamine synthetase and two GlnK proteins. *Biochim Biophys Acta* **1834**, 16–23.

28.   Stadtman ER (2001) The story of glutamine synthetase regulation. *J Biol Chem* **276**, 44357–44364.

29.   Ehlers C, Weidenbach K, Veit K, Forchhammer K, & Schmitz RA (2005) Unique mechanistic features of post-translational regulation of glutamine synthetase activity in *Methanosarcina mazei* strain Gö1 in response to nitrogen availability. *Mol Microbiol* **55**, 1841–1854.

30.   Yildiz O, Kalthoff C, Raunser S & Kühlbrandt W (2007) Structure of GlnK1 with bound effectors indicates regulatory mechanism for ammonia uptake. *EMBO J* **26**, 589–599.

31.   Helfmann S, Lü W, Litz C & Andrade SLA (2010) Cooperative binding of MgATP and MgADP in the trimeric PII protein GlnK2 from *Archaeoglobus fulgidus*. *J Mol Biol* **402**, 165–177.

32.   Litz C, Helfmann S, Gerhardt S & Andrade SLA (2011) Structure of GlnK1, a signalling protein from *Archaeoglobus fulgidus*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **67**, 178–181.33. Han J, Zhang F, Hou J, Liu X, Li M, Liu H, Cai L, Zhang B, Chen Y, Zhou J, Hu S & Xiang H (2012) Complete genome sequence of the metabolically versatile halophilic archaeon *Haloferax mediterranei*, a poly(3-hydroxybutyrate-co-3-hydroxyvalerate) producer. *J Bacteriol* **194**, 4463-4464.

34.   Pedro-Roig L, Camacho M & Bonete MJ (2013) *Haloferax mediterranei* GlnK proteins are post-translationally modified by uridylylation. *Proteomics* **13**, 1371–1374.

35.   Paul S, Bag SK, Das S, Harvill ET & Dutta C (2008) Molecular signature of hypersaline adaptation: insights from genome and proteome composition of halophilic prokaryotes. *Genome Biol* **9,** R70.

36.   Tanford C (1980) *The hydrophobic effect.* J Wiley & Sons, New York.

37.   Tanford C (1961) *Physical chemistry of Macromolecules*. J Wiley & Sons, New York.

38.   Huergo LF, Merrick M, Pedrosa FO, Chubatsu LS, Araújo LM & Souza EM (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.

39.   Huergo LF, Merrick M, Monteiro RA, Chubatsu LS, Steffens MBR, Pedrosa FO & Souza EM (2008) In vitro interactions between the PII proteins and the nitrogenase regulatory enzymes dinitrogenase reductase ADP-ribosyltransferase (DraT) and dinitrogenase reductase-activating glycohydrolase (DraG) in *Azospirillum brasilense*. *J Biol Chem* **284**, 6674–6682.

40.   Radchenko MV, Thornton J & Merrick M. (2013) P_{II} signal transduction proteins are ATPases whose activity is regulated by 2-oxoglutarate. *Proc Natl Acad Sci USA* **110**, 12948–12953.

41.   Truan D, Huergo LF, Chubatsu LS, Merrick M, Li X-D & Winkler FK (2010) A new PII protein structure identifies the 2-oxoglutarate binding site. *J Mol Biol* **400**, 531–539.

42.   Fokina O, Chellamuthu V-R, Forchhammer K & Zeth K (2010) Mechanism of 2-oxoglutarate signaling by the *Synechococcus elongatus* PII signal transduction protein. *Proc Natl Acad Sci USA* **107**, 19760–19765.

43.   Sakai H, Wang H, Takemoto-Hori C, Kaminishi T, Yamaguchi H, Kamewari Y, Terada T, Kuramitsu S, Shirouzu M & Yokoyama S (2005) Crystal structures of the signal transducing protein GlnK from *Thermus thermophilus* HB8. *J Struct Biol* **149**, 99–110.

44.   Qvist J, Ortega G, Tadeo X, Millet O & Halle B (2012) Hydration dynamics of a halophilic protein in folded and unfolded states. *J Phys Chem B* **116**, 3436–3444.

45.   Ninfa AJ & Atkinson MR (2000) PII signal transduction proteins. *Trends Microbiol* **8**, 172–179.

46.   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.

47.   Marg B-L, Schweimer K, Sticht H & Oesterhelt D (2005) A two-alpha-helix extra domain mediates the halophilic character of a plant-type ferredoxin from halophilic archaea. *Biochemistry* **44**, 29–39.

48.   Evilia C & Hou Y-M (2006) Acquisition of an insertion peptide for efficient aminoacylation by a halophile tRNA synthetase. *Biochemistry* **45**, 6835–6845.

49.   Moure VR, Razzera G, Araújo LM, Oliveira MAS, Gerhardt ECM, Müller-Santos M, Almeida F, Pedrosa FO, Valente AP, Souza EM & Huergo LF (2012) Heat stability of proteobacterial PII protein facilitate purification using a single chromatography step. *Protein Expr Purif* **81**, 83–88.

50. Reed CJ, Lewis H, Trejo E, Winston V & Evilia C (2013) Protein adaptations in archaeal extremophiles. *Archaea* **2013**, 1–14.

51. Villeret V, Clantin B, Tricot C, Legrain C, Roovers M, Stalon V, Glansdorff N, Van Beeumen J (1998) The crystal structure of *Pyrococcus furiosus* ornithine carbamoyltransferase reveals a key role for oligomerization in enzyme stability at extremely high temperatures. *Proc Natl Acad Sci USA* **95**, 2801-2806.

52. Pedro-Roig L, Lange C, Bonete MJ, Soppa J & Maupin-Furlow J. (2013) Nitrogen regulation of protein-protein interactions and transcript levels of GlnK PII regulator and AmtB ammonium transporter homologs in Archaea. *MicrobiologyOpen* **2**, 826-840.

53. Bradford MM. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254

54. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132.

55. Battye TGG, Kontogiannis L, Johnson O, Powell HR & Leslie AGW (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271–281.

56. Evans P (2005) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72–82.

57. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.

58. Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255.

59. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.

60. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC & Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221.

61. Lovell SC, Davis IW, Arendall WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS & Richardson DC (2003) Structure validation by Cα geometry: φ, ψ and Cβ deviation. *Proteins* **50**, 437–450.

62. Fraczkiewicz R & Werner B (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J Comput Chem* **19**, 319–333.

63. Pace CN, Vajdos F, Fee L, Grimsley G & Gray T (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* **4**, 2411–2423.

64. Beutler HO & Supp M (1988) Coenzymes, metabolites and other biochemical reagents. In *Methods of Enzymatic Analysis, volume II* (Bergmeyer HU, Bergmeyer J & Grassl M, eds.), pp. 328-393 VCHVerlagsgesellschaft mbH, Weinheim, Germany.

65. Baker NA, Sept, D, Joseph S, Holst MJ & McCammon JA (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA*, **98**, 10037–10041.

Table 1. X-Ray data collection and refinement statistics.

Crystal	hmGlnK2-AMP	hmGlnK2-ADP	hmGlnK2-ATP
PDB file identifier	4OZL	4OZJ	4OZN
Data collection[a]			
Synchrotron	DLS	DLS	ESRF
Beamline (synchrotron)	I-03	I-04	ID23-2
Wavelength (Å)	0.9618	0.9795	0.8726
Space group	*R*32 (H32)	*R*32 (H32)	$P3_121$
Unit cell parameters			
a, b, c (Å)	89.1, 89.1, 95.3	89.2, 89.2, 96.3	110.7, 110.7, 76.9
α, β, γ, (º)	90, 90, 120	90, 90, 120	90, 90, 120
Solvent (%)	64.2	64.5	69.5
Resolution range (Å)	59.96 - 1.49 (2.55 - 1.49)	60.26 - 1.45 (1.47-1.45)	47.90 - 2.60 (2.74 - 2.60)
Reflections, total/unique	471382/23723	196084/25401	98157/17092
Completeness (%)	100 (99.9)	96.8 (79.8)	100 (100)
Multiplicity	19.9 (19.8)	7.7 (2.7)	5.7 (5.7)
I/σ_I	37.3 (4.5)	21.9 (1.9)	8.2 (3.1)
R_{pim}[b] (%)	1.6 (31.1)	1.7 (36.5)	5.8 (19.2)
Refinement			
Resolution	59.96 - 1.49	60.26 - 1.45	47.90 - 2.60
Reflections, work/test	22537/1186	24131/1270	16237/855
R_{factor}[c]/R_{free}[d] (%)	16.14/18.42	16.76 / 18.39	21.67/24.35
r.m.s.d. from ideal			
Bond length (Å)	0.006	0.008	0.005
Bond angle (º)	1.162	1.423	1.139
Number of			
Polypeptide chains	1	1	3
Protein atoms	787	758	2376
Nucleotide molecules	1 (AMP)	1 (ADP)	3 (ATP)
Solvent molecules	62	72	28
Average B-factor (Å^2)			
Protein	24.89	28.83	45.44
Nucleotide	17.1 (AMP)	28.17 (ADP)	53.5 (ATP)
H_2O	32.24	41.60	37.69
Ramachandran plot[e] (%)			
Favoured	99.0	99.0	98.4
Allowed	1.0	1.0	1.6
Outliers	0.0	0.0	0.0

[a]Values in parentheses are the data for the highest resolution shell.

[b]$R_{p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2} \sum_i |I_{hkl, i} - \langle I_{hkl} \rangle| / \sum_{hkl} \langle I_{hkl} \rangle$

[c]$R_{factor} = \sum_{hkl}||F_{obs}|-|F_{calc}|| / \sum_{hkl}|F_{obs}|$ where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

[d]R_{free} is the same as R_{factor} but calculated for the 5% data omitted from the refinement.

[e]Calculated using RAMPAGE [61]

Supporting information. Additional supporting information may be found in the online version of this article at the publisher's web site:

Table S1. List of PII proteins with an N-terminal extension longer than 3 residues.

Table S2. List of proteins annotated as GlnK in the Comprehensive Microbial Resource
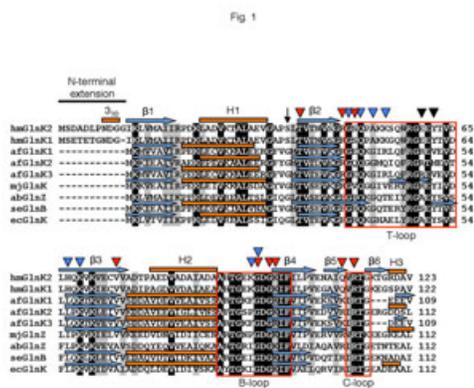
Fig. 1. Alignment of the sequences of *H. mediterranei* (hm) GlnKs with those of PII proteins of known structure from Archaea (*A. fulgidus*, af; *M. jannaschii*, mj) and from two representative bacteria, *S. elongatus* PCC7942 (se) and *E. coli* (ec). UniProt K/Protein Databank (when available) entries are: hmGlnK1, B8ZYW0; hmGlnK2, B8ZYW1/this work; afGlnK1, O29284/3O8W; afGlnK2, O28527/3NCR; afGlnK3, O28524/3T9Z; mjGlnK, Q60381/2J9C; seGlnB, P0A3F4/1QY7; ecGlnK, P0AC55/2GNK. Invariance and conservation are highlighted in black and grey shadowing, respectively. Helices and β-strand are shown as orange bars and blue arrows, respectively, above the sequences of those PII proteins for which the structure has been determined, and, are labeled when constant. The 3_{10} turn at the N-terminal extension (labeled) of hmGlnK2 is also indicated. The T-, B- and C-loops are labeled (below the stacked sequences) and are enclosed in red squares. Red triangles point at residues that contact the nucleotide in the structures of all these PII proteins. Blue triangles indicate 2OG-contacting residues in afGlnK3 and seGlnB. Black triangles indicate the positions at which postranslational modifications have been reported in some PII proteins. The vertical arrow points to Ser38 of hmGlnK2, the residue that participates in the sandwiching of the adenine ring.
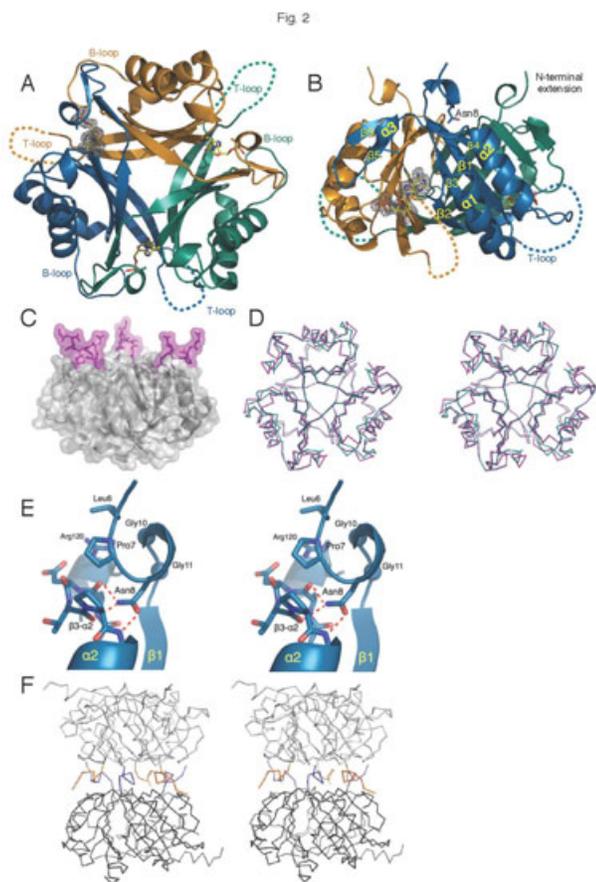
Fig. 2. The hmGlnK2 structure. (A) and (B) Overall structure (cartoon representation) of the AMP-bound form, viewed (A) along the molecular threefold axis with the N-ends close to the viewer, or (B) with the axis running vertically within the plane of the paper, with the N-ends at the top. Each subunit is represented in a different color. B-loops and T-loops are marked. Since the T-loops are not visible in the crystal, they have been arbitrarily represented as broken lines. In (B), major secondary elements are labeled. (C) Semi-transparent surface representation of a PII trimer enclosing the cartoon representation of the structure (in grey) to highlight the protruding N-terminal extensions (colored magenta, with the model in sticks and showing side-chains). (D) Stereo view of superimposition of the backbone of the model for the structure of the trimer of hmGlnK2-ATP (in green) with that for *A. fulgidus* GlnK2 complexed with ATP (in magenta, PDB entry, 3NCQ) without the T-loops and the N-terminal extensions of hmGlnK2. (E) Stereo view of the N-terminal extension, to detail its interactions with the PII body of the same subunit. Dashed red lines denote hydrogen bonds. (F) Stereo view of the hexamer formed by contacts mediated by the N-terminal extensions (colored blue for one trimer and yellow for the other trimer) in the ATP-containing crystal. The protein is shown in backbone representation in two shades of grey, one for each trimer.
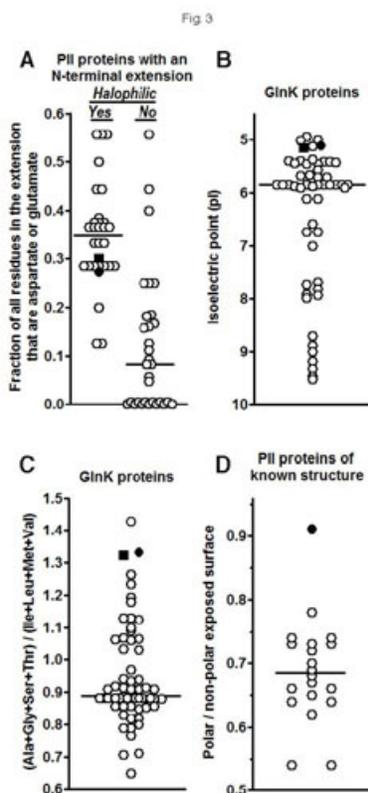
Fig. 3. Potential haloadaptation determinants of hmGlnKs. In all four panels an horizontal line indicates the median. Closed squares and circles correspond, respectively, to data for hmGlnK1 and hmGlnK2. (A) Fractional richness of negatively charged residues in the N-terminal extensions of the PII proteins predicted from the gene sequence to have such an extension of more than three residues (identified in Supplementary Table S1), classified as belonging to halophilic or to non-halophilic organisms, as indicated. (B, C) Isoelectric points (pI) estimated from the protein sequences (http://web.expasy.org/protparam), and (C) ratio of the number of small residues versus the number of large aliphatic non-polar residues, for the GlnK proteins annotated as such in the now closed Comprehensive Microbial Resource (http://www.jcvi.org/cms/research/past-projects/cmr). These proteins are listed in Supplementary Table S2. (D) Ratio of polar versus non-polar components of the exposed surfaces of the reported structures of PII proteins. PII trimers without T-loops (corresponding to *E. coli* PII residues 37-54) and ligands, were analyzed with program GETAREA [62]. The black circle corresponds to the value for hmGlnK2, and the open circles correspond to PDB entries, ordered from a higher to a lower polar/non polar exposed surface ratio, 2EG1, 2J9C, 3L7P, 3O8W, 3T9Z, 3NCP, 3MHY, 1V3R, 2GW8, 3N5B, 2XUL, 3BZQ, 2PII, 3CE8, 1HWU, 2O66, 1UL3, 1O51 and 1GNK.

.
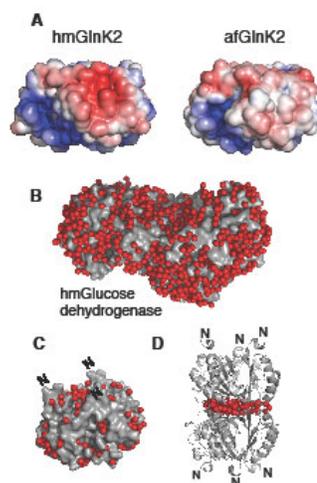
Fig. 4. Comparison of surface polarity and hydration layer of hmGlnK2 with those of relevant proteins. (A) Electrostatic surface representation of GlnK2 trimers from *H. mediterranei* (hmGlnK2) and *A. fulgidus* (afGlnK2; PDB entry 3NCR), without, respectively, the N-terminal extension (first 11 residues) or C-terminal tag (residues following Leu112). Ligands and T-loops (residues 48-65 and 37-54 of hmGlnK2 and afGlnK2, respectively) were also absent. Electrostatic potentials were calculated using APBS [65]. Polar surfaces are coloured blue (positively charged) and red (negatively charged). (B) Surface representation of the *H. mediterranei* glucose dehydrogenase dimer at 1.6 Å resolution [5], to ilustrate that its hydration layer (water molecules shown as red spheres) is much more dense than in the case of hmGlnK2, which is represented in the same way in (C), in an orientation in which the threefold axis crosses obliquely the plane of the paper. The three N-terminal extensions are labeled with an N. (D) Cartoon representation of two PII trimers in the crystal with ADP, to illustrate the cushion of water molecules (red spheres) between both trimers. The trimers are oriented in such a way that the N-terminal extensions (labeled with an N) look towards the top (one trimer) and bottom (the other trimer). The water molecules concentrate in the region between the exit of the T-loops.

Fig. 5
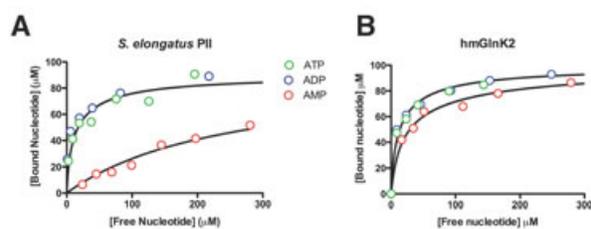
Fig. 5. Binding of ATP, ADP and AMP to *S. elongatus* PII protein and to hmGlnK2. For details on the binding assays see Materials and Methods. The concentrations of the PII proteins were 93 and 99 μM (as individual subunits) for *S. elongatus* PII and hmGlnK2, respectively. The curves are those for these total concentrations of sites, for a three-site binding model (one site per subunit of the trimer) with K_d values for the top curves in both panels of, 4, 12.5 and 47.4 μM, which are the values reported for ATP for *S. elongatus* PII [15]. In (A) the bottom curve is that for the same K_d, for the three sites, with a value of 203±20 μM. In (B) the bottom curve is that for K_d values for the three sites of 10.6, 19.3 and 133.4 μM, the values of these constants reported for ADP [15] for *S. elongatus* PII protein.

Fig. 6
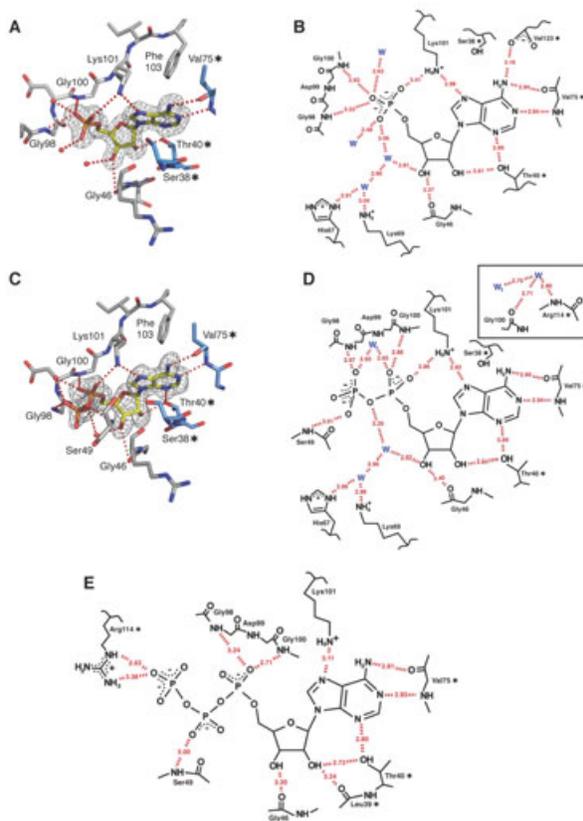
Fig. 6. Nucleotide binding to hmGlnK2. (A, C) Sticks representation of AMP (A) or ADP (C) binding. The Fo-Fc omit map for the nucleotide, contoured at 2.5σ, is shown as a grid enclosing the model for the nucleotide. Water molecules are shown as red spheres, and hydrogen bonds are represented by dashed red lines. Nucleotide and protein are shown in sticks representation, with O, N and P atoms, red, blue and orange, respectively, and C atoms yellow (nucleotide) grey (protein chain A) or cyan (protein chain B, also marked with asterisks). (B,D,E) Polar contacts of the nucleotide with the protein and the surrounding water in the crystals of hmGlnK2-AMP (A), hmGlnK2-ADP (D) and hmGlnK2-ATP (E). Asterisks mark chain B residues. Hydrogen bonds are represented by red broken thick lines with distances in Å. Water molecules are denoted by the letter W (in blue). The inset in B details some of the water sructure around the protein and the nucleotide, since W1 is shown in the main figure interacting with the nucleotide and in the inset interlinked to the protein via another fixed water molecule.

Fig. 7
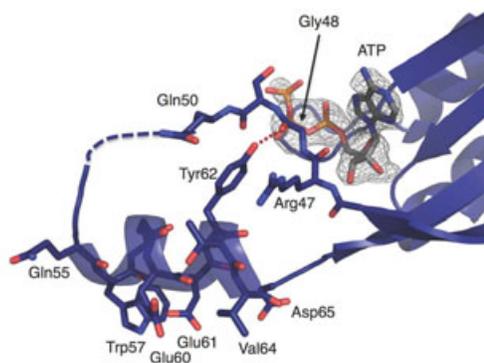
Fig. 7. The helical T-loop as seen in subunit A of the hmGlnK2-ATP complex, where the Fo-Fc omit map for the nucleotide, contoured at 2.4 σ, is shown as a grid enclosing the ATP model. The helix is shown in cartoon semi-transparent representation overlapping with the sticks-representation of the loop including its side-chains. The hydrogen bond between the main-chain O of Gly48 and Tyr62 is depicted by a red dotted line. A blue broken line represents the three missing residues (51-53) of the T-loop.

Fig. 8
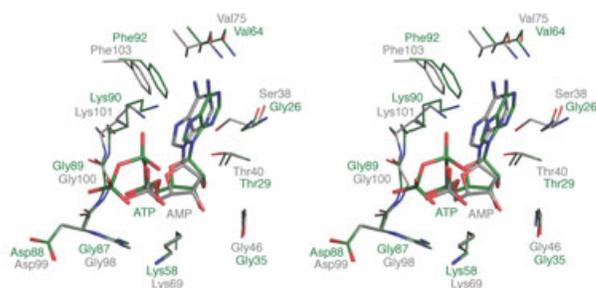
Fig. 8. Stereoview of the superimposed residues (thinner sticks) making the AMP site (grey carbons) and the ATP site (green carbons) in hmGlnK2 and *S. elongatus* PII, respectively, with the bound nucleotides shown in thicker sticks with their C and P atoms grey in the case of AMP and green for ATP. O and N atoms are shown red and blue, respectively.