Functional dissection of N-acetylglutamate synthase (ArgA) of *Pseudomonas aeruginosa* and restoration of its ancestral N-acetylglutamate kinase activity

Running title: Domain functions of acetylglutamate synthase

Authors: **Enea Sancho-Vaello[†], María L. Fernández-Murga[†,*] and Vicente Rubio[#]**

Instituto de Biomedicina de Valencia (IBV-CSIC) and Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER-ISCIII), C/ Jaime Roig 11, 46010 Valencia, Spain

Correspondent footnote

[#]Dr. Vicente Rubio

Instituto de Biomedicina de Valencia (IBV-CSIC)

C/ Jaime Roig 11

46010-Valencia, Spain

Phone: +34 96 339 17 72

Fax: +34 96 369 08 00

E-mail: rubio@ibv.csic.es

Other footnotes

[†] These authors contributed equally to this work.

[*]Present address of MLF-M, Fundación para la Investigación en el Hospital Universitario Dr.Peset, Avda. Gaspar Aguilar 90, Valencia-46017, Spain

ABSTRACT

In many microorganisms the arginine biosynthesis first step is catalyzed by classical N-acetylglutamate synthase (NAGS), an enzyme composed of N-terminal amino acid kinase (AAK) and C-terminal histone acetyltransferase (GNAT) domains that bind, respectively, the feed-back inhibitor arginine and the substrates. In NAGS three AAK domain dimers are interlinked by their N-terminal helices, conforming a hexameric ring, whereas each GNAT domain sits on the AAK domain of an adjacent dimer. Arginine inhibition of *Pseudomonas aeruginosa* NAGS was strongly hampered, abolished or even reverted to modest activation by changes in the length/sequence of the short linker connecting both domains, supporting a crucial role of this linker in arginine regulation. Linker cleavage or recombinant domain production allowed isolation of each NAGS domain. The AAK domain was hexameric and inactive, whereas the GNAT domain was monomeric/dimeric and catalytically active, although with ~50-fold increased and ~3-fold decreased $K_m^{glutamate}$ and k_{cat} , respectively, with arginine not influencing its activity. Deletion of N-terminal residues 1-12 dissociated NAGS to active dimers catalyzing the reaction with identical substrate kinetics and arginine insensitivity as the GNAT domain. Therefore, the interaction between the AAK and GNAT domains from different dimers modulates GNAT domain activity, whereas the hexameric architecture appears essential for arginine inhibition. We proved the closeness of the AAK domains of NAGS and N-acetylglutamate kinase (NAGK), the enzyme that catalyzes the next arginine biosynthesis step, shedding light on the origin of classical NAGS, by showing that a double mutation (M26K/L240K) in the isolated NAGS AAK domain elicited NAGK activity.

INTRODUCTION

In most microorganisms N-acetyl-L-glutamate synthase (NAGS), the enzyme encoded by *argA*, catalyzes the first step of arginine biosynthesis (Fig. 1A), producing N-acetyl-L-glutamate (NAG) from glutamate and acetyl coenzyme A (acetyl-CoA), and being feed-back inhibited by arginine (Fig. 1A) (8,10). Plants also use NAGS for making arginine (35). Although NAG is not a precursor of arginine in animals (2), NAGS is present in animals that make urea (8), since NAG is an essential activator of the urea cycle enzyme carbamoyl phosphate synthetase I (31) (Fig. 1B). Consequently, human NAGS deficiency is an inborn error of the urea cycle that causes clinical hyperammonaemia (7).

Although the key role and widespread distribution of NAGS in all domains of life warrants study of this enzyme, most detailed data concern the bacterial forms of this enzyme (17, 23,27,32,33) including the crystal structure of the NAGS from *Neisseria gonorrhoeae* (NgNAGS) in substrate-bound and arginine-bound forms (24,34). NgNAGS can be considered a typical example of classical bacterial NAGSs as defined by the early studied *Escherichia coli* and *Pseudomonas aeruginosa* enzymes (17, 23, 32, 33). These classical bacterial forms, encoded by *argA*, consist (Fig. 2A) of a single polypeptide of ~50 kDa mass that is composed of an N-terminal ~260-residue domain and a C-terminal ~150-residue domain, belonging, respectively to the amino acid kinase (AAK) and the histone acetyltransferase (GNAT) families (32). The NgNAGS structure proved that the GNAT domain binds both substrates (34), whereas the AAK domain was shown to host the site for the feed-back inhibitor arginine (24,34). As predicted from the homology of classical bacterial NAGS with arginine-sensitive NAG kinase (NAGK) (28), the NgNAGS structure is nucleated by a ring-like hexameric trimer of dimers of AAK domains that resembles closely the NAGK hexamer (34) (Figs. 1A and S1). NAGK catalyzes the next step

in the route of arginine synthesis in many microbes and in plants (10,35) (Fig. 1A) and is considered a paradigm for the AAK domain family (28,29). In both NAGS and arginine-sensitive NAGK a kinked α helix emerging at the N-end of each AAK domain is interlaced with the corresponding N-helix of an adjacent dimer, linking the three AAK domain dimers into the hexamer (28,34) (Fig. S1).

In bacterial NAGS and arginine-sensitive NAGK the feed-back inhibitor arginine sits in the AAK domain (24,28). The arginine sites flank the junctions between the AAK dimers, next to the kinked N-helices, with participation in each arginine site of the short N-helix segment following the kink. By binding at its site, arginine widens the hexameric ring of AAK domains (24,28). In NAGK, this binding causes inhibition because it favors an open conformation of the AAK domain, where catalysis takes place, distorting and widening the active center (28). However, in NAGS both substrates bind in the GNAT domain (34) and this inhibitory mechanism cannot apply. The globular GNAT domain of NgNAGS, connected to its cognate AAK domain by a 5-residue linker (Figs. 2B,C), sits on the AAK domain of an adjacent dimer (34) (Figs. 1A and S1A). Judged from site-directed mutagenesis studies (32), the AAK domain may play a regulatory role on NAGS activity. In agreement with this view, the NgNAGS crystal structures revealed that arginine dramatically changes the spacial relations between the AAK and GNAT domains (24). Thus, in the arginine-bound form of NgNAGS the GNAT domain interacts with the AAK domain of its own subunit and experiences a 109º rotation around its linker (Fig. 2B), drastically altering its interactions with the AAK domain on which it lies (24). Our prior studies (33) on the effects of linker shortening or lengthening by up to two residues on *P. aeruginosa* NAGS (PaNAGS) kinetic parameters support the importance of this rotation around the linker. NAGS inhibition by arginine involves a decrease in V_{max} as well as an increase in the

apparent K_m for glutamate (33). The latter effect fits the observation that in the arginine-bound form of NgNAGS two loops of the glutamate site become disordered (24). It also fits the finding that two-residue linker shortening, expected to drag away the GNAT domain from its normal position, mimics arginine in increasing the K_m for glutamate (33).

In the present work we subject to experimental corroboration using PaNAGS some key functional inferences that were based on the NgNAGS structure, as well as our earlier proposal (33) that the interdomain linker plays a paramount role on NAGS functionality. We (Fig. 2C) engineer this linker, cleave it, isolate and study the properties of the individual domains, and also produce recombinantly the two domains in isolated form, showing that the GNAT domain alone can catalyze the reaction but exhibits low affinity for glutamate, and that this isolated domain is insensitive to arginine. We also prove that the AAK domain is closely related to NAGK by restoring some NAGK activity by mutating two residues of this domain. The importance of the hexameric organization is demonstrated by rendering the enzyme dimeric by deletion of the N-terminal helix up to the expected site for its kink (Fig. 2C). Our findings dissect functionally classical bacterial NAGS and help generate an integrated picture that may contribute to understand also mammalian NAGS, reported to have the same domain organization as the bacterial enzyme (27,32). In fact, by engineering the linker of PaNAGS, we manage to convert arginine from a potent inhibitor to a modest activator, rendering the effect of arginine on enzyme activity reminiscent of the more potent activation triggered by arginine on the NAGSs of terrestrial animals including humans (3, 6,36).

MATERIALS AND METHODS

Preparation of cloned DNA sequences encoding wild-type and engineered PaNAGS forms.

We previously reported the cloning in pET22b (from Novagene) of *argA* (gene PA5204, http://cmr.jcvi.org/tigr-scripts/CMR) from *P. aeruginosa* encoding wild type (WT) NAGS with a $GSLEH_6$ tail (Fig. 2A, top) (32). This plasmid will be called here pET^{NAGS}. Single and double point mutations (Fig. 2C) at adjacent positions in the interdomain linker were introduced by site-directed mutagenesis in pET^{NAGS}, using the Quickchange kit (from Stratagene) and utilizing appropriate mutagenic forward and reverse oligonucleotides (Table S1). The same approach was used to produce the isolated recombinant AAK domain (rAAK) (Fig. 2A) by replacing codon 287 of *argA* by a stop codon. Two further rounds of site-directed mutagenesis on the latter plasmid introduced into rAAK the M26K/L240K double mutation. The preparation of PaNAGS having ^{283}EAQAF replacing ^{283}EQF was reported earlier (33). The pET22b plasmid carrying this mutated *argA* form was used for preparation of other engineered enzyme forms having a linker of increased length (Fig. 2C). Thus, a single round of site-directed mutagenesis on this plasmid allowed the change of the mutant sequence from ^{283}EAQAF to ^{283}EAQGP. The resulting mutant plasmid was mutated further to the sequence ^{283}LFQGP (*PScore* form), and this last plasmid was used for insertion of three extra residues *"en bloc"* in a single mutagenic round to give the linker sequence ^{283}LEVLFQGP (*PSsite* mutant), which is engineered to host the entire PreScission protease [an engineered derivative of human rhinovirus 3C protease (9) provided by GE Healthcare] cleavage site. All these changes were carried out with the Quickchange system as above, using the primers indicated in Table S1.

The DNA sequence encoding the GNAT domain was cloned into pET26b (from Novagene) after PCR-amplification of codons 287-432 of *argA* from pET^{NAGS} using a forward primer (Table S1) that introduced an initial ATG codon as part of an *Nde*I site, and a reverse

primer introducing a *Xho*I site [the same primer used (32) for cloning of argA into pET^{NAGS}, called "GNAT reverse" in Table S1], allowing directional insertion by ligation into the corresponding sites of pET26b digested with these enzymes. BL21 (DE3) cells (from Novagene) transformed with this plasmid failed to express the GNAT domain. This plasmid was used as a template for PCR amplification of the GNAT domain-encoding sequence in a way permitting its fusion to the C-terminus of SUMO (the protein chimera encoded by this construction is called here rGNAT) (Fig. 2A). The primers used for this amplification (Table S1) restore the stop codon at its normal site (thus, the C-terminal His_6 tag was eliminated, although the SUMO moiety carries one such tag) and introduce respective *Bsa*I and *Bam*HI sites before and after the coding sequence, allowing directional insertion into the corresponding sites of pSUMO (from LifeSensors Inc).

The same strategy used for cloning into pET26b of the GNAT domain was used for pET22b cloning of *argA* carrying a deletion of codons 1-12 (called from here on Δ1-12) (Fig. 2A). For this purpose, we used in the PCR-amplification step a forward primer corresponding to nucleotides 20-51 of the coding sequence in which four nucleotides were changed to include a *Nde*I site (Table S1). In this way, the engineered gene encodes residues 13-432 of PaNAGS preceded by a methionine and followed by the C-terminal $GSLEH_6$ tail.

PCR-amplification was carried out with a high-fidelity thermostable DNA polymerase (Deep-Vent, from New England Biolabs). The correctness of all the constructs and mutants prepared here was confirmed by DNA sequencing.

Expression and purification of the protein constructs. We produced wild-type PaNAGS and its mutant forms including Δ1-12 as reported earlier (32) for the wild enzyme. The procedure used includes growing the transformed BL21(DE3) cells at 37ºC, to $OD^{600} \approx 0.5$ in Luria-Bertani

161 medium containing 0.1 mg/ml ampicillin, then keeping the culture standing 45 min in ice,
162 followed by adding 2% (v/v) ethanol and 0.02 mM isopropyl-β-δ-thiogalactoside (IPTG) and
163 continuing the culture with aeration overnight at 15ºC. However, for production of rGNAT and
164 rAAK, either wild type or with the M26K/L240K double mutation, the induction was for 3 hours
165 at 37ºC with 1mM IPTG.

166 Cells were harvested by centrifugation and subsequent steps were at 4ºC. They were
167 disrupted by sonication in a solution of 20 mM Na phosphate pH 8, 1 mM dithiothreitol (DTT),
168 0.5 M NaCl and 20 mM imidazole. Insoluble material was centrifuged away. There was an
169 abundant production of the recombinant proteins in soluble form, allowing purification to
170 essential homogeneity (Fig. 2D; point and linker mutants are not shown) by Ni-affinity
171 chromatography. His-Spin Trap centrifugal columns (GE Healthcare) were used when small
172 amounts of protein were required (32). For larger amounts, the cell pellet from a 0.5-L culture
173 was suspended in 15 ml of sonication buffer and the centrifuged sonicate was applied to a 1-ml
174 His Trap-HP column mounted on an ÄKTA FPLC system (both from GE Healthcare), eluting the
175 His_6-tagged protein with a 30-ml linear gradient of 20 mM-500 mM imidazole-containing buffer.
176 The different proteins were placed in storage buffer [10 mM sodium phosphate pH 7.0, 15% (v/v)
177 glycerol, 1 mM EDTA, 1 mM DTT, 20 mM NaCl and 10 mM NAG; based on (23)], by either
178 centrifugal desalting through PD SpinTrap G-25 columns (from GE Healthcare) or, for larger
179 volumes, by repeated cycles of centrifugal ultrafiltration (Amicon Ultra of 10K cutoff, from
180 Millipore) and dilution using the same buffer.

181 For purification of the rAAK domain and its double mutant, which had no His_6 tag, 0.5-L
182 cultures were used, the sonication buffer was 15 ml of 20 mM Na phosphate pH 8/1mM DTT,
183 and the postcentrifugal supernatant was subjected to sequential precipitations with ammonium

sulphate at 30% and 60% saturation, dissolving the final precipitate in 15 ml of sonication buffer followed by desalting by repeated centrifugal ultrafiltration as above using the same buffer. The protein solution was then applied to a 1-ml HiTrap Q HP column (GE Healthcare) mounted on an AKTA FPLC system equilibrated with sonication buffer, followed by washing and by elution of essentially pure rAAK (Fig. 2D) with a 35-ml linear gradient of 0-1 M NaCl in the same buffer. The protein was concentrated to 25 mg/ml by centrifugal ultrafiltration.

Cleavage with PreScission protease and separation of the two domains. PreScission protease (from GE Healthcare) and the wild type or the engineered *PScore* or *PSsite* forms of PaNAGS (Fig. 2C), at the indicated concentrations, were incubated 5 hours at 15ºC in a solution containing 50 mM Tris-HCl pH 7.1, 0.15 M NaCl, 1 mM EDTA, 1 mM DTT, monitoring the cleavage by SDS-PAGE. At the end of the incubation, the mixture was either centrifuged through a His Spin Trap column or, for separation of the two domains, 0.3 ml of the digestion or of an equivalent non-digested mixture were applied to a Superdex 200HR (10/300) column (GE Healthcare) mounted on an AKTA fast protein liquid chromatography system. The column was equilibrated and run (flow rate, 0.50 ml/min) with a solution at 4ºC of 50mM Tris-HCl, pH 8.5/0.1M NaCl, monitoring the absorbance of the effluent at 280 nm and collecting 0.5-1 ml fractions over ice. The fractions corresponding to the AAK domain peak or to the GNAT domain peak from two identical chromatographic experiments were pooled together and concentrated by centrifugal ultrafiltration to ~1 mg protein/ml.

Enzyme activity assays. NAGS activity was determined colorimetrically with Ellman´s reagent, as CoA release (12) at 37ºC, as previously reported (32), in a solution of 0.2 M Tris-HCl pH 9 containing, in the standard assay, 30 mM L-glutamate (sodium salt), 4 mM acetyl-CoA, and,

when used, the indicated concentration of arginine hydrochloride. For estimation of the kinetic parameters for the substrates, the acetyl-CoA concentration was fixed at 4 mM while varying glutamate, or the glutamate concentration was fixed at 30 mM (except where indicated) and acetyl-CoA was varied. For assay, the enzyme was appropriately diluted in storage buffer lacking both NAG and NaCl and supplemented with 30 mg/ml bovine serum albumin (albumin concentration in the assays, 1.5-3 mg/ml). Reactions were carried out at least in duplicate, and blanks in which the enzyme was replaced by dilution solution were run in parallel and were subtracted. One enzyme unit produces 1 μmol CoA min^{-1}. Results (mean ± SE) were fitted with GraphPadPrism (GraphPad Software, San Diego) to either hyperbolic kinetics or to substrate inhibition kinetics as reported (33).

NAGK activity was determined at 37ºC and pH 7.5 as previously described (13), using the hydroxylamine-containing colorimetric assay of Haas and Leisinger (18). One enzyme unit is the amount of enzyme that generates 1 μmol of product in 1 min.

Other techniques. SDS-PAGE (20) was carried out in 15% polyacrylamide gels. Protein was determined by the Bradford assay (4) using bovine serum albumin as standard. The structures of *E. coli* NAGK (Protein Databank, PDB, file 1GS5) (29) and of the AAK domain of NgNAGS (PDB file 2R8V) (34) were superposed with program Coot (Crystallographic Object-Oriented Toolkit) (11) using default parameters. Figures representing protein structures were generated using PyMOL (http://www.pymol.org).

RESULTS AND DISCUSSION

Influence of the interdomain linker on arginine modulation of NAGS activity

To adscribe functions to the GNAT domain,we engineered the interdomain linker (Fig. 2A,C; and see Materials and Methods) to introduce a cleavage site for the highly specific PreScission protease by replacing the EQF sequence within the linker connecting both domains by LEVLFQGP (Fig. 2A,C). The enzyme with this replacement, called here *PSsite* form, was expressed and purified similarly to the recombinant wild-type enzyme (not shown). This form was catalytically active, with kinetic parameters not too different from those of wild-type PaNAGS (Table 1), but it failed to be inhibited by arginine whereas the wild type enzyme is nearly totally inhibited by 2 mM arginine (Fig. 3A and Table 1). This observation confirms our prior conclusion (33) that the interdomain linker is highly important for arginine modulation of NAGS activity.

Surprisingly, when we engineered the linker sequence to replace the wild-type EQF linker sequence by LFQGP (called here *PScore* form) and particularly by EAQGP (Fig. 2C), a form prepared as an intermediate in the stepwise engineering of the linker from wild-type to *PSsite* form, the enzyme was modestly but significantly activated by arginine rather than being inhibited (Fig. 3A and Table 1). This effect, which is somewhat reminiscent of the activation by arginine of the NAGSs from terrestrial animals (6,19), suggests that the linker sequence might determine whether arginine is an inhibitor or an activator. The importance not only of the linker length (33) but also of the linker sequence in determining the effect of arginine is illustrated by the observation that the replacement in the modified linker of EAQGP by EAQAF (Fig. 2C) resulted in enzyme inhibition by arginine although with less potency and at higher arginine concentrations than in the case of the wild-type sequence EQF (Fig. 3A, and Table 1). Since all these linker

changes reduced somewhat V_{max} (Table1), the linker influences also the efficiency of the enzyme as a catalyst in the absence of arginine.

We attempted to replicate without success the arginine activation of the human enzyme by replacing the Q and F of the PaNAGS wild type linker sequence (EQF) by R and M, respectively, the two residues found at these positions in human NAGS (Fig. 2C). This change prevented enzyme inhibition by arginine, with perhaps a slight trend towards activation (Fig. 3B and Table 1). When enzyme forms with only one or the other of these mutations was studied, it was found that the EQF-to-EQM change (the changed residue is underlined) virtually abolished arginine inhibition, although without any evidence of arginine-triggered activation, whereas the EQF-to-ERF change failed to abolish inhibition by arginine, although the arginine concentration needed for inhibition was increased (Fig. 3B). In contrast, the replacement by A of the E residue that is common to human and bacterial NAGSs (change, EQF to AQF, Fig. 2C) did not alter the sensitivity of the enzyme to arginine (Fig. 3B and Table 1). These results clearly show that the phenylalanine found in the bacterial linker is essential for arginine inhibition. The crystal structures of NgNAGS (24,34) with and without arginine (Fig. 2B) support the importance of this phenylalanine, which would act as an end lever, stabilizing the orientations of the GNAT domain relative to the AAK domain across the linker in the arginine-free and arginine-bound enzyme conformations. Thus, the phenylalanine bencenic ring makes in the arginine-free form extensive and close contacts with the GNAT domain (Fig. 2B, left panel), whereas in the arginine-bound form it makes extensive contacts with the AAK domain (Fig. 2B, right panel). Our observation that the F285M mutation, although abolishing arginine inhibition, had virtually no effect on the kinetic parameters of the enzyme in the absence of arginine (Table 1), strongly supports the view that the methionine cannot fulfill this lever function, failing to stabilize the arginine-bound

conformation. In any case, our results show that the mere replacement of this phenylalanine or of the Q preceding it by its human counterparts is not enough for rendering arginine an activator.

We examined the reasons for the activation by arginine of the enzyme with the EAQGP linker sequence. In the presence of 30 mM arginine the glutamate dependency of the activity (Fig. 3A, top inset) revealed that only when glutamate concentrations exceeded 20 mM was the activity higher with arginine than without arginine. Below this glutamate concentration arginine was an inhibitor, revealing that arginine can be an inhibitor or an activator depending on the glutamate concentration. The effective glutamate concentration at which the enzyme reaches its experimental activity maximum is higher with arginine than without arginine, and substrate inhibition also appears to occur at increased glutamate concentrations when arginine is present (Fig. 3A, top inset). These changes are qualitatively similar to those triggered by arginine on the kinetics of wild-type PaNAGS [Fig. 3A, bottom inset and (33)], but their magnitude is smaller with the EAQGP form than with the wild type enzyme. However, in the case of the modified enzyme, arginine does not appear to trigger a reduction in the apparent V_{max} for glutamate, whereas kinetic results with the wild type enzyme were consistent with a nearly 7-fold reduction in V_{max} triggered by as little as 2 mM arginine [Fig. 3A, insets, and (33)]. This different effect on V_{max} appears to be the main reason for the observed activation by arginine of the enzyme with the modified linker.

Influence of linker cleavage on activity and arginine inhibition

PreScission protease cleaved the *PSsite* enzyme form at its interdomain linker, but it failed to cleave wild type PaNAGS (Fig. 4A) or the *PScore* form (not shown), as shown by SDS-PAGE, which revealed bands with the expected masses for the AAK (31 kDa) and GNAT (18 kDa) domains in the digestion of the *PSsite* form. Linker cleavage did not lead to immediate

dissociation of the two domains, since only a small fraction (~25%, Fig. 4B, *unretained*) of the cleaved AAK domain was not retained by centrifugation through a Ni-affinity column, whereas the majority was retained (Fig. 4B, *retained*), being eluted together with the His_6-tag-containing GNAT domain by the application to the centrifugal column of 0.5 M imidazole.

Linker cleavage did not inactivate the enzyme, but it caused a very large increase in the concentrations of glutamate required for activity (diamonds in Fig. 5A). Given the lack of immediate dissociation of the two domains, this strongly suggests that the physical continuity across the linker is crucial to endow the enzyme with its normal affinity for glutamate. Similarly to the uncleaved form of the *PSsite* form, the cleaved form remained insensitive to arginine (data not shown), suggesting that the physical continuity between both domains provided by the linker is a requisite for arginine modulation of enzyme activity.

Gel filtration separates the AAK and GNAT domains of the cleaved enzyme

Application of the PreScission-cleaved *PSsite* enzyme form (Fig. 2A) to a gel filtration column (Fig. 4C) resulted in the separation as individual peaks of the AAK and GNAT domains (Figs. 2D and 4C). We also succeeded in producing recombinantly the AAK domain by replacing codon 287 of the pET22b-encoded PaNAGS by a stop codon (Q287X mutation) (Figs.2A,D). The AAK domain generated by PreScission cleavage or produced recombinantly was eluted identically from the column, at a volume corresponding to hexamers (Fig. 4C). Similarly, as reported already (23,32), the uncleaved enzyme was also eluted as expected for an hexamer (Fig. 4C). These findings agree with the observation made in the crystal structure of NgNAGS (24,34) (Figs. 1A and S1A) of a hexameric enzyme architecture nucleated by a hexameric ring of AAK domains that resembles closely the hexameric ring of AAK domains forming the structure of arginine-sensitive bacterial NAGK (28) (Figs. 1A and S1B). In contrast to the hexameric

oligomerization of the enzyme and the AAK domain, the GNAT domain produced by PreScission cleavage appeared late in the effluent from the gel filtration column, at a position intermediate between those expected for monomers and dimers (Fig. 4C), although somewhat closer to that of the monomer (mass estimate by interpolation in the calibration line, 25.6 kDa, corresponding to 1.4 protomers). A monomeric architecture would agree with the lack of interactions between GNAT domains in the crystal structure of the NgNAGS hexamer (24,34).

The isolated GNAT domain catalyzes the NAGS reaction and is insensitive to arginine.

In agreement with the observation made on the NgNAGS structure that the sites for both substrates are located in the GNAT domain (34), only this domain exhibited NAGS activity (Figs. 5A,B, inverted closed triangles). However, the concentrations of glutamate required for activity with the isolated GNAT domain were much higher than those for the uncleaved *PSsite* form (~25-fold higher; compare in Fig. 5A, the main panel and the inset), agreeing with the results obtained (see previous section) with the cleaved *PSsite* form in which the two domains had not been separated (Fig. 5A, diamonds). In addition, as expected from the binding of arginine in the AAK domain of PaNAGS (32) and NgNAGS (24), the isolated GNAT domain was insensitive to arginine (Fig. 5C). Identical results were obtained, within experimental error (Fig 5, open circles), for the substrate dependency of the activity and lack of arginine sensitivity with the recombinant GNAT domain that we finally succeeded in producing as a chimera with SUMO (Figs. 2A and 2D).

The deletion of the N-terminal helix renders PaNAGS dimeric

The present results indicate that the AAK domain, when connected covalently to the GNAT domain by the normal or by mutated versions of the interdomain linker (including 5-

residue longer versions as in the engineered linker of the *PSsite* form), triggers on the GNAT domain a very important increase in the apparent affinity for glutamate. The structure of NgNAGS (34) reveals that each GNAT domain is connected to two or even to three [in the arginine bound-form (24)] AAK domains, one belonging to the same subunit, to which the GNAT domain is covalently linked, and the others from an adjacent dimer on which the GNAT domain lies in the hexameric structure. Therefore, we sought to clarify whether it was the cognate AAK domain or the ones from an adjacent dimer which are responsible for triggering the increased affinity for glutamate of the GNAT domain.

We decided to dissociate the PaNAGS hexamer to dimers by deleting the N-terminal helix, which, by interlacing with the corresponding N-helix from another dimer, links the three dimers into the hexamer (34) (Fig. S1A). In this way, the connection of each GNAT domain with its cognate AAK domain of the same subunit would be preserved, while the interaction with the AAK domains from another dimer would be lost if the hexamer were dissociated to its three composing dimers. We actually deleted only residues 1 to 12 (Fig. 2A), which should correspond to the larger and more N-terminal portion of the N-terminal helix, down to the kink before the second portion of the helix (34), which is shorter and is involved in arginine binding (24,33), thus expectedly preserving the arginine binding site. The resulting protein (called Δ1-12 form) was expressed as the wild-type enzyme, it was soluble and was purified easily (Fig. 2D), and it behaved in gel filtration (Fig. 6) as dimers, as expected, although with some skew towards larger elution volumes that may indicate coexistence with a small fraction of monomers. Thus, these results reveal hexamer dissociation, and therefore abolition of interdimeric contacts. The peak of dissociated enzyme exhibited enzyme activity (Fig. 6) with a constant ratio, within experimental error, of activity versus protein concentration (monitored as OD^{280}, Fig. 6, and confirmed by

SDS-PAGE, not shown), indicating that the wole peak consists of a single species in terms of specific activity. The substrate kinetic parameters for this enzyme form (Fig. 5A, B, open squares) appear identical, within experimental error, to those of the isolated recombinant or PreScission-cleaved GNAT domain when velocities are expressed as turnover number per polypeptide chain. Thus, single concentration-activity curves for glutamate and acetyl-CoA were fitted to the pooled values for these enzyme forms (Figs. 5A,B). The curve for glutamate, adjusted to hyperbolic kinetics with substrate inhibition, corresponds to ~50-fold increase and ~2-fold decrease in K_m and k_{cat}, respectively, relative to the wild-type enzyme (Fig. 5A, inset). The hyperbola fitted to the pooled data for acetyl-CoA (Fig. 5B), shows a modest (~2-fold) increase in $K_m^{\text{Acetyl-CoA}}$ relative to wild-type. The identical kinetic properties of dimeric PaNAGS and of the isolated GNAT domain indicate that the modulatory effect of the AAK domain on the activity of the GNAT domain results from the interactions that occur in the hexamer between GNAT domain from one dimer and AAK domains from another dimer. Furthermore, arginine had no effect on the activity of the enzyme with the helix deletion (Fig. 5C). Since the portion of the helix that is deleted is not involved in arginine binding (24), arginine would be expected to bind to the dimeric enzyme. Therefore, the lack of influence of arginine on enzyme activity possibly reflects the requirement of an hexameric organization for inhibition in classical NAGS.

The AAK domain of NAGS derives from an ancestral NAGK domain, as shown by triggering NAGK activity by a double point mutation

As already indicated, the isolated AAK domain did not exhibit NAGS activity, as expected. However, the similarity of this domain of NgNAGS with the enzyme NAGK (34), with preservation of the active center crevice (Fig. 7A), led us to investigate whether this domain had any NAGK activity. In fact, a bifunctional NAGS/NAGK has been identified in *Xanthomonas*

394 *campestris* (27), although its NAGK activity is quite low when compared with its own NAGS
395 activity or with the NAGK activities of classical bacterial NAGKs (13,22).

396 We failed to detect any NAGK activity with PaNAGS, even at very high enzyme
397 concentrations (up to 0.16 mg/ml in the assay; at this concentration the *X. campestris* enzyme
398 would have consumed all the NAG in the assay). This lack of NAGK activity might have been
399 expected, since *P. aeruginosa* has a separate gene (*argB*) that encodes a highly active arginine-
400 sensitive NAGK (28). Similarly, we failed to detect any NAGK activity with the isolated
401 recombinant AAK domain of PaNAGS, even at concentrations of 0.3 mg/ml in the assay (closed
402 circle in Figs. 7B and 7C) excluding that the lack of NAGK activity of the complete enzyme were
403 due to some inhibitory effect of the GNAT domain on that activity. However, when we mutated
404 to lysine two residues of the putative NAGK active center of the AAK domain of PaNAGS
405 (M26K/L240K double mutation), rAAK became active as a NAGK. These residues were chosen
406 because they are the counterparts in PaNAGS of two invariant NAGK active center lysines (K8
407 and K217 of *E. coli* NAGK, Fig. 7A) that play key catalytic roles in acetylglutamate
408 phosphorylation (16,29). Whereas no activity was observed when only one of the two mutations
409 was introduced in rAAK (results not shown), the double mutant exhibited clear although low
410 NAGK activity (Fig. 7B, C). The activity depended hyperbolically on the ATP concentration
411 (Fig. 7B), and also presented a non-linear dependency on the concentration of NAG (Fig. 7C),
412 although the apparent affinity for this substrate was too low to approach saturation. The apparent
413 K_m^{ATP} of the double mutant (1.6 ± 0.2 mM) was in the typical range for a bacterial NAGK in the
414 same type of hydroxylamine-based assay (13,18). In contrast, the K_m^{NAG} could not be estimated
415 accurately given its high value, since the range of NAG concentrations used did not reach the
416 levels needed for saturation. In any case, from the data of Fig. 7C it appears evident that the
417 K_m^{NAG} and the activity at saturation of this substrate are >0.225 M and >0.5 U/mg, respectively.

Actually, the best hyperbolic fit for the NAG data gives apparent K_m^{NAG} and V_{max} values of ~0.8 M and ~1 U/mg, respectively. The V_{max} value is not too far from the activity exhibited by the bifunctional *X. campestris* enzyme (27). However, the K_m^{NAG} is much higher than for genuine NAGKs (13,18), possibly reflecting a high tendency towards a closed conformation of the NAG site of the AAK domain of NAGS, as revealed by the lowered position of the lid for this site (the β3-β4 hairpin, Fig. 7A) in the NgNAGS structure (24), despite the emptiness of this site, whereas in NAGK this site only closes down when NAG is bound (15).

.

How and why two-domain NAGSs emergerd?

The closeness of the AAK domains of NAGS and NAGK goes beyond the overall domain fold, extending to the similarity of the NAGK active center crevice (Fig. 7A) and even to the conservation of a key NAG-binding arginine (R66 of *E. coli* NAGK; R85 of NgNAGS, Fig. 7A; R82 of PaNAGS) (29,34). Indeed, although with low affinity, the AAK domain of NAGS binds NAG, as attested by the NAGK activity of its M26K/L240K double mutant despite the fact that these mutations do no affect NAG site residues. An invariant aspartate (D162 in *E. coli* NAGK) that in NAGKs has a key active center-organizing role by coordinating the two catalytic lysines and by binding MgATP (16,29), is also preserved in many NAGSs as a glutamate (E185 of PaNAGS), although not in NgNAGS (34). The NAGK activity of the M26K/L240K double mutant indicates that classical NAGS lost its ancestral NAGK activity primarily because of the mutation of these two catalytic lysines (16,22,29). Indeed, as expected from its NAGK activity, the bifunctional NAGS/NAGK of *X. campestris* has these two lysines preserved (27).

Since *Mycobacterium tuberculosis* NAGS merely is a GNAT domain and yet it is feed-back inhibited by arginine (12), the AAK-GNAT domain organization of classical bacterial

NAGS is clearly unessential for arginine-regulated NAG synthesis from acetyl-CoA and glutamate. The proposal of Labedan's group (37) that the classical two-domain NAGSs arose from the fusion of a NAGS similar to that found in *M. tuberculosis* and an arginine-sensitive NAGK, is supported very strongly by our finding that a mere two-residue mutation in the AAK domain of PaNAGS can render this domain a NAGK. This proposal is also supported by the closeness of the structures of NAGK (28,29) and the AAK domain of NgNAGS (34) (Fig. 7A), and by the discovery of bifunctional NAGS/NAGKs having the same domain organization as PaNAGS (27). A question remaining to be answered concerning the evolutionary process from single domains to two-domain NAGS is whether the arginine sensitivity of *M. tuberculosis* NAGS (12) was a late evolutionary acquisition or whether this trait of the GNAT component was originally present and has been lost over the course of evolution of two-domain NAGSs.

A major advantage of associating AAK and GNAT domains into a two-domain NAGS that has become patent in our studies is that, because of the important modulatory role of the AAK domain on the activity of the GNAT domain, the K_m for glutamate is brought down from the very high values of the *M. tuberculosis* enzyme (12) to the mM range that is characteristic for the canonical microbial enzymes (17,23,24,27,32). This value is in these enzymes closer to the range of glutamate concentrations present in bacteria [see for example (25)], thus increasing the efficiency of the enzyme in the catalysis of the reaction in vivo. Another potential reason for shifting from a single-domain to a two-domain NAGS may be related to NAG channelling between NAGS and NAGK, since such channelling would prevent unwanted NAG hydrolysis by aminoacylases such as those being prevalent in animals (30) or existing (although nowadays in the periplasmic space) in *P. aeruginosa* (14). This channelling requires the direct interaction of the GNAT domain of one subunit with the AAK domain of another subunit, as observed in hexameric NAGS. A reminiscence of such channelling might be the association in yeasts of

NAGS and NAGK, both having the two-domain composition of PaNAGS (27), to form a metabolon when NAG is produced by NAGS (1,26). It would be important to clarify whether such channelling occurs in this metabolon and/or in the bifunctional NAGS/NAGKs.

A third potential advantage of the two-domain organization of NAGS is the possibility of modulating easily the arginine regulation of the NAGS activity. In two-domain NAGS, arginine regulation results from primary arginine-triggered changes on the architecture of the AAK domain hexamer (24). These changes are similar to those observed in hexameric NAGK upon arginine binding to this enzyme (28). Since with this latter enzyme a signalling protein, PII, was shown to modulate the sensitivity of the enzyme to arginine inhibition (21), it cannot be excluded that analogous regulatory mechanisms may exist for modulating the arginine sensitivity of canonical hexameric NAGSs. Furthermore, the key role of the interdomain linker as a mediator of the arginine effect that is revealed by our previous (33) and present studies provides the basis for adapting arginine regulation by mere changes in the 5-residue linker sequence. This endows AAK-GNAT domain NAGSs with an enormous potential for adapting to the specific physiological needs of different organism. The best example of this ability to adapt is provided by the change in the effect of arginine on NAGS from inhibition to activation with the shift of animals from marine life to terrestrial ureotelism (19), a change that we have partially reproduced by linker manipulation in the present studies.

Acknowledgements

This work was supported by grants BFU2008-05021 of the Spanish Ministry of Science (MEC and MICINN) and Prometeo/2009/051 of the Valencian Government. ES-V and . MLF-M were supported by contracts from the Instituto de Salud Carlos III and the JAE-DOC Programme of the Consejo Superior de Investigaciones Científicas. We thank N. Gougeard (IBV-CIBERER) and

the late E. Cogollos (IBV-CSIC) for technical assistance and F. Gil-Ortiz (IBV-CSIC) for help with Figs. 2B and 7A.

References

1. Abadjieva A, Pauwels K, Hilven P, Crabeel M. 2001. A new yeast metabolon involving at least the two first enzymes of arginine biosynthesis: acetylglutamate synthase activity requires complex formation with acetylglutamate kinase. J. Biol. Chem. **276**: 42869-42880.

2. Alonso E, Rubio V. 1989. Participation of ornithine aminotransferase in the synthesis and catabolism of ornithine in mice. Studies using gabaculine and arginine deprivation. Biochem. J. **259**:131-138.

3. Bachmann C, Krähenbühl S, Colombo JP. 1982. Purification and properties of acetyl-CoA:L-glutamate N-acetyltransferase from human liver. Biochem. J. **205**:123-127.

4. 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.

5. Briozzo P, Evrin C, Meyer P, Assairi L, Joly N, Barzu O, Gilles AM. 2005. Structure of *Escherichia coli* UMP kinase differs from that of other nucleoside monophosphate kinases and sheds new light on enzyme regulation. J. Biol. Chem. **280**:25533-25540.

6. Caldovic L, Lopez GY, Haskins N, Panglao M, Shi D, Morizono H, Tuchman M. 2006. Biochemical properties of recombinant human and mouse N-acetylglutamate synthase. Mol. Genet. Metab. **87**:226-232.

7. Caldovic L, Morizono H, Tuchman M. 2007. Mutations and polymorphisms in the human N-acetylglutamate synthase (NAGS) gene. Hum. Mutat. **28**:754-759.

8. Caldovic L, Tuchman M. 2003. N-acetylglutamate and its changing role through evolution. Biochem. J. **372**:279-290.

9. Cordingley MG, Callahan PL, Sardana VV, Garsky VM, Colonno RJ. 1990. Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro. J. Biol. Chem. **265**:9062-9065.

10. Cunin R, Glansdorff N, Pierard A, Stalon V. 1986. Biosynthesis and metabolism of arginine in bacteria. Microbiol. Rev. **50**:314–352.

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

12. Errey JC, Blanchard JS. 2005. Functional characterization of a novel ArgA from *Mycobacterium tuberculosis*. J. Bacteriol. **187**:3039–3044.

13. Fernández-Murga ML, Gil-Ortiz F, Llácer JL, Rubio V. 2004. Arginine biosynthesis in *Thermotoga maritima*: characterization of the arginine-sensitive N-acetyl-L-glutamate kinase. J. Bacteriol. **186**:6142-6149.

14. Früh H, Leisinger T. 1981. Properties and localization of N-acetylglutamate deacetylase from *Pseudomonas aeruginos*a. J. Gen. Microbiol. **125:**1-10.

15. Gil-Ortiz F, Ramón-Maiques S, Fernández-Murga ML, Fita I, Rubio V. 2010. Two crystal structures of *Escherichia coli* N-acetyl-L-glutamate kinase demonstrate the cycling between open and closed conformations. J. Mol. Biol. **399**:476-490.

16. Gil-Ortiz F, Ramón-Maiques S, Fita I, Rubio V. 2003. The course of phosphorus in the reaction of N-acetyl-L-glutamate kinase, determined from the structures of crystalline complexes, including a complex with an AlF_4^- transition state mimic. J. Mol. Biol. **331**:231-244.

17. Haas D, Kurer V, Leisinger T. 1972. N-acetylglutamate synthetase of *Pseudomonas aeruginosa*. An assay in vitro and feedback inhibition by arginine. Eur. J. Biochem. **31**:290-295.

18. Haas D, Leisinger T. 1975. N-acetylglutamate 5-phosphotransferase of *Pseudomonas aeruginosa*. Catalytic and regulatory properties. Eur. J. Biochem. **52**:377-393.

19. Haskins N, Panglao M, Qu Q, Majumdar H, Cabrera-Luque J, Morizono H, Tuchman M, Caldovic L. 2008. Inversion of allosteric effect of arginine on N-acetylglutamate synthase, a molecular marker for evolution of tetrapods. BMC Biochem. **9**:24.

20. Laemmli, UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. **227:**680-685.

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 P_{II} and acetylglutamate kinase reveals how P_{II} controls the storage of nitrogen as arginine. Proc. Natl. Acad. Sci. U S A **104**:17644-17649.

22. Marco-Marín C, Ramón-Maiques S, Tavárez S, Rubio V. 2003. Site-directed mutagenesis of *Escherichia coli* acetylglutamate kinase and aspartokinase III probes the catalytic and substrate-binding mechanisms of these amino acid kinase family enzymes and allows three-dimensional modelling of aspartokinase. J. Mol. Biol. **334**:459-476.

23. Marvil DK, Leisinger T. 1977. N-acetylglutamate synthase of *Escherichia coli*: purification, characterization, and molecular properties. J. Biol. Chem. **252**:3295-3303.

24. Min L, Jin Z, Caldovic L, Morizono H, Allewell NM, Tuchman M, Shi D. 2009. Mechanism of allosteric inhibition of N-acetyl-L-glutamate synthase by L-arginine. J. Biol. Chem. **284**:4873-4880.

25. Ogahara T, Ohno M, Takayama M, Igarashi K, Kobayashi H. 1995. Accumulation of glutamate by osmotically stressed *Escherichia coli* is dependent on pH. J. Bacteriol. **177**:5987-5990.

26. Pauwels K, Abadjieva A, Hilven P, Stankiewicz A, Crabeel M. 2003. The N-acetylglutamate synthase/N-acetylglutamate kinase metabolon of *Saccharomyces cerevisiae* allows co-ordinated feedback regulation of the first two steps in arginine biosynthesis. Eur. J. Biochem. **270**:1014-1024.

27. Qu Q, Morizono H, Shi D, Tuchman M, Caldovic L. 2007. A novel bifunctional N-acetylglutamate synthase-kinase from *Xanthomonas campestris* that is closely related to mammalian N-acetylglutamate synthase. BMC Biochem. **8**:4.

28. Ramón-Maiques S, Fernández-Murga ML, Gil-Ortiz F, Vagin A, Fita I, Rubio V. 2006. Structural bases of feed-back control of arginine biosynthesis, revealed by the structures of two hexameric N-acetylglutamate kinases, from *Thermotoga maritima* and *Pseudomonas aeruginosa*. J. Mol. Biol. **356:**695–713.

29. Ramón-Maiques S, Marina A, Gil-Ortiz F, Fita I, Rubio V. 2002. Structure of acetylglutamate kinase, a key enzyme for arginine biosynthesis and a prototype for the amino acid kinase enzyme family, during catalysis. Structure. **10**:329-342.

30. Reglero A, Rivas J, Mendelson J, Wallace R, Grisolia S. 1977. Deacylation and transacetylation of acetyl glutamate and acetyl ornithine in rat liver. FEBS. Lett. **81**:13-17.

31. Rubio V, Ramponi G, Grisolia S. 1981. Carbamoyl phosphate synthetase I of human liver. Purification, some properties and immunological cross-reactivity with the rat liver enzyme. Biochim. Biophys. Acta. **659**:150-160.

32. Sancho-Vaello E, Fernández-Murga ML, Rubio V. 2008. Site-directed mutagenesis studies of acetylglutamate synthase delineate the site for the arginine inhibitor. FEBS Lett. **582**:1081-1086.

33. Sancho-Vaello E, Fernández-Murga ML, Rubio V. 2009. Mechanism of arginine regulation of acetylglutamate synthase, the first enzyme of arginine synthesis. FEBS Lett. **583**:202-206.

34. Shi D, Sagar V, Jin Z, Yu X, Caldovic L, Morizono H, Allewell NM, Tuchman M. 2008. The crystal structure of N-acetyl-L-glutamate synthase from *Neisseria gonorrhoeae* provides insights into mechanisms of catalysis and regulation. J. Biol. Chem. **283**:7176-7184.

35. Slocum RD. 2005. Genes, enzymes and regulation of arginine biosynthesis in plants. Plant. Physiol. Biochem. **43**:729-745.

36. Sonoda T, Tatibana M. 1983. Purification of N-acetyl-L-glutamate synthetase from rat liver mitochondria and substrate and activator specificity of the enzyme. J. Biol. Chem. **258**:9839-9844.

37. Xu Y, Glansdorff N, Labedan B. 2006. Bioinformatic analysis of an unusual gene-enzyme relationship in the arginine biosynthetic pathway among marine gamma proteobacteria: implications concerning the formation of N-acetylated intermediates in prokaryotes. BMC Genomics. **7**:4.

Figure legends

Fig.1. Roles of N-acetyl-L-glutamate synthase (NAGS) and of its homologous enzyme N-acetyl-L-glutamate kinase (NAGK) in arginine biosynthesis. Two arrows in succession indicate the existence of two steps that are not detailed. Double and triple green arrows and green encircled plus signs denote activation, whereas double red arrows and red encircled minus signs denote inhibition. *OTC*, ornithine transcarbamylase. *ARG1*, arginase 1. **(A)** Schematic representation of the arginine biosynthetic pathway of *P. aeruginosa* and of many other bacteria and plants. The dotted arrows for the NAGS reaction indicate an anaplerotic role of NAGS in those organisms like *P. aeruginosa* in which the N-acetyl group is recycled by transacetylation from acetylornithine to glutamate (10). However, some organisms like *E. coli* deacetylate acetylornithine hydrolytically, and in these cases NAGS makes one NAG molecule per arginine molecule synthesized (10). The structures of the NAGS from *N. gonorrhoeae* (34) (PDB file 2R8V) and of NAGK from *P. aeruginosa* (28) (PDB file 2BUF) are shown next to the steps catalyzed by them to illustrate grossly their structural similarity. They are viewed along their threefold axes, with each dimer colored differently and with both subunits of each dimer in different color hues. NAGK and the AAK domain of NAGS are shown in cartoon representation. In NAGS, to avoid occluding the view of the AAK domains, the GNAT domains are shown in surface semi-transparent representation, and those in the background are fainter. **(B)** Arginine and urea biosynthesis in urea-making terrestrial animals such as humans. Animals do not make ornithine through N-acetylated intermediates and they lack NAGK and other enzymes of the route except NAGS (2). The large triple green arrow in (B) stresses the essentiality of the activation of carbamoyl phosphate synthetase (CPS I) by NAG.

Fig. 2. Engineered and mutant forms of PaNAGS. **(A)** Schematic representation of the PaNAGS polypeptide, illustrating the constructions prepared in the present work, giving the abbreviated denomination used here for each construction. The figures above the wild type form correspond to the first and last residues of each domain. **(B)** Detail of the interdomain linker (in string representation and yellow) of a subunit of NgNAGS in arginine-free form (left panel) (PDB file 2R8V) (34) and arginine-bound form (right panel) (PDB file 3D2P) (24), showing the AAK and GNAT domains that are connected by this linker in semi-transparent surface representation and colored blue and brown-orange, respectively. The side-chain of the linker residue F286 (corresponding to F285 of PaNAGS) is shown in sticks representation with its Van der Waals suface in yellow dots. **(C)** Linker mutants, shown below the alignment of the region encompassing the last and first β strands (shown as arrows) of the AAK and GNAT domains, respectively, of NgNAGS and the corresponding sequences of human NAGS (HuNAGS) and PaNAGS. Amino acid identities and conservative replacements are highlighted in black and grey backgrounds, respectively, with lettering in white. All the mutant sequences replace the EQF sequence of the wild-type form. In the *PSsite* mutant the cleavage by PreScission protease should take place immediately before the two underlined residues. **(D)** Coomassie-stained SDS-PAGE of the purified wild type enzyme, of the rAAK and rGNAT domains and of the Δ1-12 engineered form (see panel A for the composition of these forms) as well as of the gel-filtration separated AAK and GNAT domains prepared by PreScission cleavage of the *PSsite* form. St, molecular mass standard markers, with masses given at the side.

Fig. 3. Influence of linker mutations on the effect of arginine on PaNAGS activity (expressed as a percentage of that for each form in the absence of arginine). See Fig. 2C for the key to each

mutant form. **(A)** and **(B),** mutant forms having or not having a lengthened linker, respectively, are compared with the wild-type form (EQF). The activation by arginine of the EAQGP and LFQGP forms is statistically significant for the points marked with double stars ($p<0.001$) or single stars ($p<0.01$) (tested with ANOVA, followed by the Bonferroni test for individual points; n=4-8). The insets to the right of **(A)** show the glutamate concentration dependency of the velocity for the form with the EAQGP linker sequence (top inset) and for the wild-type enzyme (bottom inset). The curves drawn in these insets are those for hyperbolic kinetics with substrate inhibition for the following K_m^{Glu}, K_I^{Glu} (both in mM units) and V_{max} (as U/mg), respectively: EAQGP without arginine, 4.9, 69 and 67; EAQGP with 30 mM arginine, 37, 103 and 136; EQF (wild type) without arginine, 5.2, 72 and 136; and EQF with 2 mM arginine, 85, ∞ (no substrate inhibition) and 22.

Fig. 4. Linker cleavage with PreScission protease and size exclusion chromatography of the digested enzyme and of the recombinant AAK domain. **(A)** SDS-PAGE analysis (Coomassie staining) reveals that wild type PaNAGS (WT) is not cleaved, whereas the enzyme engineered to include in the linker the PreScission cleavage site (*PSsite*, see Fig. 2A and C) is cleaved. PaNAGS and PreScission concentrations, 0.5 mg/ml and 167 U/ml, respectively. Note that a large excess of protease was added and that the polypeptide mass of this protease is very close to that of PaNAGS. St, protein markers of the indicated masses. **(B)** Of a digestion of *PSsite* form of PaNAGS (1 mg/ml) with 30 U/ml of PreScission protease, 95 μl were centrifuged through a 0.1 ml His Spin Trap column, followed by a 0.2 ml wash with Na phosphate 20 mM pH 8/1 mM DTT/0.5 M NaCl/20 mM imidazole, and with elution with two lots of the same buffer supplemented with 0.5 M imidazole, collecting separately each eluate from the sample, the

671 washing and the two-lot elutions. The figure shows the result of SDS-PAGE analysis of the
672 various fractions. AAK and GNAT denote the bands corresponding to these individual domains.
673 **(C)** Size exclusion chromatography of cleaved (continuous line) or uncleaved (broken line)
674 PaNAGS and of the recombinant AAK domain (dotted line). For details see Materials and
675 Methods. 0.3 mg of each protein were injected. The digestion was that shown in panel (B). The
676 upper line is the semilogarithmic plot of the masses of marker proteins (closed circles) versus
677 their elution volumes. The open symbols correspond to the protein peaks below them, for the
678 following sequence-deduced masses: (○) whole enzyme, assuming it is hexameric, 294.4 kDa;
679 (□) AAK domain, either recombinat or produced by cleavage, assuming that it is hexameric,
680 189.1 kDa; (Δ, ∇) GNAT domain, assuming that it is monomeric (18.1 kDa) (Δ) or dimeric (36.2
681 kDa) (∇). The following protein standards were used (with masses given in kDa): thyroglobulin,
682 669 (not shown), ferritin, 440; *T. maritima* acetylglutamate kinase, 182 (28); *E. coli* UMP kinase,
683 165 (5); bovine serum albumin, 66.4; carbonic anhydrase, 29; ribonuclease, 13.7.
684
685 **Fig. 5.** Dependency of NAGS activity on the concentration of both substrates (**A** and **B**) and of
686 arginine (**C**) for WT, *PSsite* form (either uncleaved or PreScission protease-cleaved) and Δ1-12
687 form of PaNAGS, and of the GNAT domain either isolated after cleavage (GNAT cleaved) or
688 produced recombinantly (rGNAT). When acetyl-CoA was varied the concentration of glutamate
689 was fixed at 100 mM, except for the WT and the uncleaved *PSsite* forms, with which it was kept
690 at 30 mM. When glutamate was varied, acetyl-CoA was fixed at 4 mM.To allow meaningful
691 comparison of the activity of various enzyme forms having different masses, velocities are given
692 as turnover numbers per polypeptide chain (units are s^{-1}). Similar concentration-dependency of
693 the *PSsite* cleaved form, the Δ1-12 form and the isolated GNAT domains is evident for each

substrate. Therefore, single curves were fitted for the results with all these forms for glutamate **(A)** and for acetyl-CoA **(B)**. The curve for glutamate corresponds to hyperbolic kinetics with substrate inhibition and respective values for K_m^{Glu}, K_I^{Glu} and apparent k_{cat} at infinite glutamate, of 240 ± 45 mM, 1254 ± 480 mM and 66 ± 8 s^{-1}. The curve for acetyl-CoA **(B)** is a hyperbola with values of $K_m^{Acetyl\text{-}CoA}$ and apparent k_{cat} at infinite acetyl-CoA of 190 ± 30 μM and 20.0 ± 0.7 s^{-1}, respectively. The K_m^{Glu}, K_I^{Glu} and $K_m^{Acetyl\text{-}CoA}$ for the WT and for the uncleaved *PSsite* forms of the enzyme are those of Table 1, and the apparent k_{cat} values (in s^{-1}) are, for [glutamate] = ∞, 111 ± 16 for WT and 69 ± 7 for the uncleaved *PSsite* form (see inset in **A**); and for [acetyl-CoA] = ∞, 65 ± 1 for WT and 46 ± 1 for the *PSsite* form (see **B**). **(C)** Influence of arginine concentration on enzyme activity. Results are expressed as a percentage of the activity of the same enzyme form in the absence of arginine. A single line corresponding to no inhibition has been fitted to the results for all forms except the wild type enzyme. Substrate concentrations in these assays were 4 mM acetyl-CoA and either 30 mM glutamate for WT and uncleaved *PSsite* forms or 100 mM for all other forms.

Fig. 6. Size exclusion chromatography of the Δ1-12 form of PaNAGS. The chromatographic profile (continuous line) is compared with that of the wild type enzyme (broken line), both injected in 0.2 mg amounts. The procedure and system are descibed in the section on "cleavage with PreScission protease and separation of the two domains" in the Materials and Methods. In the case of Δ1-12, enzyme activity was measured in 1-ml collected fractions (closed circles, bottom plot). The upper line is the semilogarithmic plot of the masses of marker proteins (closed circles) versus their elution volumes.. The open symbols correspond to the protein peaks below them, for the following sequence-deduced masses: (○) whole enzyme, assuming it is hexameric

(294.4 kDa); (◇) Δ1-12 enzyme form, assuming that it is dimeric (95.5 kDa). The protein standards used and their masses in kDa are: ferritin, 440; β-amylase, 224; *T. maritima* acetylglutamate kinase, 182 (28); aldolase, 158; alcohol dehydrogenase, 147; bovine serum albumin, 66.4; carbonic anhydrase, 29; ribonuclease, 13.7.

Fig. 7. The AAK domain of NAGS is an ancestral NAGK. **(A)** Stereo view of the superimposition of the structure of the active center of *E. coli* NAGK (in green) bound to NAG and the ATP inert analog AMPPNP (PDB file 1GS5) (29) with the corresponding region of NgNAGS (in orange-brown) (PDB 2R8V). The main chain of the structural elements (labeled) that make the site are shown. A few important residues are illustrated in the same color as the main chain (except the N atoms of the side chains, that are colored blue) and are labeled. NAG and AMPPNP are shown in sticks representation, with C, P, O and N atoms colored yellow, green, red and blue, respectively. **(B)** and **(C)**, dependency of the activity of the M26K/L240K double mutant (squares) of rAAK on ATP and NAG concentrations. The circle illustrates the lack of activity of the wild type enzyme (0.3 mg/ml) assayed under the same conditions at 20 mM ATP and 100 mM NAG concentrations. The hyperbola fitted to the points for variable ATP yields K_m^{ATP} and $V^{[ATP] = \infty}$ values of 1.6 ± 0.2 mM and 0.122 ± 0.003 U/mg. The nearly linear NAG concentration-dependency of the activity indicates that the K_m^{NAG} and the $V^{[NAG] = \infty}$ exceed, respectively, 0.225 M and 0.5 U/mg. The curve fitted over the points would correspond to the hyperbola for K_m^{NAG} and $V^{[NAG] = \infty}$ values of 0.83 M and 1.1 U/mg, respectively.