Crystal Structure of N-Succinylarginine Dihydrolase, AstB, bound to Substrate and Product, an Enzyme from the Arginine Catabolic Pathway of *Escherichia coli*

Ante Tocilj\(^1,2\), Joseph D. Schrag\(^1,2\), Yunge Li\(^1,2\), Barbara L. Schneider\(^3\), Larry Reitzer\(^3\), Allan Matte\(^1,2\) and Miroslaw Cygler\(^1,2,*\)

\(^1\)Biotechnology Research Institute, NRC, 6100 Royalmount Avenue, Montréal, Québec, Canada H4P 2R2
\(^2\)Montreal Joint Centre for Structural Biology, Montreal, Quebec, Canada
\(^3\)Department of Molecular and Cell Biology, The University of Texas at Dallas, P.O. Box 830688, Richardson, TX, 75083-0688, U.S.A.

*Corresponding author:
Dr. Mirek Cygler
Biotechnology Research Institute, NRC
6100 Royalmount Avenue
Montréal, Quebec H4P 2R2
Canada
Tel: (514) 496-6321
Fax: (514) 496-5143
Email: mirek@bri.nrc.ca

25 January 2005
Abstract

The ammonia-producing arginine succinyltransferase (AST) pathway is the major pathway in *Escherichia coli* and related bacteria for arginine catabolism as a sole nitrogen source. This pathway consists of five steps, each catalyzed by a distinct enzyme. Here we report the crystal structure of *N*-succinylarginine dihydrolase AstB, the second enzyme of the AST pathway, providing the first structural insight into enzymes from this pathway. The enzyme exhibits a pseudo 5-fold symmetric α/β propeller fold of circularly arranged ββαβ-modules enclosing the active site. The crystal structure indicates clearly that this enzyme belongs to the amidinotransferase (AT) superfamily and that the active site contains a Cys---His-Glu triad characteristic of the AT superfamily. Structures of the complexes of AstB with the reaction product and a Cys365Ser mutant with bound the *N*-succinylarginine substrate suggest a catalytic mechanism that consists of two cycles of hydrolysis and ammonia release, with each cycle utilizing a mechanism similar to that proposed for arginine deiminases. Like other members of the AT superfamily of enzymes, AstB possesses a flexible loop that is disordered in the absence of substrate and assumes an ordered conformation upon substrate binding, shielding the ligand from the bulk solvent, thereby controlling substrate access and product release.
Introduction

Arginine is an energy-rich amino acid that can supply nitrogen, carbon and energy to various bacteria in a variety of environments. Arginine can be catabolized by a surprisingly large number of routes including the arginase pathway, the arginine deaminase (ADI) pathway, the arginine succinyltransferase (AST) pathway, the arginine transaminase, oxidase and oxygenase pathways, the arginine decarboxylase (ADC) pathway as well as others (1,2). Pseudomonas aeruginosa has four of these pathways (1). These pathways often have distinctive functions. For example, the arginine deiminase pathway generates carbamoyl phosphate for substrate level phosphorylation when oxygen is limiting (1). The presence of a particular arginine catabolic pathway may increase an organism’s ability to inhabit a much broader ecological niche.

Escherichia coli and related bacteria have two such metabolic routes (EcoCyc, http://ecocyc.org/ (3)), the ADC and AST pathways (4). The AST pathway accounts for 97% of arginine catabolism, whereas the ADC pathway accounts for only 3% (4). The AST pathway converts the carbon skeleton of arginine into glutamate, with the concomitant production of ammonia and conversion of succinyl-CoA to succinate and CoA (Fig. 1). The AST pathway consists of five enzymes: arginine succinyltransferase (AstA, EC 2.3.1.109), succinylarginine dihydrolase (AstB, EC 3.-.-.-), succinylornithine transaminase (AstC, EC 2.6.1.-), succinylglutamic semialdehyde dehydrogenase (AstD, EC 1.2.1.-) and succinylglutamate desuccinylase (AstE, EC 3.5.1.-), all contained within the astCADBE operon (aruCFGDBE operon in Pseudomonas aeruginosa (5)). Nitrogen limitation induces transcription of the operon; ast mutants cannot utilize arginine as a nitrogen source and are impaired in ornithine utilization (4). Therefore, one function of the AST pathway is to provide nitrogen during nitrogen restriction. The ammonia produced is assimilated into glutamate and glutamine, which in turn
provide nitrogen for the synthesis of virtually all nitrogen-containing compounds. Entry into stationary phase also induces the ast operon and an ast mutant strain survives only poorly under conditions of carbon starvation (6,7). The observed phenotype may result from diminished generation of citric acid cycle intermediates. Slow growth lowers polyamine pools and nitrogen limitation induces a variety of polyamine catabolic operons (8). The AST pathway has also been proposed to contribute to polyamine homeostasis by controlling levels of intracellular arginine and ornithine, the substrates for putrescine synthesis (8). Because arginine catabolism and the AST pathway perform several important physiological functions in E. coli, analogues of AST intermediates have been suggested as potential antimicrobial agents against pathogenic E. coli (9).

Of the five E. coli AST enzymes two have homologs (at a level of ~35% sequence identity) with known three-dimensional structures, namely AstC (PDB code 1OAT, (10), ISFF, (11)) and AstD (1UZB, Inagaki and Tahirov, Protein Databank http://www.rcsb.org/pdb). Succinylglutamate desuccinylase AstE was predicted to be a member of the Zn-dependent carboxypeptidase family (12). Recently, Shirai and Mizuguchi, using sophisticated sequence analysis and fold recognition tools, proposed assignment of AstA and AstB to the acyl-CoA N-acyltransferase and amidinotransferase (β/α-propeller) fold families, respectively (9). N-succinylarginine dihydrolase (AstB), the second enzyme in the AST pathway, converts N-succinylarginine into N-succinylornithine with the release of ammonia and carbon dioxide (Schema 1). The residues involved in catalysis were proposed to be Asp173, His248 and Cys365, with the latter playing the role of a nucleophile (9). Here we report the crystal structure of AstB from E. coli, its Cys365Ser mutant, and their complexes with substrate and product. The protein
does indeed have the β/α-propeller fold and contains a Cys---His-Asp catalytic triad with similarity to other amidinotransferases, suggesting a similar catalytic mechanism.

**Experimental Procedures**

**Cloning, Expression, and Purification**

The *astB* gene was cloned into a derivative of the pET-15b vector (Amersham Biosciences). The Cys365Ser point mutation was introduced using QuikChange™ mutagenesis according to the manufacturer's instructions (Stratagene) and verified by DNA sequencing. The BL21(DE3) strain was transformed by the plasmid DNA and cells were grown at 37 °C to an $A_{600}$ of ~0.8 in Circle Grow medium (Bio101 Inc.). Recombinant protein expression was induced by addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 100 µM in a 1 liter culture that was maintained at room temperature for an additional 15 h. SeMet-labeled protein was prepared by transforming the *E. coli* methionine auxotroph DL41(DE3) with the plasmid DNA and the cells were grown in LeMaster medium supplemented with 25 mg/l of L-selenomethionine for SeMet labeling (13).

Cells were harvested by centrifugation (4000 x g, 4 °C, 25 min) and were re-suspended in 40 ml of lysis buffer (50 mM Tris-HCl, pH 7.5, 0.4 M NaCl, 5% (w/v) glycerol, 20 mM imidazole, 10 mM β-mercaptoethanol) containing one dissolved tablet of Complete™ protease inhibitor cocktail (Roche Diagnostics). Cells were lysed by sonication on ice for a total of five 30 s pulses with 45 s between each pulse for cooling. The lysate was then cleared by centrifugation (100,000 x g, 4 °C, 30 min). The protein supernatant was loaded on a 5 ml DEAE-Sepharose (Pharmacia) column equilibrated with lysis buffer, the flow-through fraction was collected and applied to a 5 ml nickel-nitrilotriacetic acid column (NiNTA, Qiagen), pre-equilibrated with lysis
buffer. The column was washed extensively with buffer (50 mM Tris-HCl, pH 7.5, 50 mM imidazole, 0.4 M NaCl) and bound protein was eluted with the same buffer containing 150 mM imidazole. The protein was subsequently concentrated for crystallization with a concomitant buffer exchange by ultrafiltration to 50 mM Tris-HCl, pH 7.5, 0.2 M NaCl, 5% (w/v) glycerol, 10 mM DTT.

Dynamic light scattering measurements were carried out at 22 °C on a DynaPro Plate Reader (Protein Solutions, Inc., Charlottesville, VA) at a protein concentration of 4 mg/ml. Gel filtration chromatography was performed using a Superose-12 column equilibrated in buffer (50 mM Tris-HCl, pH 7.5, 0.2 M NaCl) connected to an Akta Express FPLC system (Amersham). A sample of purified AstB was loaded at a flow rate of 0.8 ml/min and the elution volume, V_e, determined. The apparent molecular weight was calculated using a standard curve of protein markers from a gel filtration calibration kit (Sigma Chemical Co.).

**Crystallization and Data Collection**

Initial crystallization conditions were identified by sparse matrix screening using Screen I and II (Hampton Research, Liguna Niguel, CA). A triclinic crystal form was obtained from the His-tagged, SeMet-labeled protein after 5 days at 20 °C in hanging drops containing 2 µl of protein (7.3 mg/ml) in buffer (50 mM Tris-HCl, pH 7.5, 0.2 M NaCl, 5% (w/v) glycerol, 10 mM DTT, 4 mM glutamate) and 2 µl of reservoir solution (15% (w/v) PEG 10,000, 100 mM HEPES buffer pH 7.5). These P1 crystals diffract to 2.3 Å resolution and have unit cell dimensions a=55.6 Å, b=93.8 Å, c=139.4 Å, α=104.7°, β=101.5°, γ=90.0° with six monomers per asymmetric unit. Three MAD datasets (peak, inflection and remote) about the Se K-edge were collected at beamline X8C, NSLS, Brookhaven National Laboratory, using a Quantum-4 CCD area detector (ADSC, San Diego, CA) (Table 1). An additional dataset was collected on a
Micromax 007 rotating anode equipped with Osmic mirrors and an HTC image plate detector (Rigaku/MSC, The Woodlands, TX) (Table 1).

A monoclinic crystal form diffracting to ~3Å resolution was obtained from the same starting protein preparation after 15 days at 20 °C in hanging drops containing 2 µl of protein (4.8 mg/ml) in buffer (50 mM Tris-HCl pH 7.5, 0.2 M NaCl, 5% (w/v) glycerol, 10 mM DTT) and 2 µl of reservoir solution (11% (w/v) PEG 10,000, 5% (w/v) glycerol, 0.1 M cacodylate pH 6.6, 0.1 M KCl, 0.25 M MgCl₂). These crystals belong to the space group P2₁ with \(a=108.6 \text{ Å}, b=106.7 \text{ Å}, c=175.5 \text{ Å}, \beta=99.9^\circ\) and contain eight monomers per asymmetric unit.

Sparse-matrix screening of AstB in the presence of 5 mM \(N\)-succinylarginine resulted in an orthorhombic crystal form of the complex. The best crystals grew in hanging drops containing 2 µl of protein (6 mg/ml) in buffer (0.2 M NaCl, 5% (v/v) glycerol, 10 mM DTT, 5 mM \(N\)-succinylarginine) and 2 µl of reservoir solution (8.5% (w/v) PEG 10,000, 0.1 M cacodylate buffer pH 6.4, 0.2 M calcium acetate). They belong to the space group P2₁2₁2₁ with \(a=54.9 \text{ Å}, b=166.9 \text{ Å}, c=186.0 \text{ Å}\) and contain four monomers per asymmetric unit. Isomorphous crystals were obtained for the AstB Cys365Ser mutant complexed with the substrate. Complete datasets for the wild type enzyme and the Cys365Ser mutant co-crystallized with \(N\)-succinylarginine were collected to a resolution of 1.95 Å and 1.7 Å, respectively, at beamline X25, NSLS, Brookhaven National Laboratory on a Quantum-315 CCD area detector (ADSC, San Diego, CA). When native AstB was co-crystallized with the substrate we observed only the product in the crystal structure.

All crystal forms were soaked in a cryoprotectant solution consisting of mother liquor supplemented with 20% (w/v) glycerol, picked up in a nylon loop, and flash cooled at 100 K in
the N₂ (gas) cold stream (Oxford Cryosystems, Oxford, UK). Datasets were integrated and scaled using either HKL2000 (14) or d*trek (15).

**AstB activity assay**

N-succinylarginine was synthesized as described previously (4). Following the method described by Schneider et al. (4) the N-succinylarginine dihydrolase activity was measured in a coupled assay with glutamate dehydrogenase. Purified AstB (32 µg of native or 122 µg of AstB Cys365Ser mutant) was incubated for 1 h at 30 °C with 10 mM N-succinylarginine in 100 mM Tris-HCl, pH 7.5, in a total volume of 1 ml. Following incubation, 200 µl of the reaction mixture was added to 800 µl of 1 mM ADP (Fluka), 1.7 mM α–ketoglutarate (Sigma) and 2.9 mM NADH. Oxidation of NADH to NAD⁺ occurs in a reaction catalyzed by glutamate dehydrogenase:

\[
\text{NADH} + \alpha\text{-ketoglutarate} + \text{NH}_4^+ \rightarrow \text{glutamate} + \text{H}_2\text{O} + \text{NAD}^+ 
\]

This reaction was followed by measuring the decrease in absorbance at 340 nm after addition of 4.9 U (100 µg) of glutamate dehydrogenase (Sigma Chemical Co.). Control incubations without enzyme or substrate were performed in parallel. One unit (U) is defined as the amount of enzyme required to form 1 µmol of product per min at 30 °C.

**Structure solution and refinement**

**P1 crystal form**

This crystal form has 6 monomers per asymmetric unit. These crystals showed a two-fold non-crystallographic axis slightly off the crystallographic axis, generating pseudo C2 symmetry that initially confused the structure solution. The Rsym in space group C2 (~0.08) was
very similar to that for data processed in space group P1 (0.06). MAD phasing was initially performed in space group C2 where three independent monomers were expected per asymmetric unit based on the Matthews coefficient of 2.7 Å³/Da (16). A total of 26 Se sites were located from a three wavelength MAD calculation using the program SOLVE (17) and were used to calculate an electron density map. Density modification with the program RESOLVE (18) resulted in a figure of merit of 0.73. This electron density map was sufficiently clear to build a partial model that was ~60% complete. However, the quality of the electron density map varied significantly from one molecule in the asymmetric unit to the other and subsequent refinement using CNS (19) stalled at an R-factor of 0.48. To eliminate the possibility of systematic errors in the diffraction data as the source of the difficulty, a second data set was collected on this crystal form using a rotating anode source. The diffraction limit for this dataset was similar to that of the synchrotron dataset, suggesting that this crystal was of better quality. The self-rotation function calculated using this dataset merged in space group P1 suggested the presence of non-crystallographic, rather than crystallographic two-fold symmetry. The partial model of an AstB molecule built previously from the C2 electron density map was used as a search model for molecular replacement in the second, P2₁ crystal form described below. Despite the lower resolution of this crystal form we were able to extend the model to encompass residues 2-440. This improved model was subsequently used to locate six independent molecules in the asymmetric unit of the P1 crystal form using the program MOLREP (20). From this point the refinement using CNS (19) decreased the R-factor rapidly, confirming the choice of P1 space group. These data are 92% complete to 2.25Å resolution, with partial data extending to 1.9 Å resolution. The refinement using all available data converged to an R-factor of 0.213 and R-free of 0.251 (Table 1). The final model includes six independent monomers, each containing
residues Asn2-Ala19 and His31-Arg441. Residues Gly20-Arg30 are disordered and were not modeled. Difference electron density maps showed a strong peak in each molecule with clear octahedral coordination by oxygen atoms from the surface loop Ala340-Ser346, suggesting a bound metal ion. The ion-oxygen distances were in the range 2.6-3.1 Å with the majority between 2.7-2.9 Å. Based on these distances the site is most likely occupied by a sodium or potassium ion. This density was modeled as a potassium ion, and yielded reasonable B-factors during subsequent refinement. This ion likely plays a structural role as it is far from the active site region.

**P2₁ crystal form**

This crystal form diffracts only to 3Å resolution and has 8 molecules in the asymmetric unit. Initially, molecular replacement using a partial model (see above) located seven monomers in the asymmetric unit. The electron density was sufficient to extend and partially refine the model. Upon completion of the refinement of the model in the P1 crystal form, this refined model was used to locate all eight monomers in the asymmetric unit by molecular replacement (20). This structure was refined using CNS (19) to an R-factor of 0.26 and R-free of 0.28. The loop Ala19-Arg32 was disordered in seven of the eight molecules. The eighth molecule showed this loop in a closed conformation. As the resolution of this data is low and there are no substantial difference between these models and the higher resolution P1 form, these coordinates were not deposited.

**P2₁2₁2₁ crystal form**

The structures of the AstB-substrate complexes were solved by molecular replacement and refined using the program CNS (19). Each of the four independent molecules contained
residues 2-445. The loop Gly20-Arg30, poorly ordered in other structures, is well ordered in every molecule. A ligand molecule is bound to each monomer of AstB. In addition, ~600 solvent molecules were positioned in the electron density. The final R-factor is 0.217 and R-free is 0.245 for the complex of AstB with the \( N \)-succinylornithine reaction product and for the Cys365Ser mutant co-crystallized with the substrate the R-factor is 0.202 and R-free is 0.225 (Table 1). Coordinates of wild-type AstB in space group P1, AstB-succinylornithine complex and the Cys365Ser mutant of AstB with bound succinylarginine (both in space group P2_12_12_1) have been deposited in the RCSB Protein Data bank (21) with PDB codes 1YNF, 1YNH and 1YNI, respectively.

**Results and Discussion**

**Purification, mutagenesis and characterization**

Wild-type *E. coli* AstB protein and the Cys365Ser mutant were purified to apparent homogeneity as assessed by SDS-PAGE and native PAGE. The protein forms dimers in solution as determined by dynamic light scattering and gel filtration chromatography. Purified wild-type AstB was highly active and had a specific activity of 5.3 U/mg as determined by measuring the release of ammonia upon conversion of \( N \)-succinylarginine to \( N \)-succinylornithine. A crude extract of nitrogen-limited (i.e., fully induced) wild type *E. coli* contained 0.025 U/mg total protein activity (4). In contrast, the purified Cys365Ser mutant had a specific activity of only 0.065 U/mg, indicating a crucial role for this cysteine residue in catalysis.
Overall structure of the monomer

The AstB molecule consists of a single globular domain of 447 amino acids with an α/β topology. The domain forms a propeller composed of five repeats (modules) of a ↓β↑α↑β motif arranged circularly around 5-fold pseudo symmetry axis (Fig. 2). This fold has been observed previously and is called the α/β 5-stranded propeller in the CATH classification (version 2.5.1) (22;23) and the pentein β/α-propeller in the SCOP classification (24). The three strands of each module form a mixed β-sheet with the first, N-terminal strand of the repeat lying near the central axis of the propeller. The two innermost strands of the module are anti-parallel, with the α-helix forming a crossover connection to the third β-strand, which is parallel to the second. The α-helix is parallel to this last β-strand and is out of the plane of the β-sheet, on the outside of the propeller (Fig. 2). The connections between the modules vary in length and contain either a short 310-helix or α-helix. The first module, which begins with the N-terminus of the protein, differs somewhat from the other modules. It starts at the middle β-strand, followed by an α-helix and a third β-strand, while the innermost β-strand comes from a C-terminal segment, which follows the fifth module. This organization results in the first (Ala2) and the last (Thr373) residue of the propeller being adjacent to one another. The last ~75 residues (Glu374-Arg441) fold into a two-helix hairpin stacked against the edges of the first and second modules of the propeller. The loops connecting the secondary structure elements within each module on one end of the sheet are short whereas the loops on the opposite end are much longer. The connections between the modules are on the latter side and are also comparatively long. Thus one side of the propeller is relatively flat while the opposite side has a more complex topography.

The independent molecules within the asymmetric unit of each crystal form and across the different crystal forms are very similar overall. The root-mean-squares deviation (rmsd)
between the Cα atoms of molecules across different crystal forms is 0.2-0.4 Å. The only significant difference between the molecules is the conformation of a loop, Ala19-Gln34, which is largely disordered in the apo protein but becomes well ordered in the presence of bound substrate or product.

**Dimer formation**

Gel filtration and dynamic light scattering studies both indicate that AstB forms homodimers in solution. This is consistent with the presence of homodimers in each of the three crystal forms (Fig. 2c). Each dimer has approximate dimensions of 92 Å x 62 Å x 54 Å, with the two monomers being related by a non-crystallographic 2-fold symmetry axis. The dimer interface is formed by a bundle of three short helices. The residues contributing to dimerization are from the second module (residues Asn133-Ser139 and Ala164-Leu171), the helix of the third module (residues Glu216-Leu224) and the long loop that connects the first and second modules (residues Arg74-Phe78 and Trp96). These residues include several isoleucines, phenylalanines, alanines and a proline, giving this surface a partially hydrophobic character. In addition to numerous van der Waals interactions the dimer is further stabilized by hydrogen bonds, some of which are bridged by water molecules. The surface area buried upon dimer formation calculated using the method of Lee and Richards (25) with a 1.4 Å probe radius is 900 Å², which corresponds to 6% of the total surface area of each monomer. The two substrate binding sites within the dimer are positioned on the same side of the elongated dimer but each involves residues from only one monomer, suggesting that the active sites within each monomer of AstB function independently.
Substrate binding site

Attempts to capture the substrate N-succinylarginine by soaking it into the P1 and P2₁ crystal forms were unsuccessful. As this could have been a result of enzymatic hydrolysis of the substrate in the crystal, a Cys365Ser mutant, in which a serine replaced the cysteine nucleophile, was constructed and expressed to eliminate the catalytic activity of the enzyme. Nevertheless, using the mutant AstB we were still unable to detect the enzyme-bound substrate either by soaking the crystals or by cocrystallization under conditions producing the P1 or P2₁ crystal forms. Further screening for suitable crystallization conditions of this inactive AstB mutant in the presence of N-succinylarginine yielded new crystallization conditions that resulted in an orthorhombic crystal form belonging to the space group P2₁2₁2₁. An electron density map calculated from diffraction data collected from this crystal revealed a well-defined substrate molecule bound in the active site (Fig. 3a). As a result of substrate binding all of the residues in the Ala19-Arg32 loop, which were disordered in the native crystal structures, were now clearly defined in electron density.

The substrate binds to the Cys365Ser mutant enzyme on the propeller face that contains the long crossover loops between the modules (see above). The binding site is shaped as a ~15 Å deep tunnel that leads from the surface toward the protein center and rests on the residues from the ends of the innermost strands of the five β-sheets (Fig. 2a,b). The sides of the tunnel are made of residues from the various crossover loops, Ser102-Trp107, His137-Arg138 and Asn359-Gly361, and the bottom is lined with Ala109, Asn110, Ala177, Val251 and Asn306. The entrance to the substrate binding tunnel is shielded by the mobile Ala19-Arg32 loop. The substrate is oriented with its guanidinium group at the bottom of the tunnel and the succinate carboxylate closest to the surface at the entrance to the tunnel (Fig. 2, 3a). All nitrogen and
oxygen atoms of the substrate are involved in direct hydrogen bonds to the enzyme (Fig. 3a, 4a). Hydrogen bonds between the succinyl carboxyl group and ordered water molecules provide additional bridging interactions to the protein. The carboxylate group of the arginine moiety, in the middle of the extended substrate molecule, forms two salt bridges, to Arg212 (two H-bonds) and to Arg138 (one H-bond), and also forms a hydrogen bond to Asn25. The guanidinium group of the substrate is oriented through a salt bridge between its NH1 and NH2 atoms to Asp250 and is hydrogen bonded through NE to OD1 of Asn110. Further, the amide NH (former amino terminus of arginine) is hydrogen bonded to the carbonyl oxygen of Asn359 and the neighboring carbonyl oxygen of the substrate forms a hydrogen bond to the sidechain of His137. Finally, the succinyl carboxyl group interacts with residues from the mobile loop: its OD1 atom is hydrogen bonded to the NH group of Leu21 and the OH of Ser28 while the OD2 is hydrogen bonded to the NH group of Ala19 and a bridging water molecule. Of importance is also the interaction of the Trp107 sidechain stacked against the alkyl chain of the substrate’s arginine moiety. In the substrate-bound state the mobile loop forms a lid over the substrate completely burying it within the protein and contributes several hydrogen bonds to the substrate. In the Cys365Ser mutant, in which the serine replaces the cysteine nucleophile, the Ser365 sidechain is directed away from the substrate and the Ser365 hydroxyl makes two hydrogen bonds to Gly362. Although the active site His248 is 3.3 Å from the guanidinium moiety of the substrate, the imidazole ring is nearly perpendicular to the plane of the guanidinium group indicating that this histidine forms no hydrogen bond to the bound substrate (Fig. 4a).

With the exception of the contacts involving the missing guanidinium moiety, all of the previously described enzyme-substrate interactions are also observed in the wild type enzyme-
product complex (Fig. 3b). Here, the sidechain of the Cys365 nucleophile is directed toward the substrate as expected from its catalytic role.

**Sequence and Structural Similarity**

Sequence analysis using PSI-BLAST (26) identified AstB homologs in 23 bacterial species. No homologs in other kingdoms were found. The sequence identity between *E. coli* AstB and the other bacterial orthologs varies from 85% sequence identity for *Salmonella typhimurium* to 42% for *Zymomonas mobilis*, indicating a high degree of sequence conservation for AstB among these species. Mapping the positions of conserved residues identified in the sequence alignments onto the three-dimensional structure of AstB shows that the strictly conserved residues cluster predominantly around the substrate binding site. Almost all residues that are within a distance of 8 Å from the bound substrate are fully conserved in these sequences and the remaining residues are highly conserved. Indeed, this sphere of high conservation extends to ~12 Å from the substrate (Fig. 2b).

Although sequence comparison using BLAST (26) with sequences of proteins of known three-dimensional structure contained within the PDB showed no homologs, a search for structurally similar proteins using the program DALI (27) identified statistically significant matches for experimentally determined structures of N-dimethylarginine dimethylaminohydrolase (PDB code 1H70, (28)), arginine glycine amidinotransferase (1JDW; (29)), two arginine deiminases (1LXY (30) and 1RXX (31)) and two ribosome anti-association factors eIF6 (1G61 and 1G62 (32)). The first four proteins are enzymes belonging to the amidinotransferase superfamily (AT, as classified within SCOP (24)) with the structural similarity extending nearly throughout the entire protein. The substrates in the reaction catalyzed by arginine deiminases (Schema 2) are similar to that of AstB and both enzymes utilize water
molecule(s) to release ammonia (33). The eIF6 factors that are structurally similar to AstB are much smaller, with only ~230 residues, and contain five repeats of the basic ββαβ motif forming a minimal pentein propeller structure.

The amidinotransferase (AT) superfamily of enzymes (PF02274, PFAM database (34)) presently contains over 130 members from various bacterial species. The enzymes with known activities include glycine amidinotransferases (EC 2.1.4.1) involved in creatine biosynthesis, inosamine amidinotransferases (EC 2.1.4.2) involved in streptomycin biosynthesis, and arginine deiminases (EC 3.5.3.6) that convert arginine to citrulline. All these enzymes catalyze amidine group transfer or hydrolysis with the first step of the mechanism involving nucleophilic attack by a cysteine residue on the substrate. In addition to the cysteine, catalytic residues also include a histidine and an aspartate or a glutamate. Structure-based alignment of AstB with four other enzymes from the AT superfamily with known structures showed less than 10% sequence identity. Nevertheless, based on structural similarity, conservation of the catalytic residues, and the common type of reaction, it is now clear that the family of enzymes sharing sequence similarity with E. coli AstB also belongs to the AT superfamily.

Using the software FUGUE (35) Shirai and Mizuguchi recently constructed a model for the structure of E. coli AstB (9) and suggested that this protein is a member of the AT superfamily. They correctly predicted that the fold of AstB is a ββαβ propeller with five modules. However, their more detailed predictions were only partially correct. Indeed, while their assignment of Cys365 and His248 as part of the active site agrees with our structure, the third catalytic residue is Glu174 and not Asp173 as they predicted. Similarly, their prediction that Asp119 and Asp122 form hydrogen bonds to the guanidinium group of the substrate is
incorrect as these two residues are part of the surface loop and are more than 20 Å away from the substrate.

**Catalytic mechanism**

The structure of the Cys365Ser mutant complexed with N-succinylarginine and the wild type enzyme complexed with the N-succinylornithine product identified the location of the substrate binding site and the disposition of the substrate and the product relative to the catalytic residues. These sidechains include Cys365, His248 and Glu174. The comparison of AstB with other members of the AT superfamily (see above) shows a similar disposition of their catalytic residues with respect to the guanidinium moiety of the substrate (Fig. 4). In agreement with its predicted role as a nucleophile attacking the carbon of the guanidinium moiety mutation of Cys365 to serine severely compromised the enzyme’s activity. These observations indicate that AstB uses a catalytic mechanism similar to those of amidinotransferases and deiminases (30;31;36).

The reactions catalyzed by arginine deiminase (ADI) and succinylarginine dihydrolase differ in that ADI removes one NH₂ from the guanidinium moiety of the arginyl chain and replaces it by a carbonyl oxygen atom derived from a water molecule, whereas AstB carries the reaction further by removing the second NH₃ group, and releasing CO₂, leaving an ornithine sidechain in the product. Detailed catalytic mechanisms have been proposed for the arginine deaminases (30;31). Surprisingly, comparison of the sidechains in the vicinity of the guanidinium moiety of ADI (PDB code 1LXY (30)) and AstB shows nearly identical environments (Fig. 4) raising a question as to why the succinylarginine dihydrolase does not stop at converting N-succinylarginine to N-succinylcitrulline but carries the reaction further through a second hydrolysis event. The key catalytic residues Cys---His-Glu are conserved, as are the
Asp250 (Asp271 in ADI) and Arg212 (Arg232 in ADI) that form salt bridges to the guanidinium and carboxylate groups of the substrate, respectively. The only difference in the vicinity of the guanidinium group is the sidechain of residue 110, which in AstB is an asparagine whereas in ADI it is an aspartate (Asp161 in 1LXY) (Fig. 4). This sidechain in AstB forms one hydrogen bond to the NE atom of the guanidinium moiety whereas the equivalent residue, Asp161, in ADI forms two hydrogen bonds to NE and NH2. A review of available structures, reinforced by sequence alignment within the AT, ADI and AstB families, shows that in other enzymes that substitute a carbonyl oxygen for NH2 this sidechain is always an aspartate, whereas in the AstB dihydrolase family it is always an asparagine. We speculate therefore that the residue that is hydrogen bonded to the NE atom of the arginyl moiety of the substrate determines the outcome of the reaction. When this residue is an aspartate, as in ADI (Asp161), this sidechain forms two hydrogen bonds to the guanidinium moiety in the substrate and to the corresponding citrulline atoms in the product. A detailed reaction mechanism for ADI enzymes has been previously proposed (30;31).

We propose that AstB employs a mechanism similar to that of ADI but with two ADI-like hydrolytic reaction cycles to replace NH1 and NH2 by oxygens. In AstB the sidechain of Asn110 is hydrogen bonded to the arginyl NE (but not NH2). The first reaction cycle would convert N-succinylarginine to N-succinylcitrulline, replacing the NH1 atom with oxygen. This carbonyl oxygen would be in close proximity to the negatively charged acidic group of Asp250. A 180° rotation around the NE-CZ bond would relieve the unfavorable C=O … Asp250 contact, would bring this carbonyl oxygen atom into hydrogen bonding distance of the ND2 atom of Asn110 and would place the NH2 atom into the position previously occupied by NH1, thereby preparing the stage for the second hydrolysis cycle. The presence of an aspartate rather than
asparagine at this position in ADI enzymes would prevent such a rotation and lead to a release of the citrulline product.

The previously proposed mechanism for ADI suggests that the substrate’s NH1 forms a hydrogen bond to the histidine (His269 in ADI). The structure of the Cys365Ser mutant complexed with the $N$-succinylarginine substrate indicates that His248 does not form a hydrogen bond to the bound substrate’s guanidinium group as it is nearly perpendicular to the plane of the guanidinium. Rather, such a hydrogen bond would form only after the transfer of a proton to the arginyl CZ, when this atom attains sp$^3$ hybridization. The resulting tetrahedral coordination would direct this NH1 toward the NE of His248 forming a transient hydrogen bond that would aid in the release of ammonia, NH$_3$. Such a tetrahedral intermediate state has been observed in the structure of arginine deiminase (30) and shows the plausibility of a hydrogen bond between NH1 and the histidine (Fig. 4b).

Based on the proposed model for the reaction mechanism, we expect that succinylcitrulline would also be a good substrate for AstB. According to the proposed model the replacement of Asn110 by an aspartate should convert the dihydrolase activity of the wild-type enzyme to a deiminase activity in the mutant, leading to the formation of $N$-succinylcitrulline rather than $N$-succinylornithine. Similarly, the mutation of the corresponding Asp in arginine deiminase to an asparagine should convert the enzyme into an arginine dihydrolase. We are now testing these predictions experimentally.

**Conclusions**

The crystal structure revealed that AstB has the $\alpha/\beta$ propeller fold and belongs to the AT protein superfamily. The catalytic center is comprised of residues Cys365--His248-Glu174 positioned near the bottom of a long cavity extending from one side of the protein near the
propeller axis. Although the triad is superficially reminiscent of the catalytic triad of cysteine proteases, the disposition of the cysteine and histidine sidechains in these two classes of enzymes is quite different. In cysteine proteases Cys…His are within a hydrogen-bonding distance and the role of the histidine is to deprotonate the nucleophilic cysteine directly. In the AT superfamily cysteine and histidine are separated by more than 5Å and are positioned on opposite sides of the substrate with the histidine acting on the substrate or the transition state.

The substrate and product bind in a polar cleft of AstB shielded from the solvent by a 13 residue-long (Ala19-Arg32) loop unique to this family. This loop is disordered in the apo form of the enzyme suggesting that AstB exists in an open conformation in the absence of a bound ligand. Comparison of the substrate-free and substrate-bound structures shows that the flap closes over the entrance to the substrate binding tunnel and buries the N-succinylarginine. Subsequently, the flap must open to allow N-succinylornithine to depart.

Further studies of the enzymatic mechanism and conformational mobility of AstB, as well as other enzymes of the AST pathway may lead to the design of small molecule therapeutics that inhibit these enzymes.

Acknowledgements

We would like to thank Dr. T. Sulea for helpful discussion on the enzymatic mechanism, Mr. R. Larocque for help with cloning and Dr. S. Raymond for assistance in computational matters. Diffraction data for this study were measured at beamlines X8C and X25 of the National Synchrotron Light Source. Financial support for the NSLS comes principally from the Office of Biological and Environmental Research and of Basic Energy Sciences of the US Department of Energy, and the National Center for Research Resources of the National Institutes of Health. We thank Leonid Flaks (beamline X8C) and Michael Becker (beamline X25) for assistance in data
collection. This work was supported by Canadian Institutes for Health Research grant 200103GSP-90094-GMX-CFAA-19924 to MC and NSF grant MCB-0323931 to LR
References

1. Cunin, R., Glansdorff, N., Pierard, A., and Stalon, V. (1986) *Microbiol. Rev.* **50**, 314-352
2. Stalon, V., Vander, W. C., Momin, P., and Legrain, C. (1987) *J. Gen. Microbiol.* **133**, 2487-2495
3. Karp, P. D., Riley, M., Saier, M., Paulsen, I. T., Collado-Vides, J., Paley, S. M., Pellegrini-Toole, A., Bonavides, C., and Gama-Castro, S. (2002) *Nucleic Acids Res.* **30**, 56-58
4. Schneider, B. L., Kiupakis, A. K., and Reitzer, L. J. (1998) *J. Bacteriol.* **180**, 4278-4286
5. Itoh, Y. (1997) *J. Bacteriol.* **179**, 7280-7290
6. Fraley, C. D., Kim, J. H., McCann, M. P., and Matin, A. (1998) *J. Bacteriol.* **180**, 4287-4290
7. Kiupakis, A. K. and Reitzer, L. (2002) *J. Bacteriol.* **184**, 2940-2950
8. Reitzer, L. (2003) *Annu. Rev. Microbiol.* **57**, 155-176
9. Shirai, H. and Mizuguchi, K. (2003) *FEBS Lett.* **555**, 505-510
10. Shah, S. A., Shen, B. W., and Brünger, A. T. (1997) *Structure* **5**, 1067-1075
11. Liu, W., Peterson, P. E., Carter, R. J., Zhou, X., Langston, J. A., Fisher, A. J., and Toney, M. D. (2004) *Biochemistry* **43**, 10896-10905
12. Makarova, K. S. and Grishin, N. V. (1999) *J. Mol. Biol* **292**, 11-17
13. Hendrickson, W. A., Horton, J. R., and LeMaster, D. M. (1990) *EMBO J.* **9**, 1665-1672
14. Otwinowski, Z. and Minor, W. (1997) *Methods Enzymol.* **276**, 307-326
15. Pflugrath, J. W. (1999) *Acta Crystallogr. D. Biol Crystallogr.* **55**, 1718-1725
16. Matthews, B. W. (1968) *J. Mol. Biol.* **33**, 491-497
17. Terwilliger, T. C. and Berendzen, J. (1999) *Acta Crystallogr.* **D55**, 849-861
18. Terwilliger, T. C. (2000) *Acta Crystallogr.* **D56**, 965-972
19. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr.* D54, 905-921

20. Vagin, A. and Teplyakov, A. (1997) *J. Appl. Crystallogr.* 30, 1022-1025

21. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res.* 28, 235-242

22. Pearl, F. M., Lee, D., Bray, J. E., Sillitoe, I., Todd, A. E., Harrison, A. P., Thornton, J. M., and Orengo, C. A. (2000) *Nucleic Acids Res.* 28, 277-282

23. Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B., and Thornton, J. M. (1997) *Structure* 5, 1093-1108

24. Andreeva, A., Howorth, D., Brenner, S. E., Hubbard, T. J., Chothia, C., and Murzin, A. G. (2004) *Nucleic Acids Res.* 32, D226-D229

25. Lee, B. and Richards, F. M. (1971) *J. Mol. Biol.* 55, 379-400

26. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* 25, 3389-3402

27. Holm, L. and Sander, C. (1995) *Trends Biochem. Sci.* 20, 478-480

28. Murray-Rust, J., Leiper, J., McAlister, M., Phelan, J., Tilley, S., Santa, M. J., Vallance, P., and McDonald, N. (2001) *Nat. Struct. Biol.* 8, 679-683

29. Humm, A., Fritsche, E., Steinbacher, S., and Huber, R. (1997) *EMBO J.* 16, 3373-3385

30. Das, K., Butler, G. H., Kwiatkowski, V., Clark, A. D., Jr., Yadav, P., and Arnold, E. (2004) *Structure. (Camb.)* 12, 657-667

31. Galkin, A., Kulakova, L., Sarikaya, E., Lim, K., Howard, A., and Herzberg, O. (2004) *J Biol. Chem.* 279, 14001-14008
32. Groft, C. M., Beckmann, R., Sali, A., and Burley, S. K. (2000) Nat. Struct. Biol. 7, 1156-1164

33. Lu, X., Galkin, A., Herzberg, O., and Dunaway-Mariano, D. (2004) J. Am. Chem. Soc. 126, 5374-5375

34. Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L. L., Studholme, D. J., Yeats, C., and Eddy, S. R. (2004) Nucl. Acids Res. 32, D138-D141

35. Shi, J., Blundell, T. L., and Mizuguchi, K. (2001) J. Mol. Biol. 310, 243-257

36. Shirai, H., Blundell, T. L., and Mizuguchi, K. (2001) Trends Biochem. Sci. 26, 465-468
### Table 1 X-ray crystallographic data.

| Compound          | Se-met | Se-met | Se-met | Native<sup>2</sup> | Cys365 Ser<sup>2</sup> |
|-------------------|--------|--------|--------|---------------------|-------------------------|
| **Data collection** |        |        |        |                     |                         |
| **Space group**   | P1     | P1     | P2<sub>1</sub> | P2<sub>1</sub>;2<sub>1</sub>;2<sub>1</sub> | P2<sub>1</sub>;2<sub>1</sub>;2<sub>1</sub> |
| **Cell dimensions** |        |        |        |                     |                         |
| a (Å)             | 55.8   | 55.6   | 108.6  | 54.8                | 54.8                    |
| b (Å)             | 94.2   | 93.8   | 106.7  | 167.1               | 167.0                   |
| c (Å)             | 139.7  | 139.4  | 173.5  | 186.2               | 186.0                   |
| α (°)             | 75.4   | 104.7  | 90.0   | 90.0                | 90.0                    |
| β (°)             | 78.4   | 101.5  | 99.9   | 90.0                | 90.0                    |
| γ (°)             | 89.7   | 90.0   | 90.0   | 90.0                | 90.0                    |
| Z                 | 6      | 6      | 8      | 4                   | 4                       |
| **Resolution (Å)**| 30-2.25<sup>3</sup> | 30-2.3 | 30-3.0 | 30-1.95             | 30-1.7                  |
| **Wavelength (Å)**| 1.5418 | 0.9787 | 0.9796 | 0.9640              | 0.9799                  |
| **Unique reflections** | 136833 | 225632<sup>4</sup> | 227525<sup>4</sup> | 227839<sup>4</sup> | 146953                  |
| **Average redundancy** | 1.9    | 1.9    | 1.9    | 2.0                 | 4.0                     |
| **Completeness, %** | 92.0<sup>3</sup> | 96.5   | 96.6   | 97.4                | 98.1                    |
| **I/σ(I)**        | 8.8    | 16.0   | 14.5   | 15.5                | 10.2                    |

**Refinement Statistics<sup>5</sup>**

|                        |        |        |        |        |        |
|------------------------|--------|--------|--------|--------|--------|
| **R<sub>work</sub> (%)** | 21.3   | 26.0   | 21.7   | 20.2   |        |
| **R<sub>free</sub> (%)**   | 25.1   | 28.0   | 24.5   | 22.5   |        |
| **Number of atoms**     |        |        |        |        |        |
| Protein                | 20124  | 27464  | 13733  | 13723  |        |
| Ligand                 | -      | -      | 63     | 75     |        |
| Solvent                | 1402   | -      | 605    | 567    |        |
| Ions (K<sup>+</sup>)    | 6      | -      | 4      | 4      |        |
| **Ramachandran plot**  |        |        |        |        |        |
| Allowed                | 2256   | 3096   | 1546   | 1548   |        |
| Generously allowed     | 0      | 8      | 4      | 4      |        |
| Disallowed             | 6      | 0      | 0      | 0      |        |
| **R.M.S. deviation**   |        |        |        |        |        |
| Bonds (Å)              | 0.006  | 0.010  | 0.006  | 0.005  |        |
| Angles (°)             | 1.20   | 1.18   | 1.25   | 1.24   |        |

<sup>1</sup>R<sub>sym</sub> = (Σ | I<sub>obs</sub> - I<sub>avg</sub> |) / Σ I<sub>avg</sub>.

<sup>2</sup>Co-crystallised with N-succinylarginine.

<sup>3</sup>Overall completeness of data is 92% to 2.25 Å resolution, but all data to 1.9 Å resolution were used in the refinement. Overall completeness of all data is 63.9%.

<sup>4</sup>Friedel pairs were not merged.

<sup>5</sup>Statistics shown refer to the contents of the asymmetric unit.

<sup>6</sup>R<sub>work</sub> = (Σ | F<sub>obs</sub> - F<sub>calc</sub> |) / ΣF<sub>obs</sub>
Figure Legends

**Figure 1.** The AST pathway (after EcoCyc (3))

**Figure 2.** Ribbon representation of AstB.  
**a)** Stereo view of AstB approximately along the 5-fold pseudo-symmetry axis from the side opposite to the bound substrate (shown in van der Waals representation and colored blue). Each module is colored in succession: red, cyan, magenta, green and blue. The C-terminal $\alpha$-helical hairpin extension is colored yellow. The positions of N- and C-termini are marked;  
**b)** $\text{C}^{\alpha}$ backbone of AstB with the residues colored by conservation level. Magenta – highly conserved, marine – semi-conserved, gray – others. The ligand is shown in van der Waals representation;  
**c)** Ribbon drawing of the AstB dimer viewed along the pseudo 2-fold axis. This and subsequent figures were generated using the program PyMol (http://www.pymol.org/).

**Figure 3.** The final 3Fo-2Fc $\sigma_a$-weighted electron density map contoured at 1$\sigma$ around  
**a)** N-succinylornithine product in the substrate binding site of AstB;  
**b)** N-succinylarginine substrate complexed with the AstB Cys365Ser mutant. The ligand and the surrounding residues are drawn in a ball-and-stick representation. Nitrogen atoms are shown in blue and oxygen atoms in red. The hydrogen bond between the ligand and protein atoms are marked by green dashed lines.

**Figure 4.** Structural comparison of *E. coli* AstB with a representative amidinotransferase.  
**a)** The active site residues of the AstB Cys365Ser mutant with the *N*-succinylarginine substrate. In this figure Ser365 was replaced by the native Cys365 taken from the native structure. The oxygen atoms are red, nitrogen – blue, sulfur – yellow and carbon – white. The hydrogen bond between the ligand and protein atoms are marked by green dashed lines;
b) The active site of arginine deaminase (PDB codes 1LXY or 1S9R) with the reaction product in a similar orientation to that shown in a).
Figure 1.
Schema 1. Reaction catalyzed by \(N\)-succinylarginine dihydrolase.

Schema 2. Reaction catalyzed by arginine deiminase.
Crystal structure of N-succinylarginine dihydrolase, AstB, bound to substrate and product, an enzyme from the arginine catabolic pathway of Escherichia coli
Ante Tocilj, Joseph D. Schrag, Yunge Li, Barbara L. Schneider, Larry Reitzer, Allan Matte and Miroslaw Cygler

*J. Biol. Chem.* published online February 9, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M413833200

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