Published in final edited form as:
J Biol Chem. 2008 March 14; 283(11): 7176–7184. doi:10.1074/jbc.M707678200.

The Crystal Structure of *N*-Acetyl-L-glutamate Synthase from *Neisseria gonorrhoeae* Provides Insights into Mechanisms of Catalysis and Regulation*,†

Dashuang Shi[‡,1], **Vatsala Sagar**[§], **Zhongmin Jin**[¶], **Xiaolin Yu**[‡], **Ljubica Caldovic**[‡], **Hiroki Morizono**[‡], **Norma M. Allewell**[§], and **and Mendel Tuchman**[‡]

[‡]Children's Research Institute, Children's National Medical Center, The George Washington University, Washington, D.C. 20010

[§]Department of Chemistry and Biochemistry, College of Chemical and Life Sciences, University of Maryland, College Park, Maryland 20742

[¶]Southeast Regional Collaborative Access Team, Advanced Photon Source, Argonne National Laboratory, Argonne, Illinois 60439

Abstract

The crystal structures of *N*-acetylglutamate synthase (NAGS) in the arginine biosynthetic pathway of *Neisseria gonorrhoeae* complexed with acetyl-CoA and with CoA plus *N*-acetylgluta-mate 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 and 100Å, respectively, and a height of 110Å. Each AAK domain interacts with the cognate domains of two adjacent monomers across two 2-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, Arg^{316}, Arg^{425}, and Ser^{427}, anchor *N*-acetylglutamate in a position at the active site to form hydrogen bond interactions to the main chain nitrogen atoms of Cys^{356} and Leu^{314}, and hydrophobic interactions to the side chains of Leu^{313} and Leu^{314}. The mode of binding of acetyl-CoA and CoA is similar to other NAT family proteins. The AAK domain, although catalytically inactive, appears to bind argi-nine. This is the first reported crystal structure of any NAGS, and it provides insights into the catalytic function and arginine regulation of NAGS enzymes.

*This work was supported by Public Health Service Grants DK064913 (to M. T.) and DK067935 (to D. S.) from the NIDDK, National Institutes of Health and Grant HD 32652 from the National Institute of Child Health and Human Development. This work was also supported in part by the U. S. Department of Energy under Contract W-31-109-Eng-38. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

†The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1 and S2.

[1] To whom correspondence should be addressed: Children's Research Institute, Children's National Medical Center, 111 Michigan Ave., N.W., Washington, D.C. 20010-2970. Tel.: 202-476-5817; Fax: 202-476-6014; dshi@cnmcresearch.org..

The atomic coordinates and structure factors(code2R8V,3B8G,and2R98)have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

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)[2] 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 2.3.1.35) to form an acetyl cycle (1, 2). In these organisms, NAGS seems to play only an anaplerotic role, because 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). *Neisseria 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 micro-organisms that are inhibited by arginine, this amino acid activates mammalian NAGS. Deficiency of NAGS and consequentially, of NAG, causes hyperammonemia because of 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 showing 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-related 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.

[2]The abbreviations used are: NAGS, *N*-acetyl-L-glutamate synthase; AAK, amino acid kinase; AcCoA, acetyl-CoA; NAG, *N*-acetyl-L-glutamate; NAGK, *N*-acetyl-L-glutamate kinase; NAT, *N*-acetyltransferase; OAT, ornithine acetyltransferase; SeMet, selenomethionine; MAD, multiple-wavelength anomalous dispersion.

EXPERIMENTAL PROCEDURES

Enzyme Production and Crystallization

The *argA* gene was PCR-amplified from the *N. gonorrhoeae* genomic DNA (ATCC 53420) using the primers CGGCATATGAACGCGCCCGACAGCTTTGT and CTTCTCGAGATATCAGCGGTGCAGGCGAC 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 *Neisseria 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 *Escherichia coli* BL21(DE3) cells (Invitrogen) and purified in a two-step procedure using nickel affinity 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% polyethylene glycol 3350, and 100 mM sodium citrate at pH 5.8. The selenomethionine (SeMet) substituted protein was prepared using Overnight Auto-induction System 2 (Novagen) as described previously (15, 16). The SeMet derivative was purified by the same procedure as the wild type protein. On average, 84% of methionines were substituted by SeMet as determined by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (15, 16). The SeMet derivative crystals were grown under conditions identical to the wild type crystals. The fluorescence scan of SeMet 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 Southeast Regional Collaborative Access Team 22-ID equipped with MAR300 CCD at the Advanced Photon Source 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 NIDDK, National Institutes of Health. The crystals were cryo-protected by supplementing the well solution with 30% ethylene glycerol. MAD data sets at three wavelengths 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. The data collection parameters are listed in Table 1. The values for the packing density calculation (19) suggest that there is one monomer in the asymmetric unit, and the solvent content of the crystal is 54% (v/v).

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 data set using the 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 parameters (27) were included and refined using REFMAC5 (28). The translation/libration/screw groups were selected using the TLSMD server (29). Statistics for the final refined model are given in Table 1. Refinements for the CoA and NAG complexed data set and a rescaled MAD remote data set were carried out with a similar procedure. The three refined structures are essentially identical with a root mean square deviation 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 (Figs. 1 and 2*A*).

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. 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 $MgCl_2$, 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% $FeCl_3$, 8% trichloroacetic acid, 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 Healthcare) 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 K_2HPO_4-KH_2PO_4, 150 mM KCl, 20% glycerol (v/v), pH 7.5, was used as the mobile phase. Molecular weight markers, consisting of thyroglobin, ferritin, catalase, bovine serum albumin, ovalbumin, and myoglobin, were used for column calibration. Blue dextran 2000 and vitamin B_{12} 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 mass of 336 kDa, corresponding to the predicted molecular weight of a hexamer (Fig. 3).

Modeling of the Ternary Complex

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 ~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 m_M of arginine.

The purified enzyme has no detectable NAGK activity. This is consistent with the presence of a separate gene encoding NAGK (accession number 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_6 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 with those in structures at similar resolution in the Data Bank. 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 Glu^{145} 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 (Glu^{284}-Ala^{285}-Phe^{286}) connects the N-terminal AAK domain (Asp^5–Lys^{283}) and the C-terminal GCN5-related NAT domain (Val^{287}–Arg^{436}) (Fig. 4*A*). 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. 4*B* and supplemental 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 root mean square deviation 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 (supplemental Fig. S2), as in unliganded *Mycobacterium tuberculosis* NAGK (Protein Data Bank 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. 4*C*). 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 Ala^{355} and Cys^{356}. DALI data base searching indicates that the structure of the NAT domain of *N. gonorrhoeae* NAGS is most closely related to GCN5 histone acetyltransferase (Protein Data Bank code 1M1D) with a z score of 16.4 (42). Superimposition of the two structures results in a root mean square deviation 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 3-fold crystallographic symmetry axis coincides with the molecular 3-fold axis, whereas the 2-fold crystallographic symmetry axis aligns with the molecular 2-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 diameters of ~100 and 20 Å, respectively, and a height of 110 Å (Fig. 5*A*). The diameter of the inner ring is smaller than in reported argi-nine-sensitive NAGK structures, which have inner ring diameters of 26–32 Å, depending on whether arginine is bound (12).

The three dimers and their intersubunit interactions are best viewed perpendicularly to the 3-fold axis. From this view, the AAK domains form a zigzag equatorial belt, whereas the NAT domains are polar (Fig. 5*A*). Each AAK domain interacts with the AAK domain of two other subunits. In the crystal, the molecular hexamers are packed together with their 3-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. 5, *B* and *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. 5*C*). This interface consists exclusively of residues from strand 5, helix 4, helix 5, and the loop between strands 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 Leu^{125}, Ser^{127}, and Asn^{129} 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 Gln^{96}, Arg^{103}, Ser^{104}, Glu^{107}, Asn^{129}, Gln^{161}, and His^{175}, and the hydrophobic residues of Val^{126}, Phe^{160}, and Ala^{164} 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 nitrogen helices from adjacent monomers (*e.g.* K1 and K5; Fig. 5*B*). The long N-terminal extension that 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 (Pro^{16}). 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 Tyr^{16}, corresponding to Tyr^{15} and Tyr^{21} in *T. maritima* and *P. aeruginosa* NAGK, respectively, is found in this segment. Residues Phe^{7}, Phe^{11}, Ala^{14}, Ile^{18}, and His^{10} interlace the two adjacent N-terminal helices, linking two adjacent monomers across a 2-fold axis. In addition, the polar side chains of Arg^{12}, Arg^{19}, and Arg^{22} on the other side of the helix hydrogen bond to Asp^{46}, Gln^{52}, Ser^{122}, and Asp^{164} 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. Because the hexamer appears to be essential for an argi-nine 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 and 1340 $Å^2$, 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 3-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. 5*D*). Interactions at this interface are not as extensive as interfaces within a dimer and mainly involve polar residues. The side chains of Arg^{134}, Asp^{140}, Asp^{143}, Arg^{151}, and Lys^{152} and the main chain nitrogen and/or oxygen atoms of Ile^{136}, Val^{138}, and Gly^{141} from the AAK domain interact with Asn^{393}, Thr^{394}, Asn^{426}, and His^{428} 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 Arg^{134}, Arg^{151}, and Lys^{152} from the AAK domain of the adjacent subunit (Fig. 5*D* and Table 2).

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, because it lacks several key residues. For example, Lys^{8} (*E. coli* NAGK numbering), which is conserved in all NAGK sequences including the bifunctional *X. campestris* NAGS/K, is replaced by Gly^{29} in *N. gonorrhoeae* NAGS, and the space occupied by the Lys side chain in NAGK is occupied by the side chain of Met^{186} in *N. gonor rhoeae* NAGS. Similarly, the conserved Asn^{160}-Ala^{161}-Asp^{162} motif (*E. coli* NAGK numbering) is replaced by Asp^{185}-Met^{186}-Val^{187}, and the essential Lys217 is substituted by Leu^{241}. These differences would be sufficient to render inactive the NAGK domain of *N. gonorrhoeae* NAGS (9).

Active Site and Proposed Cata lytic Mechanism

The well defined electron densities of bound AcCoA, CoA, and NAG clearly identify the catalytic site within the NAT domain (Figs. 1 and 2*A*). AcCoA and CoA bind to the enzyme in a conformation similar to that of other GCN5-related NAT proteins, using the loop joining strand 22 and helix 12 to coordinate the pyrophosphate group (Table 2). The sequence Gln^{364}-Glu^{365}-Gly^{366}-Gly^{367}-Tyr^{368}-Gly^{369} in this loop conforms to the (Arg/ Gln)-Xaa-Xaa-Gly-Xaa-(Gly/ Ala) motif for AcCoA recognition and binding in the GCN5-ralated NAT superfamily members (43). However, unique to the present structure, the side chains of Arg^{134}, Arg^{151}, and Lys^{152} from the adjacent AAK domain are also involved in the interactions with AcCoA (Fig. 5*D*). The adenosine moiety of AcCoA is located on the surface, and its conformation is fixed to the side chain of Glu^{397} via strong hydrogen bonds. AcCoA bends ~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 Ala^{355}, Leu^{357}, and Phe^{399}. The acetyl carbonyl atom hydrogen bonds to the main chain nitrogen atoms of Cys^{356} and Leu^{357}, orienting the acetyl group toward the cavity created by the β-bulge of Ala^{355}– Cys^{356} (Fig.

1). Similar hydrogen bonding interactions between the acetyl carbonyl group and the NH group of residues corresponding to Leu^{357} 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. Ser^{392} 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 Ser^{392} appears to be fixed by hydrogen bonding to the main chain nitrogen atom of Thr^{395}. The pantetheine moiety of AcCoA and CoA interacts with the protein through hydrogen bonds to Val^{359} (NH) and Leu^{357} (C=O), in a manner similar to an anti-parallel β-sheet, whereas its carbonyl oxygen atoms form hydrogen bonds with the side chains of Thr^{395} and Gln^{364}.

Because the acetyl group of AcCoA is completely buried within the protein 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 Arg^{316}, Arg^{425}, and Ser^{427} and the main chain nitrogen atoms of Cys^{356} and Leu^{314} to anchor its two carboxyl groups and the hydrophobic side chains of Leu^{314} and Leu^{391} to interact with its hydrophobic stem. Similar to the acetyl group in AcCoA, the acetyl carbonyl atom of NAG hydrogen bonds to the main chain nitrogen atoms of Cys^{356} and Leu^{357}. The side chain of Arg^{416}, whose position is fixed by hydrogen bonding with the side chain of conserved Glu^{353}, 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 Arg^{316} moves significantly. In the ternary complex, the side chain of Arg^{316} shifts ~8 Å, so that it swings toward the active site (Fig. 2*A*). 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 (Cys^{356}) 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. 2*B*). 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 nitrogen atoms of Cys^{356} and Leu^{357}. When the products (CoA and NAG) form, the sulfur atom of CoA moves ~0.9 Å toward the side chain of Ser^{392}, 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 Arg^{316} moves away to allow NAG to dissociate from the protein. The main chain carbonyl oxygen atom of Leu^{391}, which hydrogen bonds to amino nitrogen atom of NAG and close to the sulfur atom of CoA, may help to transfer the proton from the amino nitrogen atom to the sulfur 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 approximately 440 amino acid residues (8). The present *N. gonorrhoeae* structure clearly indicates that the NAT domain is essential for catalysis, whereas the AAK domain has probably a regulatory function. This assertion is consistent with the observations in various NAGS enzymes indicating that the NAT domain is essential for catalysis, whereas 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 argininosuccinate 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).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. David Davies for facilitating use of the diffraction equipment in the Molecular Structure Section at the NIDDK, National Institutes of Health and Dr. Fred Dyda for help in data collection.

REFERENCES

1. Cunin R, Glansdorff N, Pierard A, Stalon V. Microbiol. Rev. 1986; 50:314–352. [PubMed: 3534538]
2. Slocum RD. Plant Physiol. Biochem. 2005; 43:729–745.
3. Caldovic L, Tuchman M. Biochem. J. 2003; 372:279–290. [PubMed: 12633501]
4. Martin PR, Mulks MH. J. Bacteriol. 1992; 174:2694–2701. [PubMed: 1339419]
5. Grisolia S, Cohen PP. J. Biol. Chem. 1953; 204:753–757. [PubMed: 13117851]
6. Bachmann C, Colombo J, P. Jaggi K. Adv. Exp. Med. Biol. 1982; 153:39–45. [PubMed: 7164912]
7. Caldovic L, Morizono H, Gracia Panglao M, Gallegos R, Yu X, Shi D, Malamy MH, Allewell NM, Tuchman M. Biochem. Biophys. Res. Commun. 2002; 299:581–586. [PubMed: 12459178]
8. Caldovic L, Morizono H, Yu X, Thompson M, Shi D, Gallegos R, Allewell NM, Malamy MH, Tuchman M. Biochem. J. 2002; 364:825–831. [PubMed: 12049647]
9. Qu Q, Morizono H, Shi D, Tuchman M, Caldovic L. BMC Biochem. 2007; 8:4. [PubMed: 17425781]
10. Errey JC, Blanchard JS. J. Bacteriol. 2005; 187:3039–3044. [PubMed: 15838030]
11. Ramon-Maiques S, Marina A, Gil-Ortiz F, Fita I, Rubio V. Structure. 2002; 10:329–342. [PubMed: 12005432]

12. Ramon-Maiques S, Fernandez-Murga ML, Gil-Ortiz F, Vagin A, Fita I, Rubio V. J. Mol. Biol. 2006; 356:695–713. [PubMed: 16376937]
13. Dyda F, Klein DC, Hickman AB. Annu. Rev. Biophys. Biomol. Struct. 2000; 29:81–103. [PubMed: 10940244]
14. He H, Ding Y, Bartlam M, Sun F, Le Y, Qin X, Tang H, Zhang R, Joachimiak A, Liu J, Zhao N, Rao Z. J. Mol. Biol. 2003; 325:1019–1030. [PubMed: 12527305]
15. Shi D, Caldovic L, Jin Z, Yu X, Qu Q, Roth L, Morizono H, Hathout Y, Allewell NM, Tuchman M. Acta Crystallo-graph. Sect. F Struct. Biol. Cryst. Commun. 2006; 62:1218–1222.
16. Shi D, Yu X, Roth L, Morizono H, Hathout Y, Allewell NM, Tuchman M. Acta Crystallogr. F Struct. Biol. Crystalliz. Comm. 2005; 61:676–679.
17. Otwinowski Z, Minor W. Methods Enzymol. 1997; 276:307–326.
18. Collaborative Computational Project, Number 4. Acta Crystallogr. Sect. D Biol. Crystallogr. 1994; 50:760–763. [PubMed: 15299374]
19. Matthews BW. J. Mol. Biol. 1968; 33:491–497. [PubMed: 5700707]
20. Uson I, Sheldrick GM. Curr. Opin. Struct. Biol. 1999; 9:643–648. [PubMed: 10508770]
21. Schneider TR, Sheldrick GM. Acta Crystallogr. Sect. D Biol. Crystallogr. 2002; 58:1772–1779. [PubMed: 12351820]
22. Terwilliger TC. Acta Crystallogr. Sect. D Biol. Crystallogr. 2003; 59:38–44. [PubMed: 12499537]
23. Emsley P, Cowtan K. Acta Crystallogr. Sect. D Biol. Crystal-logr. 2004; 60:2126–2132.
24. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Crosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice L, Simonson MT, Warren GL. Acta Crystallogr. Sect. D Biol. Crystallogr. 1998; 54:905–921.
25. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Acta Crystallogr. Sect. A. 1991; 47:110–119. [PubMed: 2025413]
26. Brünger AT. Nature. 1992; 355:472–475. [PubMed: 18481394]
27. Winn MD, Isupov MN, Murshudov GN. Acta Crystal-logr. Sect. D Biol. Crystallogr. 2001; 57:122–133.
28. Murshudov GN, Vagin AA, Dodson EJ. Acta Crystallogr. Sect. D Biol. Crystallogr. 1997; 55:240–255. [PubMed: 15299926]
29. Painter J, Merritt EA. Acta Crystallogr. Sect. D Biol. Crystal-logr. 2006; 62:439–450.
30. Haas D, Leisinger T. Eur. J. Biochem. 1975; 52:377–393. [PubMed: 240684]
31. Chacón P, Wriggers W. J. Mol. Biol. 2002; 317:375–384. [PubMed: 11922671]
32. Humphrey W, Dalke A, Schulten K. J. Mol. Graphics. 1996; 14:33–38.
33. DeLano WL. The PyMOL Molecular Graphics System. 2002
34. Marvil DK, Leisinger T. J. Biol. Chem. 1977; 252:3295–3303. [PubMed: 16890]
35. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. J. Appl. Crystallogr. 1993; 26:283–291.
36. Marina A, Alzari PM, Bravo J, Uriarte M, Barcelona B, Fita I, Rubio V. Protein Sci. 1999; 8:934–940. [PubMed: 10211841]
37. Ramon-Maiques S, Marina A, Uriarte M, Fita I, Rubio V. J. Mol. Biol. 2000; 299:463–476. [PubMed: 10860751]
38. Marco-Marin C, Gil-Ortiz F, Rubio V. J. Mol. Biol. 2005; 352:438–454. [PubMed: 16095620]
39. Briozzo P, Evrin C, Meyer P, Assairi L, Joly N, Barzu O, Gilles AM. J. Biol. Chem. 2005; 280:25533–25540. [PubMed: 15857829]
40. Kotaka M, Ren J, Lockyer M, Hawkins AR, Stammers DK. J. Biol. Chem. 2006; 281:31544–31552. [PubMed: 16905770]
41. Marco-Marin C, Gil-Ortiz F, Perez-Arellano I, Cervera J, Fita I, Rubio V. J. Mol. Biol. 2007; 367:1431–1446. [PubMed: 17321544]
42. Holm L, Sander C. J. Mol. Biol. 1993; 233:123–138. [PubMed: 8377180]
43. Neuwald AF, Landsman D. Trends Biochem. Sci. 1997; 22:154–155. [PubMed: 9175471]
44. Marmorstein R, Roth SY. Curr. Opin. Genet. Dev. 2001; 11:155–161. [PubMed: 11250138]

45. Vetting MW, Roderick SL, Yu M, Blanchard JS. Protein Sci. 2003; 12:1954–1959. [PubMed: 12930994]
46. Yan Y, Harper S, Speicher DW, Marmorstein R. Nat. Struct. Biol. 2002; 9:862–869. [PubMed: 12368900]
47. Wong LJ, Wong SS. Biochemistry. 1983; 22:4637–4641. [PubMed: 6626521]
48. Sinclair JC, Sandy J, Delgoda R, Sim E, Noble ME. Nat. Struct. Biol. 2000; 7:560–564. [PubMed: 10876241]
49. Xu Y, Glansdorff N, Labedan B. BMC Genomics. 2006; 7:4. [PubMed: 16409639]
50. Xu Y, Labedan B, Glansdorff N. Microbiol. Mol. Biol. Rev. 2007; 71:36–47. [PubMed: 17347518]
51. Labedan B, Xu Y, Naumoff DG, Glansdorff N. Mol. Biol. Evol. 2004; 21:364–373. [PubMed: 14660694]

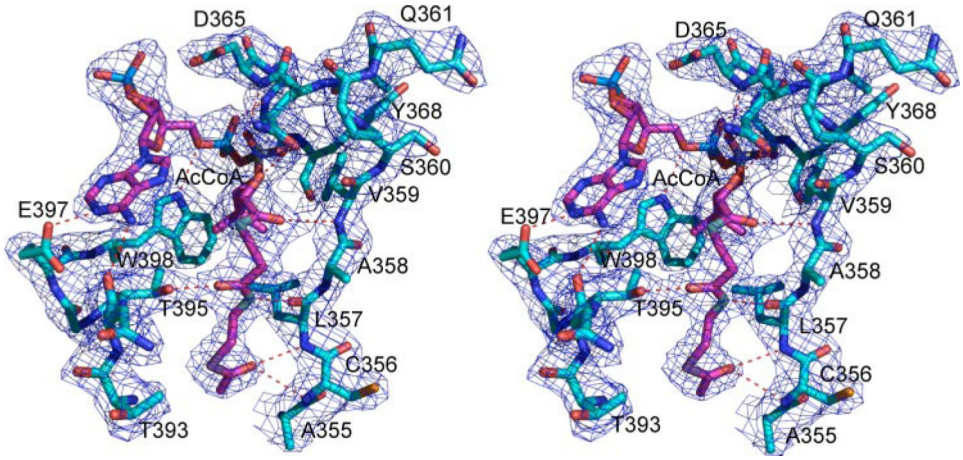

FIGURE 1. Stereo view of contours of the electron density map ($2F_o - F_c$) (1. 0 σ shown in blue cage) for the AcCoA binding site

The carbon atoms of AcCoA are shown as *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*.

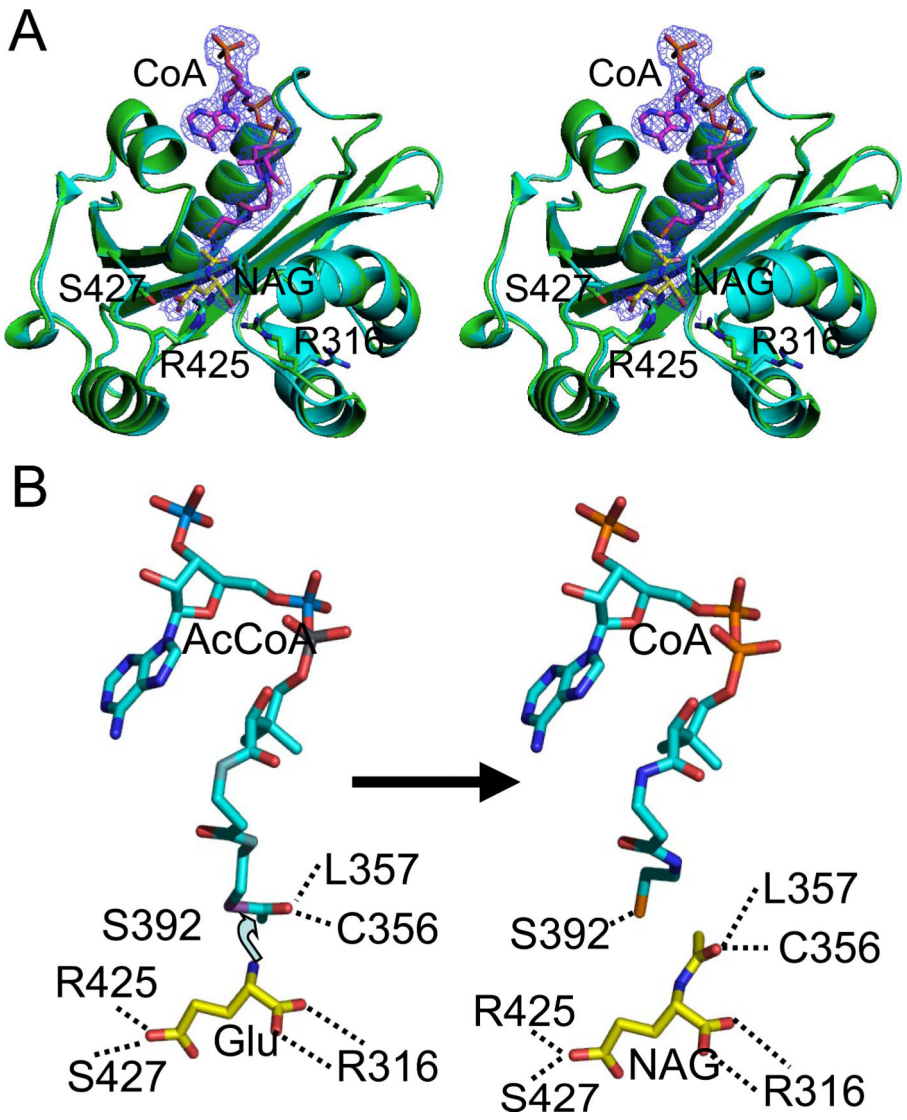

FIGURE 2. Ribbon diagram of the NAT domain for the AcCoA and CoA and NAG complexed structures and the reaction model

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 ($2F_o - F_c$) (1.0 σ shown as *blue cage*) corresponds to bound CoA and NAG. The carbon atoms of AcCoA are shown as *pink sticks*, whereas the carbon atoms of NAG are shown as *yellow sticks*. The side chains of Arg^{316} and Arg^{425}, which are involved in NAG or glutamate binding, are shown as *sticks*. *B*, reaction model of *N. gonorrhoeae* NAGS. Theα-amino nitrogen atom of glutamate is 2.5 Å from the carbon atom of the acetyl group, in a position to attack the acetyl group to begin the reaction.

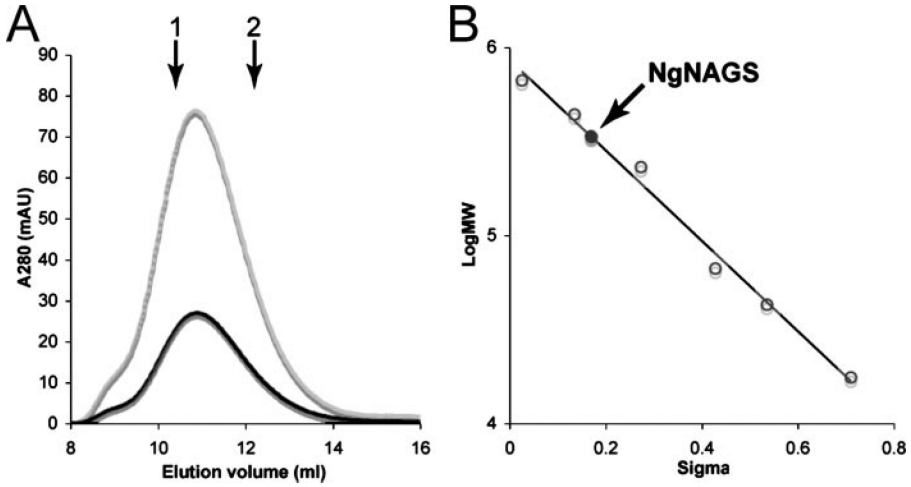

FIGURE 3. Analytical gel chromatography of *N. gonorrhoeae* NAGS

A, elution profiles of *N. gonorrhoeae* NAGS at two different concentrations: 1.7 mg/ml (*black*) and 5 mg/ml (gray). The *down arrows* indicate elution volumes of the standard proteins ferritin (*arrow 1*) and catalase (*arrow 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 the formula $\sigma = (V_e - V_o)/(V_i - V_e)$, where V_e is the elution volume of standard proteins and *N. gonorrhoeae* NAGS, V_o is the void volume, and V_i is the included volume.

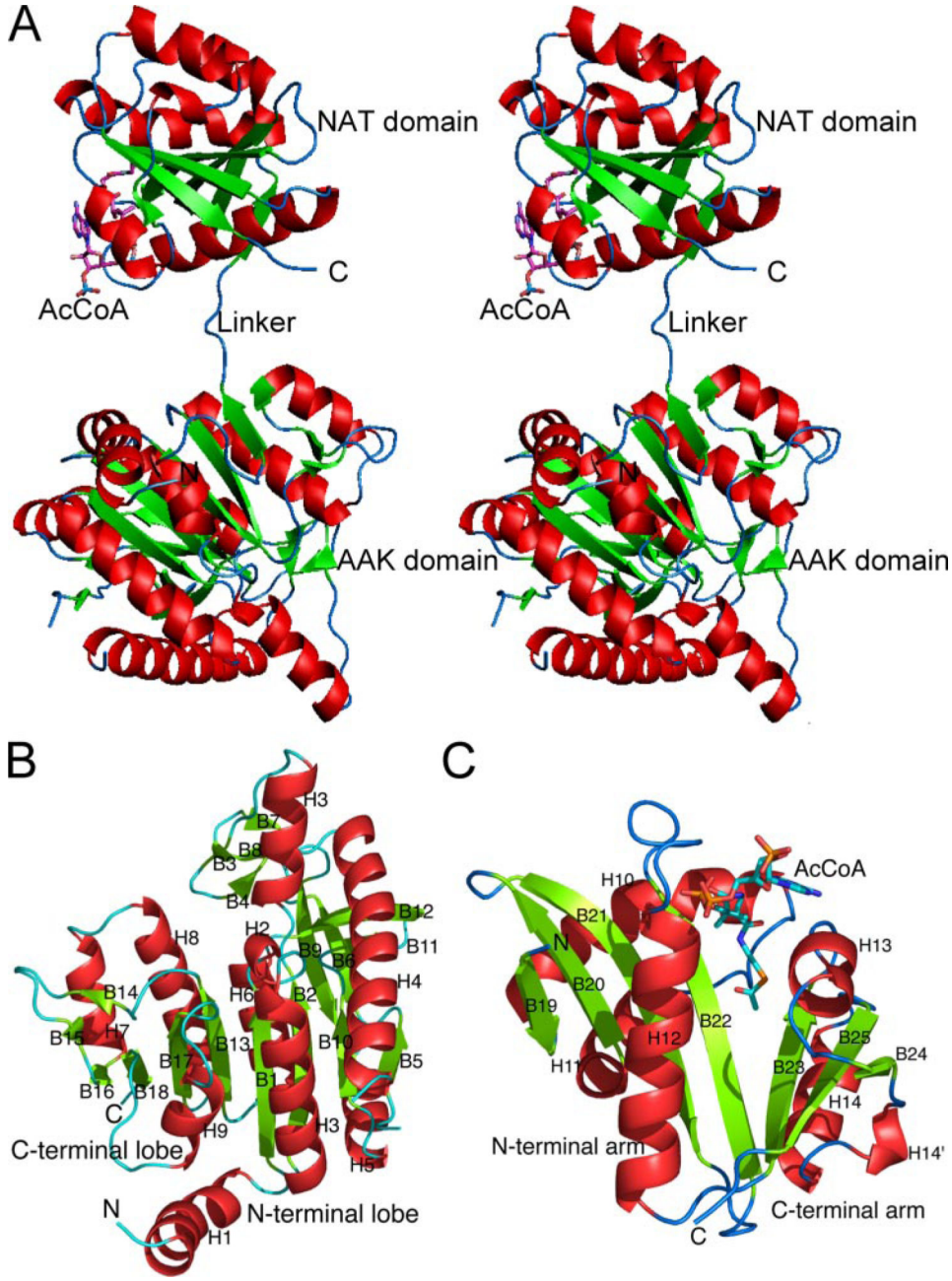

FIGURE 4. Ribbon diagram of a monomer (A, shown in stereo) and its AAK (B) and NAT (C) domains

The *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.

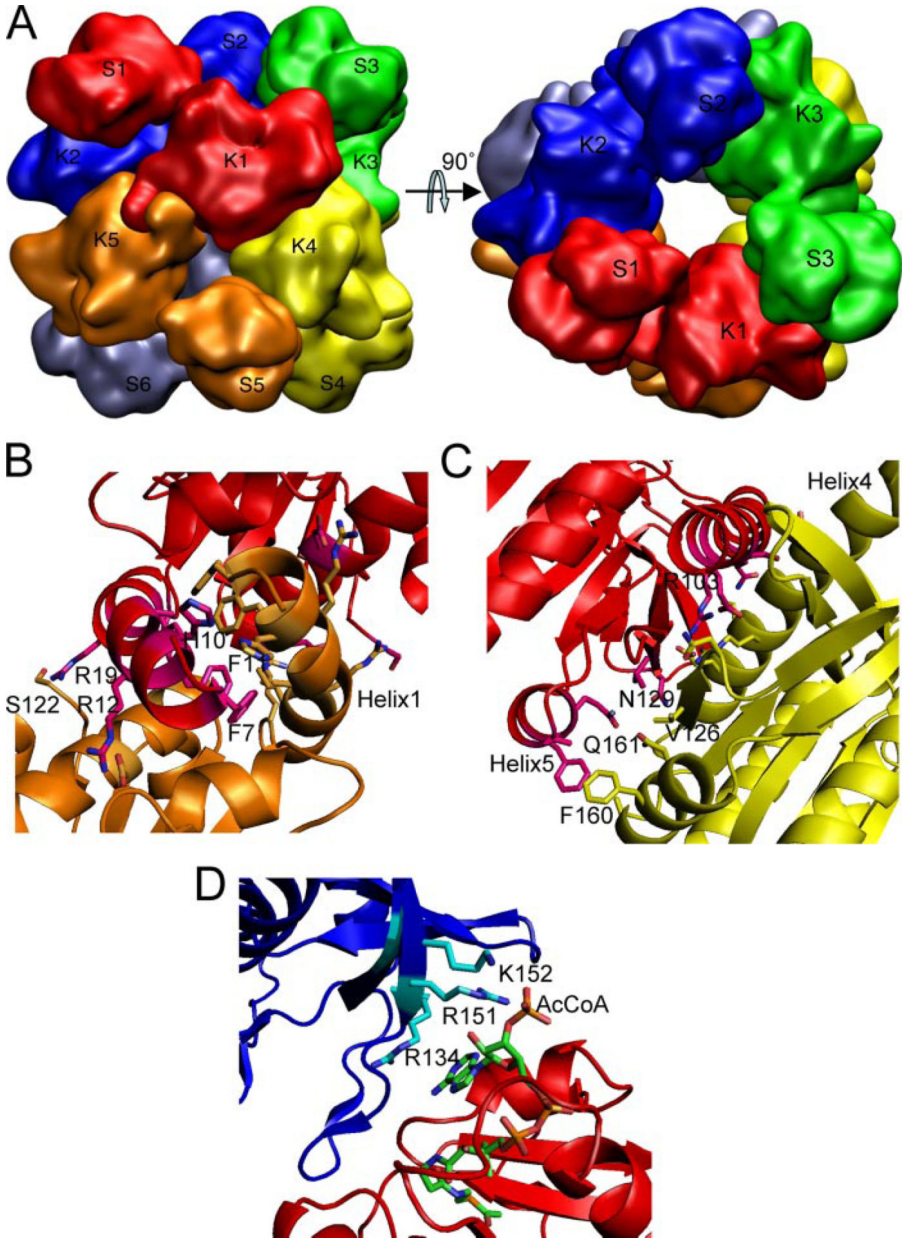

FIGURE 5. Molecular hexamer of *N. gonorrhoeae* NAGS and its interfaces

A, simplified model showing the quaternary interactions of the subunits. *B*, dimer interfaces across the 2-fold axes between AAK domains K1 and K5. *C*, dimer interface between AAK domain K1 and K4. *D*, interface across the 3-fold axis between the NAT domain S1 and the AAK domain K2. Different subunits are shown in different colors. The side chains of residues that are involved in interface interactions are shown in sticks.

TABLE 1

Data collection and refinement statistics

	Native data		MAD data (AcCoA)		
	AcCoA	**CoA and NAG**	**Peak**	**Inflection**	**Remote**
Data collection					
Space group	P312	P312	P312	P312	P312
Wavelength (Å)	1.00000	1.54178	0.97936	0.98950	0.97472
Resolution (Å)	30-2.5 (2.59-2.50)[a]	30-2.6 (2.69-2.60)	30-2.5 (2.59-2.50)	30-2.5 (2.59-2.50)	30-2.5 (2.59-2.50)
Unit cell parameters (Å)	*a* = *b* = 96.7	*a* = *b* 98.7	*a* = *b* 99.0	*a* = *b* 99.1	*a* = *b* 99.0
	c = 89.8	*c* = 89.8	*c* = 89.3	*c* = 89.4	*c* = 89.3
Measurements	188,244	123,451	162,927	161,449	162,319
Unique reflections	17,297 (1,712)	15,494 (1,489)	33,771 (3,369)	33,925 (3,417)	33,830 (3,387)
Redundancy	10.9 (8.8)	8.0 (5.8)	4.8 (4.7)	4.8 (4.2)	4.8 (4.6)
Completeness (%)	99.1 (98.8)	99.2 (96.9)	99.9 (100.0)	99.9 (99.8)	99.9 (100)
$<I/\sigma(t)>$	30.6 (2.0)	13.1 (1.5)	25.7 (2.5)	21.4 (1.2)	25.7 (2.2)
R_{merg}(%)[b]	9.7 (58.4)	9.5 (92.7)	8.0 (46.6)	8.6 (74.5)	8.1 (52.5)
Refinement					
Resolution range (Å)	30–2.5 (2.57–2.50)	30–2.6 (2.66–2.60)			30–2.4 (2.46–2.40)[c]
No. of protein atoms	3227	3227			3227
No. of solvent atoms	117	127			92
No. of hetero atoms	51	61			51
Root mean square deviation of bond lengths (Å)	0.011	0.012			0.010
Root mean square deviation of bond angle (°)	2.0	2.4			2.1
R_{work}(%)[d]	19.2 (28.2)	20.0 (38.1)			19.4 (27.1)
R_{free} (%)[e]	25.2 (29.4)	28.2 (42.6)			24.4 (32.8)
Average *B* factor (Å^2)	27.6	35.1			45.5
Ramachandran plot (%)					
Favored	89.6	87.5			87.2
Allowed	9.8	12.0			12.3
Generous	0.3	0			0.3
Disallowed	0.3	0.5			0.3

[a] Figures in parentheses apply to the highest resolution shell.

[b] $R_{merg} = \Sigma h \Sigma i |I(h,i) - <I(h)>| \Sigma h \Sigma i I(h,i)$, 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*.

[c] The refinement data set was prepared by reprocessing MAD remote data set to scale the Friedel pair together to 2.4 A resolution with R_{merg} of 9.2% and 19,707 unique reflections.

[d] $R_{work} = \Sigma h |F_{obs} - F_{calc}| / \Sigma h F_{obs}$.

[e] $R_{free} = \Sigma h|F_{obs} - F_{calc}|/\Sigma h\, F_{obs}$ for 5% of the reserved reflections.

TABLE 2

Interactions between AcCoA, CoA, or NAG and protein atoms

Ligand atoms	Protein atoms	Distance (Å)	
		Bound AcCoA	Bound CoA and NAG
AcCoA or CoA			
O	Cys^{356} N	3.33	
	Leu^{357} N	3.07	
SBV	Ser^{392} OG	3.82	3.26
NBS	Leu^{357} O	2.59	2.59
OBR	Thr^{395} OG1	2.57	2.64
OBM	Val^{359} N	3.33	3.45
OBK	Gln^{364} OE1	3.03	3.35
OBD	Asp^{365} N	2.86	2.90
OBC	Gly^{369} N	2.79	2.99
OAZ	Glu^{370} N	3.04	3.24
OAY	Gly^{366} N	3.49	3.50
	Gly^{367} N	2.87	2.76
O5*	Trp^{398} NE1	3.14	3.20
O2*	Arg^{151} NE[a]	3.70	3.69
O3*	Arg^{151} NE[a]	3.56	
	Arg^{151} NH_2[a]		3.53
OAQ	Lys^{152} NZ[a]	3.56	3.39
N1	Glu^{397} OE2	2.53	2.35
N6	Asn^{394} O	2.52	3.49
NAG			
O	Arg^{316} NH_2		3.35
	Cys^{356} N		2.55
OXT	Arg^{316} NH1		2.81
	Leu^{314} N		2.93
OE1	Arg^{416} NH_2		3.64
OE2	Ser^{427} OG		2.07
	Arg^{425} NH1		2.37
N2	Leu^{391} O		3.16
O7	Cys^{356} N		2.93
	Leu^{357} N		2.98

[a] Arg^{134}, Lys^{151}, and Arg^{152} are residues from the K domain of an adjacent monomer (symmetry operation: $-y + 1, x - y, z$).