THE CRYSTAL STRUCTURE OF N-ACETYL-L-GLUTAMATE SYNTHASE FROM NEISSERIA GONORRHOEAE PROVIDES INSIGHTS INTO MECHANISMS OF CATALYSIS AND REGULATION.*

Dashuang Shi†‡§, Vatsala Sagar‡, Zhongmin Jin‖, Xiaolin Yu†, Ljubica Caldovic†, Hiroki Morizono†, Norma M. Allewell‡ and Mendel Tuchman†

From the †‡§ Children’s Research Institute, Children’s National Medical Center, The George Washington University, 111 Michigan Avenue, Washington, DC 20010, USA, ‡Department of Chemistry and Biochemistry, College of Chemical and Life Sciences, 2300 Symons Hall, University of Maryland, College Park, MD 20742, USA, and ‖SER-CAT, APS, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, USA

Running Title: Structure of N-acetylglutamate synthase

Address correspondence to: Dashuang Shi, Children’s Research Institute, Children’s National Medical Center, 111 Michigan Avenue, N.W., Washington, D.C. 20010-2970, Tel. 202-476-5817; Fax. 202-476-6014; E-Mail: dshi@cnmcresearch.org

The crystal structures of N-acetylglutamate synthase (NAGS) in the arginine biosynthetic pathway of Neisseria gonorrhoeae complexed with acetyl coenzyme A (AcCoA), and with coenzyme A (CoA) plus N-acetylglutamate (NAG), have been determined at 2.5 and 2.6 Å resolution, respectively. The monomer consists of two separately folded domains, an amino acid kinase (AAK) domain and an N-acetyltransferase (NAT) domain connected through a 10 Å linker. The monomers assemble into a hexameric ring that consists of a trimer of dimers with 32 point symmetry, inner and outer ring diameters of 20 Å and 100 Å, respectively, and a height of 110 Å. Each AAK domain interacts with the cognate domains of two adjacent monomers across two two-fold symmetry axes and with the NAT domain from a second monomer of the adjacent dimer in the ring. The catalytic sites are located within the NAT domains. Three active site residues, Arg316, Arg425 and Ser427, anchor NAG in a position at the active site to form hydrogen bond interactions to the main chain N atoms of Cys356 and Leu314, and hydrophobic interactions to the side chains of Leu313 and Leu314. The mode of binding of AcCoA and CoA is similar to other NAT family proteins. The AAK domain while catalytically inactive, appears to bind arginine. This is the first reported crystal structure of any NAGS and it provides insights into the catalytic function and arginine regulation of NAGS enzymes.

The first committed step of arginine biosynthesis in most microorganisms and plants is the acetylation of L-glutamate by N-acetylglutamate synthase (NAGS, EC 2.3.1.1) to produce N-acetylglutamate (NAG). NAG is then converted by NAG kinase (NAGK) to NAG-phosphate, which is subsequently converted to N-acetylornithine. In fungi, plants and many bacteria, NAG can also be generated via transfer of the acetyl group from N-acetylornithine to glutamate by ornithine acetyltransferase (OAT, EC 2.3.1.1) to form an acetyl cycle (1-2). In these organisms, NAGS seems to play only an anaplerotic role, since most NAG is produced by OAT. In organisms that possess OAT, feedback inhibition by arginine primarily targets NAGK, although NAGS is also inhibited to some extent. In organisms without OAT, arginine inhibits NAGS but not NAGK (3). N. gonorrhoeae, which possesses both NAGS and OAT, can be classified as a member of the former group. Moreover, the OAT in N. gonorrhoeae appears to be bifunctional, and able to use both N-acetylornithine and AcCoA as acetylation donors (4), even though there is limited sequence similarity between NAGS and bifunctional OAT genes (3).

In mammals, NAGS produces NAG which is an obligatory allosteric activator of carbamylphosphate synthetase I in the urea cycle (5). Interestingly, and in contrast to NAGS of microorganisms that are inhibited by arginine, this amino acid activates mammalian NAGS. Deficiency of NAGS and consequentially, of
NAG, causes hyperammonemia due to a secondary deficiency of carbamylphosphate synthetase I, resulting in blocked ureagenesis (6).

In contrast to other enzymes of the arginine pathway and the urea cycle which show significant sequence similarity across phyla, the similarity between mammalian, fungal and bacterial NAGS is relatively low (7-9). However, it is believed that most present day NAGS genes were derived from the fusion of two ancestral genes encoding NAGK and N-acetyltransferase (NAT). The discovery of a bifunctional enzyme capable of catalyzing the first two reactions of arginine biosynthesis (9) and the existence of a “short” version of NAGS in some bacteria (10) lend support to this hypothesis.

Several NAGK- and many NAT-related structures have been determined (11-14). However, no NAGS structure from any organism has been previously reported, probably because most NAGS proteins are unstable and not amendable to crystallization (15). We reported herein the first three-dimensional structures of NAGS, and show that they consist of two linked but independent folding domains similar to those of AAK and GCN5-related NAT, respectively. Two structures have been determined, one with AcCoA bound, and a second with NAG and CoA bound. These structures provide insights into the mechanism of catalysis and the allosteric effect of arginine.

**EXPERIMENTAL PROCEDURE**

**Enzyme production and crystallization** - The argA gene was PCR amplified from the *N. gonorrhoeae* genomic DNA (ATCC 53420) using the primers CCGGCAATATGAAACGCAGCCGACAGCTTTGT and CTTCTCGAAGATATCAGCGGTGCGGCGGAC, and cloned into a pET28a expression vector (Novagen). The inserted gene was sequenced and was found to have three point variations that lead to V312I, D336N and P427S amino acid substitutions. All three point variations are present in the wild type NAT domain sequence of *N. meningitidis* NAGS (accession codes: NP_274872, AAF42210 or CAM07850). Thus, the resulting construct agrees with the published genomic NAGS sequences of the AAK domain from *N. gonorrhoeae* and the NAT domain from *N. meningitidis*.

The protein was expressed in *E. coli* BL21(DE3) cells (Invitrogen) and purified in a two-step procedure using Ni-afinity and DEAE columns (GE Healthcare). The protein was concentrated to ~10 mg/ml and incubated for 30 min. in the presence of ~23 mM AcCoA and with or without ~100 L-glutamate after which CsCl was added to 100 mM. The hanging drop method was used for crystallization screening. The best crystals were obtained by mixing 2 µl of the protein mixture with 2 µl of a well solution containing 100 mM CsCl, 6% PEG 3350 and 100 mM sodium citrate at pH 5.8. The selenomethionine (Se-Met) substituted protein was prepared using Overnight Auto-induction System 2 (Novagen) as described previously (15-16). The Se-Met derivative was purified by the same procedure as the wild-type protein. On average, 84% of methionines were substituted by Se-Met as determined by MALDI-TOF/TOF mass spectrometry (15-16). The Se-Met derivative crystals were grown under conditions identical to the wild-type crystals. The fluorescence scan of Se-Met derivative crystals showed a strong signal at the selenium absorption edge. The crystals grown in the presence of AcCoA and L-glutamate were found to bind the reaction products CoA and NAG.

**Data collection and processing** – Multiple-wavelength anomalous dispersion (MAD) and diffraction data from the AcCoA bound crystal were collected from a single crystal at 100 K on beamline SER-CAT 22-ID equipped with MAR300 CCD at the Advanced Photon Source (APS) at Argonne National Laboratory. The data for CoA and NAG bound crystal were collected on a Rigaku rotating anode generator equipped with R-axis IV imaging plate at the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health. The crystals were cryoprotected by supplementing the well solution with 30% ethylene glycerol. MAD datasets at three wavelength were collected at the selenium absorption edge, the inflection point, and a high energy remote position. The data were indexed, integrated and scaled with the software package HKL2000 (17) and reduced using the program TRUNCATE in the CCP4 suite (18). The crystals belonged to the trigonal space group P312 with unit cell parameters: *a* = *b* = 98.7, *c* = 89.8 Å.
Structure determination and refinement - The structure was solved by MAD phasing using the SHELX package (20-21). Four of the six possible selenium positions were identified using the program SHELXD. The model was built using the program RESOLVE (22). Approximately 70% of the sequence can be auto-traced from the electron density map. The program COOT was used for viewing electron density maps and building the initial model (23). Bound AcCoA was clearly visible as a well-defined electron density and was built into the model. The initial structure was subsequently refined against the native dataset using the program CNS package (24). The model was subjected to 1000 steps of simulated annealing refinement with a starting temperature of 2500 K, followed by several rounds of model rebuilding with program O (25) or COOT (23). The process of refinement was monitored by the free R factor of 5% of the data (26). In the later stages of refinement, translation/libration/screw (TLS) parameters (27) were included and refined using REFMAC5 (28). The TLS groups were selected using the TLSMD server (29). Statistics for the final refined model are given in Table II. Refinements for the CoA and NAG complexed dataset, and a rescaled MAD remote dataset were carried out in a similar procedure. The three refined structures are essentially identical with a rmsd of less than 0.40 Å for the superimposition of all equivalent Cα atoms. The bound AcCoA, CoA and NAG, as well as the conformation of the surrounding active site residues are all well defined (Fig.1, Fig. 2A).

Enzymatic assay – NAGS and NAGK activities were assayed using a modification of the method described by Qu et al. (9). Synthase activity was measured in 50 mM Tris-HCl buffer, pH 8.5, containing 100 mM NaCl, 10 mM L-glutamate and 2.5 mM AcCoA and 0.16 μg of enzyme in a total reaction volume of 100 μl. The assay was performed at 30 °C for 5 min and quenched with 100 μl of 30% trichloroacetic acid (TCA). The product, NAG, was quantified using liquid chromatography-mass spectroscopy as described previously (7-8). NAGK activity was measured using a colorimetric assay (30) with 1.0 μg of enzyme in 100 μl of 100 mM Tris-HCl, pH 8.5, 100 mM NaCl, 100 mM NAG, 20 mM ATP, 40 mM MgCl₂ and 400 mM hydroxylamine at room temperature for 20 min. The reaction was quenched and color was developed by adding 100 μl of an aqueous solution containing 5% FeCl₃, 8% TCA and 0.3 M HCl. The absorbance of the colored product was measured at 540 nm. Inhibition of NAGS activity was measured over a range (0.01-10 mM) of L-arginine concentrations.

Analytical gel filtration chromatography – A superdex 200HR (10/30) column mounted on an AKTA fast protein liquid chromatography system (GE Health) was used for the analytical gel filtration experiment, which was carried out at 298 K at a flow rate of 1.0 ml/min. A solution, consisting of 50 mM K₂HPO₄- KH₂PO₄, 150 mM KCl, 20 % glycerol (v/v) pH 7.5, was used as the mobile phase. Molecular weight markers, consisting of thyroglobulin, ferritin, catalase, bovine serum albumin, ovalbumin and myoglobin, were used for column calibration. Blue dextran 2000 and vitamin B₁₂ were used as the void and included markers. The UV detector monitored the effluent at 280 nm, which revealed the elution of the NAGS enzyme at a calculated molecular weight of 336 KDa corresponding to the predicted molecular weight of a hexamer (Fig. 3).

Modeling of the ternary complex and arginine binding site – The ternary glutamate/AcCoA/NAGS complex model was built using the coordinates of AcCoA from the AcCoA bound structure and NAG without the acetyl group from the CoA and NAG bound structure.

The figures were drawn using the programs Situs 2.3 (31), VMD (32), Pymol (33) and O (25).

RESULTS AND DISCUSSION

Enzymatic activity – The purified enzyme displayed specific NAGS activity of 29.1 μmol/min/mg protein under the conditions used. This level of activity is approximately 10 fold lower than the NAGS activity of the bifunctional
NAGS/K enzyme from *Xanthomonas campestris* under the same assay conditions and 4 fold lower than that of NAGS from *E. coli* (34).

Even though *N. gonorrhoeae* uses the cyclic acetyl pathway to produce arginine, its NAGS activity is inhibited by arginine with 50% inhibition in the presence of 0.15 mM of arginine.

The purified enzyme has no detectable NAGK activity. This is consistent with the presence of a separate gene encoding NAGK (accession No: YP_207961) in the *N. gonorrhoeae* genome.

**Structure of NAGS** - The final structural models of *N. gonorrhoeae* NAGS consist of 423 amino acids, 1 AcCoA and 117 water molecules for the AcCoA bound binary complex, and 423 amino acids, 1 CoA, 1 NAG and 127 water molecules for the CoA and NAG bound ternary complex. The *N*-terminal his<sub>6</sub> tag with a thrombin recognition site (20 amino acid residues), the first four amino residues of the protein, and residues 114-122 were not modeled because their electron density was ambiguous. All the main chain, side chain and planar group parameters assessed with PROCHECK (35) are comparable to those in structures at similar resolution in the PDB. The main chain conformations of 88.6% of the residues are within the most favorable regions in the Ramachandran plot for the AcCoA bound structure. Only residue Glu145 is in an energetically unfavorable conformation, which is maintained by an extensive hydrogen bonding network.

The AAK and NAT domains of the monomer are connected by a 10 Å linker – The monomer of *N. gonorrhoeae* NAGS consists of two distinct domains. A three-amino acid (10 Å) linker (Glu284-Ala285-Phe286) connects the N-terminal AAK domain (Asp5-Lys283) and the C-terminal GCN5-related NAT domain (Val287-Arg436) (Fig. 4A). These two domains have no other contacts within a single monomer. The N-terminal domain has a typical AAK fold with a central eight-stranded parallel β sheet sandwiched between α helices (Fig. 4B, and supplementary materials Fig. S1). Other protein structures with a similar fold include NAGK (11-12), carbamate kinase (36), carbamate kinase-like CPS (37), UMP kinase (38-39), aspartokinase (40) and 5-glutamate kinase (41). Superimposition of this domain with the closely related *Thermotoga maritima* and *Pseudomonas aeruginosa* NAGK structures yields a rmsd of 1.7 and 1.6 Å for the 238 and 234 equivalent Cα atoms, respectively. The amino acid sequence of the AAK domain of *N. gonorrhoeae* NAGS has 25.9% and 27.4% sequence identity to *T. maritima* and *P. aeruginosa* NAGK, respectively. The major structural difference among the different AAK family members resides in their N-terminal lobe which is involved in binding different substrates and formation of different dimer interfaces. On the other hand, the structures of the C-terminal lobe, which is involved in binding the common substrate, ATP, are similar (38, 41). The structural differences among the NAGKs from different organisms consist mainly of variations in the conformation of the β3-β4 loop which results in differences in the degree of closure of the active site. In the AAK domain of *N. gonorrhoeae* NAGS, the region corresponding to the active site of NAGK proteins is closed (see supplementary materials Fig. S2), as in unliganded *Mycobacterium tuberculosis* NAGK (PDB accession code: 2AP9).

The C-terminal domain of the monomer has a typical GCN5-related NAT fold, with central anti-parallel β sheets flanked by α helices (13). The central β sheet is divided into two arms that form a “V” shape, with AcCoA binding between the two arms (Fig. 4C). The N-terminal arm of this domain consists of a four-stranded anti-parallel β sheet flanked by three α helices. The C-terminal arm is composed of a three-stranded anti-parallel β sheet flanked by two α helices. The two arms are joined by four hydrogen bonds between two parallel strands (β22 and β23) and are divided by a bulge formed by residues Ala355 and Cys356. DALI database searching indicates that the structure of the NAT domain of *N. gonorrhoeae* NAGS is most closely related to GCN5 histone acetyltransferase (PDB code; 1M1D) with a z score of 16.4 (42). Superimposition of the two structures results in a rmsd of 1.7 Å for 121 equivalent Cα atoms.

*NAGS* monomers assemble into a hexameric ring with equatorial AAK and polar NAT domains – Crystals of *N. gonorrhoeae* NAGS (space group P312) have only one monomer in an asymmetric unit. The three-fold crystallographic symmetry...
axis coincides with the molecular three-fold axis, while the two-fold crystallographic symmetry axis aligns with the molecular two-fold axis, so that the hexamer has exactly 32 point group symmetry. As a result, the hexamer can be considered to be composed from either two trimers or three dimers.

Within each trimer, the AAK domain of one subunit interacts with the NAT domain of an adjacent subunit. The stacked trimers in a hexamer form a ring with outer and inner diameter of approximately 100 Å and 20 Å, respectively, and a height of 110 Å (Fig. 5A). The diameter of the inner ring is smaller than in reported arginine sensitive NAGK structures, which have inner ring diameters of 26-32 Å, depending on whether arginine is bound (12).

The three dimers and their inter-subunit interactions are best viewed perpendicularly to the three-fold axis. From this view, the AAK domains form a zigzag equatorial belt while the NAT domains are polar (Fig. 5A). Each AAK domain interacts with the AAK domain of two other subunits.

In the crystal, the molecular hexamers are packed together with their three-fold axes along the c axis so that they form an extended tube across the crystal with alternating AAK and NAT layers.

**Monomer-monomer interactions across dimer interfaces** – Two types of dimer interfaces occur in a hexamer (Fig. 5B-C). One is similar to the dimer interface of the homodimeric *E. coli* NAGK structure and is exemplified by the interaction between K1 and K4 (Fig. 5B). This interface consists exclusively of residues from strand 5, helix 4, helix 5 and the loop between strand 11 and 12 in the N-terminal lobes of both subunits. The interface is quite flat with strand 5 in the middle and helices on both sides. Strand 5 interacts with the corresponding strand from the adjacent subunit through hydrogen bonds, in a pattern similar to antiparallel β strand interactions. The main chain atoms of Leu125, Ser127 and Asn129 are involved in this interaction forming a 16-stranded β sheet that spans the whole dimer. Interactions between the corresponding α helices from each subunit are mainly hydrophilic. The side chains of the polar residues of Gln96, Arg103, Ser104, Glu107, Asn129, Gln161 and His175, and the hydrophobic residue of Val126, Phe160 and Ala164 are involved in this dimer interface. Extensive interactions that include central secondary structure elements with many hydrogen bonding and hydrophobic anchor points across this dimer interface suggest that this dimer is stable and likely to exist in solution.

The second dimer interface is similar to that of the arginine-sensitive NAGK structure (12). It is formed by interlacing N-helices from adjacent monomers (e.g. K1 and K5, Fig. 5C). The long N-terminal extension which forms this helix is present in arginine-sensitive NAGKs and NAGSs, but not arginine-insensitive NAGKs. This helix is divided into two segments by a bend at a highly conserved proline (Pro16). The N-terminal 15 residue segment is amphipathic, with hydrophobic residues on one side and polar residues on the other. The C-terminal segment, which is generally 5 residues long, is likely to be directly involved in binding arginine at a site similar to that of arginine-sensitive NAGKs (12). The conserved Tyr16, corresponding to Tyr15 and Tyr21 in *T. maritima* and *P. aeruginosa* NAGK, respectively, is found in this segment. Residues Phe7, Phe11, Ala14, Ile18 and His10 interlace the two adjacent N-terminal helices, linking two adjacent monomers across a two-fold axis. In addition, the polar side chains of Arg12, Arg19 and Arg22 on the other side of the helix hydrogen bond to Asp46, Gln52, Ser122 and Asp164 from the adjacent monomer, reinforcing the interactions between monomers. This N-terminal helix appears to be important for the formation of the hexamer and arginine allosteric regulation. Since the hexamer appears to be essential for an arginine response, it is likely to represent the native structure in solution (12), which has been confirmed by our gel-filtration chromatography study (Fig. 3).

The buried surface area of the K1-K4 and K1-K5 interfaces are 1206 Å² and 1340 Å², in size respectively, calculated with a probe radius of 1.4 Å using program AREAIMOL at CCP4 suite (18). These values are similar to those for the arginine-sensitive NAGK structures (12). The extensive interactions at both dimeric interfaces indicate that the hexameric structure is likely to be stable.

**AAK and NAT domain interactions around the three-fold axis** – There are no interactions between
the AAK and NAT domain within the same monomer. However, the NAT domain of each monomer is positioned to interact with the AAK domain of a monomer from an adjacent dimer (Fig. 5D). Interactions at this interface are not as extensive as interfaces within a dimer and mainly involve polar residues. The side chains of Arg134, Asp140, Asp143, Arg151, Lys152, and the main chain nitrogen and/or oxygen atoms of Ile136, Val138 and Gly141 from the AAK domain interact with Asn393, Thr394, Asn426 and His428 from the NAT domain. Interestingly, the bound AcCoA also takes part in this interface interaction, through atoms O2*, O3* and OAQ which interact with the side chains of Arg134, Arg151 and Lys152 from the AAK domain of the adjacent subunit (Fig. 5D, Table II).

ATP- and NAG-like binding sites in the AAK domain – Although there is significant sequence similarity between NAGK proteins and the AAK domain of N. gonorrhoeae NAGS (20-27% sequence identity), the latter does not display detectable NAGK activity. Structural alignment of N. gonorrhoeae NAGS with known structures of NAGK clearly indicates that the corresponding NAGK-like site in N. gonorrhoeae NAGS would not be able to bind ATP or NAG and catalyze the phosphorylation of NAG, since it lacks several key residues. For example, Lys8 (E. coli NAGK numbering) which is conserved in all NAGK sequences including the bifunctional X. campestris NAGS/K is replaced by Gly29 in N. gonorrhoeae NAGS and the space occupied by the Lys side chain in NAGK is occupied by the side chain of Met186 in N. gonorrhoeae NAGS. Similarly, the conserved Asn160-Ala161-Asp162 motif (E. coli NAGK numbering) is replaced by Asp185-Met186-Val187 and the essential Lys217 is substituted by Leu241. These differences would be sufficient to render inactive the NAGK domain of N. gonorrhoeae NAGS (9).

Active site and proposed catalytic mechanism – The well-defined electron densities of bound AcCoA, CoA and NAG clearly identify the catalytic site within the NAT domain (Fig 1, Fig. 2A). AcCoA and CoA bind to the enzyme in a conformation similar to that of other GCN-5-related NAT proteins, using the loop joining strand 22 and helix 12 to coordinate the pyrophosphate group (Table II). The sequence Gln364-Glu365-Gly366-Gly367-Tyr368-Gly369 in this loop conforms to the {Arg/Gln/-X-X-Gly-X-{Gly/Ala} motif for AcCoA recognition and binding in the GCN5-related NAT superfamily members (43). However, unique to the present structure, the side chains of Arg134, Arg151, and Lys152 from the adjacent AAK domain are also involved in the interactions with AcCoA (Fig. 5D). The adenosine moiety of AcCoA is located on the surface and its conformation is fixed to the side chain of Glu397 via strong hydrogen bonds. AcCoA bends approximately 180° at the pyrophosphate group to render the CoA chain parallel to the adenosine ring and bends further 90° at the CBO atom of pantetheine, so that its acetyl group extends deeply into the protein core. The acetyl group is completely buried within a hydrophobic pocket created by the side chains of Ala355, Leu357 and Phe399. The acetyl carbonyl atom hydrogen bonds to the main chain N atoms of Cys356 and Leu357, orienting the acetyl group towards the cavity created by the β-bulge of Ala355-Cys356 (Fig. 1). Similar hydrogen bonding interactions between the acetyl carbonyl group and the NH group of residues corresponding to Leu357 have been reported in other GCN5-related NAT proteins (44). However, a highly conserved tyrosine, which interacts with the sulfur or carbonyl atom of AcCoA in other NAT protein structures (14, 45), is not found in N. gonorrhoeae NAGS. Ser392 is nearby; but its side chain oxygen atom is too far (3.7 Å) to interact strongly with the sulfur atom of AcCoA. Instead, it forms a hydrogen bond with this sulfur atom in CoA (3.3 Å). In addition, the conformation of the side chain of Ser392 appears to be fixed by hydrogen bonding to the main chain N atom of Thr395. The pantetheine moiety of AcCoA and CoA interacts with the protein through hydrogen bonds to Val359 (NH) and Leu357 (C=O), in a manner similar to an anti-parallel β-sheet, while its carbonyl oxygen atoms form hydrogen bonds with the side chains of Thr395 and Gln364.

Since the acetyl group of AcCoA is completely buried within the protein’s core, the potential acceptor, L-glutamate, must enter the protein through the opposite side of the protein by a way of a channel that provides the acetyl group of AcCoA with access to the solvent. The CoA
and NAG bound ternary structure confirms that L-glutamate or NAG binds to the active site using the polar side chains of Arg316, Arg425 and Ser427, and the main chain N atoms of Cys356 and Leu314 to anchor its two carboxyl groups, and the hydrophobic side chains of Leu314 and Leu391 to interact its hydrophobic stem. Similar to the acetyl group in AcCoA, the acetyl carbonyl atom of NAG hydrogen bonds to the main chain N atoms of Cys356 and Leu357. The side chain of Arg416, whose position is fixed by hydrogen bonding with the side chain of conserved Glu353, may also be important in providing a positive charge environment for L-glutamate or NAG binding.

Even though no significant general conformational differences appear to exist between the binary complexes with AcCoA, CoA, and the ternary complex with both NAG and CoA bound, the side chain of Arg316 moves significantly. In the ternary complex, the side chain of Arg316 shifts ~8 Å, so that it swings towards the active site (Fig 2A). This side chain movement appears to be important for the catalytic process.

Two different catalytic mechanisms have been proposed for the NAT enzymes. One consists of a two-step ping-pong process in which the acetyl group of AcCoA is transferred to a Cys group of the protein before being transferred to the substrate (46-48). In the present structure, however, the distance of the only potential acceptor (Cys356) is 8.4 Å away from the sulfur atom of AcCoA. It is therefore unlikely to form a covalently linked intermediate. Hence, it is more likely that N. gonorrhoeae NAGS uses the alternative proposed mechanism which is a one-step catalysis, as do most members of the GCN5-related NAT family (13). The binary and ternary N. gonorrhoeae NAGS structures that have been determined enable a model of the enzyme-catalyzed reaction between AcCoA and L-glutamate to be developed (Fig. 2B). In this model, the α-amino nitrogen atom of L-glutamate is 2.5 Å from the carbon atom of the acetyl group, poised to attack the acetyl group directly and to form an S-configured tetrahedral intermediate. According to this model, the carbonyl bond is polarized during the reaction by two strong hydrogen bonding interactions with main chain N atoms of Cys356 and Leu357. When the products (CoA and NAG) form, the sulfur atom of CoA moves ~0.9 Å towards the side chain of Ser392, allowing it to be protonated prior to release. The plane of the acetyl group of NAG rotates almost 90° from the acetyl group of AcCoA. At the same time, the side chain of Arg316 moves away to allow NAG to dissociate from the protein. The main chain carbonyl O atom of Leu391, which hydrogen bonds to amino N atom of NAG and close to the S atom of CoA, may help to transfer the proton from the amino N atom to the S atom of CoA.

Relationship to other NAGS proteins – Four different classes of NAGS are known: “classical” E. coli-like NAGS; bifunctional vertebrate-like NAGS-K; “short” NAGS and argH-argA fusion NAGS (49). Both “classical” and vertebrate-like NAGS sequences consist of an AAK domain and NAT domain with about 440 amino acid residues (8). The present N. gonorrhoeae structure clearly indicates that the NAT domain is essential for catalysis while the AAK domain has probably a regulatory function. This assertion is consistent with the observations in various NAGS enzymes which indicate that the NAT domain is essential for catalysis, while the AAK domain may be missing, or if present may or may not have NAGK activity, or may even be part of an argH encoded arginosuccinate lyase (49). Based on the above results, the AAK domain in N. gonorrhoeae NAGS does not participate in the glutamate binding contrary to previous hypotheses (12, 50). This observation is reminiscent of aspartate transcarbamylase in which the essential catalytic trimer can function alone, complexed with active or inactive dihydroorotase, or regulatory units or even fused to other proteins such as carbamylphosphate synthetase and dihydroorotase (51).

Acknowledgments - We thank Dr. David Davies for facilitating use of the diffraction equipment in the Molecular Structure Section at the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institute of Health and Dr. Fred Dyda for help in data collection. Data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting
institutions may be found at www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.
REFERENCES

1. Cunin, R., Glansdorff, N., Pierard, A., and Stalon, V. (1986) Microbiol. Rev. 50, 314-352.
2. Slocum, R. D. (2005) Plant Physiol. Biochem. 43, 729-45.
3. Caldivic, L., and Tuchman, M. (2003) Biochem. J. 372, 279-290.
4. Martin, P. R., and Mulks, M. H. (1992) J. Bacteriol. 174, 2694-2701.
5. Grisolia, S., and Cohen, P. P. (1953) J. Biol. Chem. 204, 753-757.
6. Bachmann, C., Colombo, J., P. and Jaggi, K. (1982) Adv. Exp. Med. Biol. 153, 39-45.
7. Caldivic, L., Morizono, H., Gracia Panglao, M., Gallegos, R., Yu, X., Shi, D., Malamy, M. H., Allewell, N. M., and Tuchman, M. (2002) Biochem. Biophys. Res. Commun. 299, 581-586.
8. Caldivic, L., Morizono, H., Yu, X., Thompson, M., Shi, D., Gallegos, R., Allewell, N. M., Malamy, M. H., and Tuchman M. (2002) Biochem. J. 364, 825-831.
9. Qu, Q., Morizono, H., Shi, D., Tuchman, M., and Caldivic, L. (2007) BMC Biochem. 8, 4.
10. Errey, J. C., and Blanchard, J. S. (2005) J. Bacteriol. 187, 3039-3044.
11. Ramon-Maiques, S., Marina, A., Gil-Ortiz, F., Fita, I., and Rubio, V. (2002) Structure 10, 329-342.
12. Ramon-Maiques, S., Fernandez-Murga, M. L., Gil-Ortiz, F., Vagin, A., Fita, I., and Rubio, V. (2006) J. Mol. Biol. 369, 695-713.
13. Dyda, F., Klein, D. C., and Hickman, A. B. (2000) Annu. Rev. Biophys. Biomol. Struct. 29, 81-103.
14. He, H., Ding, Y., Bartlam, M., Sun, F., Le, Y., Qin, X., Tang, H., Zhang, R., Joachimiak, A., Liu, J., Zhao, N., and Rao, Z. (2003) J. Mol. Biol. 325, 1019-1030.
15. Shi, D., Caldivic, L., Jin, Z., Yu, X., Qu, Q., Roth, L., Morizono, H., Hathout, Y., Allewell, N. M., and Tuchman, M. (2006) Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 62, 1218-1222.
16. Shi, D., Yu, X., Roth, L., Morizono, H., Hathout, Y., Allewell, N. M., and Tuchman, M. (2005) Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 61, 676-679.
17. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307-326.
18. Collaborative Computational Project, Number 4 (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 760-763
19. Matthews, B. W. (1968) J. Mol. Biol. 33, 491-497.
20. Uson, I., and Sheldrick, G. M. (1999) Curr. Opin. Struct. Biol. 9, 643-648.
21. Schneider, T. R., and Sheldrick, G. M. (2002) Acta Crystallogr. Sect. D Biol. Crystallogr. 58, 1772-1779.
22. Terwilliger, T. C. (2003) Acta Crystallogr. Sect. D Biol. Crystallogr. 59, 38-44.
23. Emsley, P., and Cowtan, K. (2004) Acta Crystallogr. Sect. D Biol. Crystallogr. 60, 2126-2132.
24. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Crosse-Kunsteve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J. Rice, L. Simonson, M. T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905-921.
25. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110-119.
26. Brünger, A. T. (1992) Nature (London) 355, 472-475.
27. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) Acta Crystallogr. Sect. D Biol. Crystallogr. 57, 122-133.
28. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. Sect. D Biol. Crystallogr. 55, 240-255.
29. Painter, J., and Merritt, E. A. (2006) Acta Crystallogr. Sect. D Biol. Crystallogr. 62, 439-450.
29. Haas, D., and Leisinger, T. (1975) Eur. J. Biochem. 52, 377-393.
31. Chácon, P., and Wriggers, W. (2002) J. Mol. Biol. 317, 375-384.
32. Humphrey, W., Dalke, A. and Schulten, K. (1996) J. Mol. Graphics 14, 33-38.
33. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, http://www.pymol.org.
34. Marvil, D. K., and Leisinger, T. (1977) J. Biol. Chem. 252, 3295-3303.
35. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr.
26, 283–291.
36. Marina, A., Alzari, P. M., Bravo, J., Uriarte, M., Barcelona, B., Fita, I., and Rubio, V. (1999) *Protein Sci.* **8**, 934-940.
37. Ramon-Maiques, S., Marina, A., Uriarte, M., Fita, I., and Rubio, V. (2000) *J. Mol. Biol.* **299**, 463-476.
38. Marco-Marin, C., Gil-Ortiz, F., and Rubio, V. (2005) *J. Mol. Biol.* **352**, 438-454.
39. Briozzo, P., Evrin, C., Meyer, P., Assairi, L., Joly, N., Barzu, O., and Gilles, A. M. (2005) *J. Biol. Chem.* **280**, 25533-25540.
40. Kotaka, M., Ren, J., Lockyer, M., Hawkins, A. R., and Stammers, D. K. (2006) *J. Biol. Chem.* **281**, 31544-31552.
41. Marco-Marin, C., Gil-Ortiz, F., Perez-Arellano, I., Cervera, J., Fita, I., and Rubio, V. (2007) *J. Mol. Biol.* **367**, 1431-1446.
42. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233**, 123-138.
43. Neuwald, A. F., and Landsman, D. (1997) *Trends Biochem. Sci.* **22**, 154-155.
44. Marmorstein, R., and Roth, S. Y. (2001) *Curr. Opin. Genet. Dev.* **11**, 155-161.
45. Vetting, M. W, Roderick, S. L., Yu, M., and Blanchard, J. S. (2003) *Protein Sci.* **12**, 1954-1959.
46. Yan, Y., Harper, S., Speicher, D. W., and Marmorstein, R. (2002) *Nat. Struct. Biol.* **9**, 862-869.
47. Wong, L. J., and Wong, S. S. (1983) *Biochemistry* **22**, 4637-4641.
48. Sinclair, J. C., Sandy, J., Delgoda, R., Sim, E., and Noble, M. E. (2000) *Nat. Struct. Biol.* **7**, 560-564.
49. Xu, Y., Glansdorff, N., and Labedan, B. (2006) *BMC Genomics* **7**, 4.
50. Xu, Y., Labedan, B., and Glansdorff, N. (2007) *Microbiol. Mol. Biol. Rev.* **71**, 36-47.
51. Labedan, B., Xu, Y., Naumoff, D. G., and Glansdorff, N. (2004) *Mol. Biol. Evol.* **21**, 10-37.3.
FOOTNOTES

*This work was supported by Public Health Service grants DK064913 (to M. T.) and DK067935 (to D. S.) from the National Institute of Diabetes, Digestive and Kidney Diseases and HD 32652 from the National Institute of Child Health and Human Development.

The atomic coordinates and structure factors (codes 2R8V, 3B8G and 2R98) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (www.pdb.org).

§To whom correspondence should be addressed: Children’s Research Institute, Children’s National Medical Center, 111 Michigan Ave., NW, Washington, DC 20010. Tel.:202-476-5817; Fax: 202-476-6014; E-mail: dshi@cnmcresearch.org.

1Abbreviations used are: AAK, amino acid kinase; AcCoA, acetyl coenzyme A; NAG, N-acetyl-L-glutamate; NAGK, N-acetyl-L-glutamate kinase; NAGS, N-acetyl-L-glutamate synthase; NAT, N-acetyltransferase; OAT, ornithine acetyltransferase; PDB, Protein Data Bank; RMS, root mean square; UMPK, uridylate kinase.

FIGURE LEGENDS

Figure 1. Stereo view of contours of the electron density map (2Fo-Fc) (1.0 σ shown in blue cage) for the AcCoA binding site. The carbon atoms of AcCoA are shown in pink sticks. The carbon atoms of the protein are shown as light-blue sticks. Hydrogen bonds between AcCoA and protein are indicated by red dashed lines.

Figure 2. Ribbon diagram of the NAT domain for the AcCoA, and CoA and NAG complexed structures and the reaction model. A, stereo diagram of the superimposition of the NAT domain of the AcCoA complexed structure (shown in light blue) and the CoA and NAG complexed structure (shown in green). The electron density map (2Fo-Fc) (1.0 σ shown in blue cage) corresponds to bound CoA and NAG. The carbon atoms of AcCoA are shown as pink sticks while the carbon atoms of NAG are shown as yellow sticks. The side chains of Arg316 and Arg425 which are involved in NAG or glutamate binding, are shown as sticks. B, reaction model of N. gonorrhoeae NAGS. The α-amino nitrogen atom of glutamate is 2.8 Å from the carbon atom of the acetyl group, in a position to attack the acetyl group to begin the reaction.

Figure 3. Analytical gel chromatography of N. gonorrhoeae NAGS. A, elution profiles of N. gonorrhoeae NAGS at two different concentrations: 1.7 mg/ml (black) and 5 mg/ml (gray). Down arrows indicate elution volumes of the standard proteins ferritin (1) and catalase (2). B, determination of molecular weight of N. gonorrhoeae NAGS. Plot of sigma factor versus log(molecular weight) of standard proteins (open symbols) and N. gonorrhoeae NAGS (solid symbol). The sigma factor (σ) of N. gonorrhoeae NAGS and standard proteins were calculated using formula σ = (Ve-Vo)/(Vi-Ve), where Ve is elution volume of standard proteins and N. gonorrhoeae NAGS, V_o is void volume and V_i is included volume.

Figure 4. Ribbon diagram of a monomer (A, shown in stereo), and its AAK (B) and NAT (C) domains. Green arrows indicate the direction of strands in β-sheets, α-helices are in red and β-sheets are in green. AcCoA is represented as a stick model.

Figure 5. Molecular hexamer of N. gonorrhoeae NAGS and its interfaces. A, simplified model showing the quaternary interactions of the subunits. B, dimer interfaces across the two-fold axes between AAK domain K1 and K4. C, dimer interface between AAK domain K1 and K5. D, interface across the three-fold axis between the NAT domain S1 and the AAK domain K2. Different subunits are shown in different colors. The side chains of residues which are involved in interface interactions are shown in sticks.
### Table I. Data Collection and Refinement Statistics

| Table I. Data Collection and Refinement Statistics | Native data | MAD data |
|--------------------------------------------------|-------------|----------|
| **Data collection** | AcCoA P312 | AcCoA P312 |
| **Space group** | CoA and NAG P312 | Remote P312 |
| **Ligands** | **Peak** | **Inflection** |
| AcCoA | 0.97936 | 0.98950 |
| CoA and NAG | 0.97472 | 0.97472 |
| **Resolution (Å)** | 30-2.5 (2.59-2.50) | 30-2.5 (2.59-2.50) |
| **Wavelength (Å)** | 1.00000 | 1.54178 |
| **Unit-cell parameters (Å)** | | |
| a = b = 96.7 | a = b = 98.7 | a = b = 99.0 |
| c = 89.8 | c = 89.8 | c = 89.3 |
| | | c = 89.4 |
| | | c = 89.3 |
| **Unique reflections** | 17,297 (1,712) | 33,925 (3,417) |
| **Redundancy** | 10.9 (8.8) | 4.8 (4.2) |
| **Completeness (%)** | 99.1 (98.8) | 99.9 (99.8) |
| **<I/σ(I)>** | 30.6 (2.0) | 8.6 (74.5) |
| **R<sub>merge</sub> (%)<sup>b</sup>** | 9.7 (58.4) | 8.1 (52.5) |
| **No. of protein atoms** | 3227 | 3227 |
| **No. of solvent atoms** | 117 | 92 |
| **No. of hetero atoms** | 51 | 51 |
| **Rmsd of bond lengths (Å)** | 0.011 | 0.010 |
| **Rmsd of bond angle (°)** | 2.0 | 2.1 |
| **R<sub>work</sub> (%)<sup>d</sup>** | 19.2 (28.2) | 19.4 (27.1) |
| **Average B factor (Å<sup>2</sup>)** | 27.6 | 45.5 |
| **Ramachandran plot (%)** | | |
| Favored | 89.6 | 87.2 |
| Allowed | 9.8 | 12.3 |
| Generous | 0.3 | 0.3 |
| Disallowed | 0.3 | 0.3 |

<sup>a</sup> Figures in brackets apply to the highest-resolution shell.

<sup>b</sup> \( R_{\text{merge}} = \sum \sum |I(h,i) - <I(h)>| / \sum \sum |I(h,i)| \), where \( I(h,i) \) is the intensity of the \( i \)th observation of reflection \( h \), and \(<I(h)>\) is the average intensity of redundant measurements of reflection \( h \).

<sup>c</sup> The refinement dataset was prepared by reprocessing MAD remote data set to scale the Friedel pair together to 2.4 Å resolution with \( R_{\text{merge}} \) of 9.2% and 19,707 unique reflections.

<sup>d</sup> \( R_{\text{work}} = \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}} \).

<sup>e</sup> \( R_{\text{free}} = \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}} \) for 5% of the reserved reflections.
Table II. Interactions between AcCoA, CoA or NAG and protein atoms.

| Ligand atoms | Protein atoms | Distance (Å) |
|--------------|---------------|--------------|
| AcCoA or CoA | Bound AcCoA   | Bound CoA and NAG |
| O            | Cys356 N      | 3.33         |
|              | Leu357 N      | 3.07         |
| SBV          | Ser392 OG     | 3.82         |
| NBS          | Leu357 O      | 2.59         |
| OBR          | Thr395 OG1    | 2.57         |
| OBM          | Val359 N      | 3.33         |
| OBK          | Gln364 OE1    | 3.03         |
| OBD          | Asp365 N      | 2.86         |
| OBC          | Gly369 N      | 2.79         |
| OAZ          | Glu370 N      | 3.04         |
| OAY          | Gly366 N      | 3.49         |
|              | Gly367 N      | 2.87         |
| O5*          | Trp398 NE1    | 3.14         |
| O2*          | Arg151 § NE   | 3.70         |
| O3*          | Arg151 § NE   | 3.56         |
|              | Arg151 § NH2  | 3.53         |
| OAQ          | Lys152 § NZ   | 3.56         |
| N1           | Glu397 OE2    | 2.53         |
| N6           | Asn394 O      | 2.52         |
| NAG          | Arg316 NH2    | 3.35         |
|              | Cys356 N      | 2.55         |
| OXT          | Arg316 NH1    | 2.81         |
|              | Leu314 N      | 2.93         |
| OE1          | Arg416 NH2    | 3.64         |
| OE2          | Ser427 OG     | 2.07         |
|              | Arg425 NH1    | 2.37         |
| N2           | Leu391 O      | 3.16         |
| O7           | Cys356 N      | 2.93         |
|              | Leu357 N      | 2.98         |

Arg134 §, Lys151 § and Arg152 § are residues from the K domain of an adjacent monomer (symmetry operation: -y+1, x-y, z).
Figure 1
Figure 2
Figure 3
The crystal structure of N-acetyl-L-glutamate synthase from Neisseria gonorrhoeae provides insights into mechanisms of catalysis and regulation

Dashuang Shi, Vatsala Sagar, Zhongmin Jin, Xiaolin Yu, Ljubica Caldovic, Hiroki Morizono, Norma M. Allewell and Mendel Tuchman

*J. Biol. Chem.* published online January 9, 2008

Access the most updated version of this article at doi: 10.1074/jbc.M707678200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2008/01/10/M707678200.DC1