The 1.75 Å Crystal Structure of Acetyl-CoA Synthetase Bound to Adenosine-5′-propylphosphate and Coenzyme A†

Andrew M. Gulick,*,‡,§ Vincent J. Starai,|| Alexander R. Horswill,||,⊥ Kristen M. Homick,‡ and Jorge C. Escalante-Semerena||

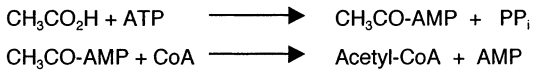
Hauptman-Woodward Medical Research Institute, Buffalo, New York 14203-1149, Department of Structural Biology, The State University of New York, Buffalo, New York 14203-1149, and Department of Bacteriology, University of Wisconsin, Madison, Wisconsin 53706-1567

Received November 12, 2002; Revised Manuscript Received January 14, 2003

ABSTRACT: Acetyl-coenzyme A synthetase catalyzes the two-step synthesis of acetyl-CoA from acetate, ATP, and CoA and belongs to a family of adenylate-forming enzymes that generate an acyl-AMP intermediate. This family includes other acyl- and aryl-CoA synthetases, firefly luciferase, and the adenylation domains of the modular nonribosomal peptide synthetases. We have determined the X-ray crystal structure of acetyl-CoA synthetase complexed with adenosine-5′-propylphosphate and CoA. The structure identifies the CoA binding pocket as well as a new conformation for members of this enzyme family in which the ∼110-residue C-terminal domain exhibits a large rotation compared to structures of peptide synthetase adenylation domains. This domain movement presents a new set of residues to the active site and removes a conserved lysine residue that was previously shown to be important for catalysis of the adenylation half-reaction. Comparison of our structure with kinetic and structural data of closely related enzymes suggests that the members of the adenylate-forming family of enzymes may adopt two different orientations to catalyze the two half-reactions. Additionally, we provide a structural explanation for the recently shown control of enzyme activity by acetylation of an active site lysine.

Acetyl-CoA synthetase (Acs,[1] acetate:CoA ligase, EC 6.2.1.1) is present in all organisms from bacteria to humans where it plays a key role in central metabolism. The activity of this enzyme is important for maintaining adequate levels of acetyl-CoA, a metabolite of many anabolic and catabolic processes. Acs is a member of a family of AMP-forming enzymes that catalyze two-step reactions in which an acyl-AMP intermediate is formed and pyrophosphate is released in the initial half-reaction (Scheme 1). In Acs, the acetyl-AMP intermediate is the substrate for the second half-reaction in which CoA displaces AMP yielding acetyl-CoA, the final product of the reaction (*1*). This AMP-forming family of enzymes includes firefly luciferase, acyl- and aryl-CoA synthetases, and the adenylation domains of the nonribosomal peptide synthetases (NRPSs) (*2*−*9*). The NRPS enzymes are involved in the biosynthesis of anticancer agents, such as bleomycin, and the peptide and *β*-lactam antibiotics (*10*−*12*) and often have multiple catalytic domains expressed as a single multidomain protein that can be several thousand residues in length (*9*, *13*). We have determined the structure of Acs bound to CoA and an analogue of the acyl-adenylate intermediate and propose that the structure reported here represents the enzyme conformation at the start of the second half-reaction.

Scheme 1



The sequences of members of the AMP-forming family of enzymes are 20−40% identical, and share certain conserved sequence motifs (*2*, *8*, *9*). Ten conserved regions of the NRPS adenylation domains have been described (*9*); seven of these are well-conserved in the acyl-CoA synthetases. These include a serine- and glycine-rich loop (termed region A3), a sequence containing a Thr−Glu dipeptide, where the glutamate is invariant (A5), a DX_6GXR sequence in which the glycine is invariant (A8), and a region containing a conserved Gly−Lys dipeptide (A10). While the A3 and A10 regions are involved in the adenylation half-reaction (*3*, *5*, *6*, *14*−*16*), the structure of Acs reported here explains the conservation of the A5 and A8 regions.

The crystal structures of firefly luciferase and PheA, the phenylalanine adenylation domain of gramicidin synthetase

† This work was supported in part by NIH Grant GM62203 (J.C.E.-S.) and by a grant from the John R. Oishei Foundation (A.M.G.). This work is based upon research conducted at the Cornell High Energy Synchrotron Source (CHESS), which is supported by the National Science Foundation under Award DMR 97-13424, using the Macromolecular Diffraction at CHESS (MacCHESS) facility, which is supported by Award RR-01646 from the National Institutes of Health, through its National Center for Research Resources.

* To whom correspondence should be addressed. Phone: (716) 856-9600, ext 327. Fax: (716) 852-6086. E-mail: gulick@hwi.buffalo.edu.

‡ Hauptman-Woodward Medical Research Institute.

§ The State University of New York.

|| University of Wisconsin.

⊥ Current address: Department of Chemistry, The Pennsylvania State University, University Park, PA 16802.

[1] Abbreviations: Acs, acetyl-CoA synthetase; NRPS, nonribosomal peptide synthetase; TCEP, tris(2-carboxyethyl)phosphine hydrochloride; MES, 2-(*N*-morpholino)ethanesulfonate; MIR, multiple isomorphous replacement; PheA, phenylalanine activating domain of the gramicidin synthetase A peptide synthetase; rms, root-mean-square.

10.1021/bi0271603 CCC: $25.00
Published on Web 02/19/2003

A (*4*, *17*), reveal that members of this family have a large N-terminal domain and an ~110-residue C-terminal domain. The active site is located at the interface between these two domains, and the orientation of the C-terminal domain is dependent on the contents of the active site. The C-terminal domain of the unliganded luciferase was shown to be in an open conformation, while the PheA structure, which was determined in the presence of phenylalanine and AMP, contains a C-terminal domain that is rotated by ~90° relative to that of luciferase. Very recently, the structure of DhbE, a free-standing adenylation domain from *Bacillus subtilis*, was determined (*18*). The conformation of DhbE in the absence of ligands and in the presence of the adenylate of the substrate 2,3-dihydroxybenzoate is very similar to the conformation observed in the PheA structure.

These structures have provided insight into the specificity determining elements of the acyl-substrate binding pocket (*19*, *20*). Additionally, the structural work has suggested that the conserved lysine residue of the A10 motif plays an important role in catalysis. The side chain amine of Lys^{517} of PheA interacts with a carboxylate oxygen of the phenylalanine substrate and with the 4′- and 5′-oxygens of the AMP. This residue was proposed to align properly the substrates in the active site and also maintain the orientation of the C-terminal domain. Mutagenesis results support this conclusion; replacement of this Lys residue in the valine-activating domain of surfactin synthetase I with a glutamine results in a 20-fold reduction in the rate of the adenylation half-reaction (*21*).

While the current structures provide insight into the adenylation half-reaction, there has not been much progress in determining the protein conformation or the residues involved in the thioester-forming step of the catalytic mechanism. We report here the 1.75 Å crystal structure of Acs from *Salmonella enterica* bound to both CoA and adenosine-5′-propylphosphate, an inhibitor of the related propionyl-CoA synthetase (*6*, *22*). The structure reveals a third conformation for adenylate-forming enzymes and suggests that the enzyme uses a large conformational change of the C-terminal domain to form the complete active site for the second half-reaction. We also report the structure of acetylated Acs determined at 2.3 Å resolution and provide an explanation for the inhibition of Acs by acetylation of Lys^{609} of the C-terminal domain (*23*).

METHODS

Protein Crystallization and Data Collection. Acs was prepared as described previously, and the acetylated form of the enzyme was prepared from a *cobB* deficient genetic background (*23*, *24*). Acs was crystallized at 14 °C by the hanging drop method. Protein (~10 mg/mL) was combined with 1 mM CoA (Sigma), 1 mM TCEP, and 1 mM adenosine-5′-propylphosphate. A droplet containing a 1:1 ratio of protein and reservoir solution was incubated over a reservoir of 10−14% PEG 8000, 10% ethylene glycol, and 50 mM MES (pH 6.5). After a period of 3−5 days, crystallization was induced by microseeding with a cat whisker. Crystals appeared within 2 days and grew to a maximum size of ~0.2 mm × 0.2 mm × 0.5 mm. Attempts to grow crystals in the absence of either adenosine-5′-propylphosphate or CoA were unsuccessful.

To determine the structure by MIR, native and heavy atom data sets were collected at −170 °C using a Rigaku RU-H3RHB rotating anode equipped with a Cu anode, Osmic Max-Flux confocal focusing mirrors (MSC), and an R-AXIS IV image plate system (Rigaku). Crystals were frozen by serially transferring them from the hanging drops to solutions containing 16, 20, and 24% ethylene glycol in the mother liquor of 12% PEG 8000, 50 mM MES (pH 6.5), and TCEP, CoA, and adenylate-5′-propylphosphate (1 mM each). For heavy atom soaks, the crystals were transferred to 50 μL of mother liquor containing the desired heavy atom. Data were processed with DENZO/SCALEPACK and written in the unmerged format (*25*).

Determination of the Structure of Acs. The structure of Acs was determined with SOLVE/RESOLVE using MIR with four heavy atom derivatives (*26*, *27*); anomalous differences were collected for all derivatives. The derivative data sets were scaled to the native data, and Patterson maps were visually inspected for significant peaks. The heavy atom positions were refined with SOLVE, and solvent flattening was performed with both DM (*28*, *29*) and RESOLVE. Surprisingly, the Lu, UO_2, and Pb derivatives all bound in the same position in the unit cell; differences in minor peaks were apparent, and all three derivatives, in addition to the Pt derivative, were included in the MIR phasing. The mean figure of merit for 31 901 reflections to 2.7 Å was 0.50 prior to solvent flattening and 0.74 after RESOLVE (*30*). The resultant maps were of sufficient quality to build. The automated fitting of RESOLVE was used to identify several stretches of residues (*31*). Once enough residues could be positioned, the structure of PheA (*4*) was manually rotated onto the new model and used as a guide.

High-Resolution Synchrotron Data Collection of Acs Crystals. Initial refinement with CNS (*32*) was performed after 527 residues were positioned in each subunit. The crystallographic *R*-factor dropped from 45.7 to 37.4% (R_{free} from 45.0 to 42.3%) using data to 2.7 Å. A high-resolution data set was then obtained at beamline F2 of the Cornell High Energy Synchrotron Source. Data were collected with an ADSC Quantum4 detector at a distance of 160 mm using a wavelength of 0.98074 Å. Two hundred degrees of data were collected using 0.5° oscillations and 25 s exposures.

The final model contains residues 5−647 of each Acs molecule with one disordered region (residues 623−631 of chain A and 623−627 of chain B) and one molecule of adenosine-5′-propylphosphate and coenzyme A bound to each chain. The two molecules overlap with an rms deviation of 0.15 Å for 633 Cα atoms. There are 846 solvent molecules and five molecules of ethylene glycol included in the final model. All of the residues lie in the most favored (90.5%) or allowed (9.3%) regions of the Ramachandran plot with the exception of Thr^{438} which has ϕ and ψ angles of 70° and −118°, respectively. The electron density is unambiguous for this residue in both subunits, and it is unclear why this residue adopts this unfavorable conformation.

The *acs* gene used in this study was cloned prior to the completion of the genomic sequencing of *S. enterica* (*33*). The cloned gene contained one mutation, located at the codon for Arg^{174}, such that our expression clone encoded a cysteine residue at this position. The structural work presented here was done with the R174C mutant protein. We have since cloned and expressed the wild-type protein and have col-

Table 1: Crystallographic Data for Native and Heavy Atom Derivative Data Sets[a]

	native[b]	CHESS[b]	acetylated	K_2PtCl_4	$LuCl_3$	$UO_2(OAc)_2$	$Pb(OAc)_2$
resolution (Å)	2.3 (2.38−2.3)	1.75 (1.81−1.75)	2.3 (2.38−2.3)	2.3 (2.38−2.3)	2.7 (2.8−2.7)	2.7 (2.8−2.7)	2.7 (2.8−2.7)
R_{merge}[c] (%)	2.7 (5.9)	7.1 (41.5)	10.5 (38.4)	6.3 (21.2)	7.1 (17.1)	10.3 (27.4)	10.0 (26.2)
completeness (%)	89.8 (61.4)	93.5 (75.1)	95.9 (78.3)	98.7 (88.2)	99.7 (97.9)	89.3 (77.4)	99.8 (98.4)
I/σ	39.5 (16.6)	14.2 (2.0)	7.9 (1.7)	17.8 (4.1)	11.3 (3.8)	7.0 (2.0)	7.6 (2.4)
no. of observations	129201 (7663)	402947 (5608)	177381 (9516)	293936 (17479)	118411 (9564)	70880 (5608)	135732 (10770)
no. of reflections	48517 (3241)	114217 (2926)	51552 (4037)	53262 (4623)	33054 (3349)	29741 (2926)	33330 (3394)
soak condition[d]				1 mM, 42 h	2 mM, 91 h	4 mM, 20 h	10 mM, 20 h
R_{iso}[e] (%)				10.7	18.7	22.3	19.6

[a] Values for the highest-resolution shell are given in parentheses. [b] The native data set was obtained using an in-house X-ray source and was used for MIR phasing and initial model building. The high-resolution CHESS data set was collected on a crystal grown from unacetylated protein and was used for final refinement. [c] $R_{merge} = \sum(|I_{hi} - I_h|)/\sum I_{hi}$, where I_{hi} and I_h are individual and mean intensities of all equivalent reflections, respectively. [d] Concentration of heavy atom and length of time for soak. [e] $R_{iso} = \sum(|F_h - F_p|)/\sum F_p$, where F_p and F_h are the native and derivative structure factors, respectively.

lected a moderate-resolution (2.5 Å) data set on crystals of the wild-type protein. The density for Arg^{174} is unambiguous, and there are no other differences between the wild-type and R174C structures.

Structural Determination of the Acetylated Form of Acs. The acetylated Acs crystallized under conditions similar to those of the unacetylated protein. Data were collected as described above for the low-resolution native data sets. Crystals of the acetylated protein diffracted to 2.3 Å. The structure was determined by difference Fourier methods. All water molecules, ligands, and heteroatoms were removed from the final unacetylated Acs structure prior to refinement against the new data set. While the electron density for the adenosine-5′-propylphosphate and nucleotide portion of CoA were unambiguous, the orientation of the panthetheine group of CoA was less well-ordered and the position of the 5′-diphosphate is changed slightly when compared to that of the unacetylated protein. Additional density was present throughout the pantetheine channel; however, the density was not of sufficient quality to warrant inclusion in the refined model. We believe this represents the limited resolution of the acetylated Acs crystal structure and not a change in the binding pocket for CoA; the acetylated enzyme is able to catalyze the thioester-forming reaction and therefore binds correctly to CoA (*23*). The moderate-resolution data set for the wild-type protein shows similar features in the electron density for the CoA molecule.

Deposition of Model Coordinates. The atomic coordinates and structure factors for both the unacetylated and acetylated Acs structures have been deposited with the Protein Data Bank (1NNM and 1NNN, respectively).

RESULTS

Determination of the Structure of Acs. The Acs proteins were purified (*23*) and crystallized in the monoclinic space group $P2_1$ with the following dimensions: $a = 60.0$ Å, $b = 143.0$ Å, $c = 71.7$ Å, and $\beta = 91.32°$. The structure of Acs was determined by MIR using four heavy atom derivatives (Table 1). There are two subunits of Acs in the asymmetric unit. The final model contains residues His^5−Ala^{647}; the four N- and C-terminal residues are disordered in each molecule. Additionally, there is one disordered loop near the C-termini, positioned at residues 623−631. Final refinement statistics against high-resolution synchrotron data are shown in Table 2. Unbiased electron density maps for the ligands prior to their inclusion in refinement are shown in Figure 1.

Table 2: Refinement Statistics[a]

	unacetylated	acetylated
resolution range (Å)	30.0−1.75	30.0−2.3
R_{cryst} (%) (overall/highest-resolution shell)[b]	18.4/28.1	19.9/25.0
R_{free} (%) (overall/highest-resolution shell)[b]	21.1/30.9	23.9/28.3
Wilson *B*-factor (Å^2)	18.9	23.0
average *B*-factor (Å^2)		
overall	21.1 (10835)[c]	21.7 (10384)[c]
protein	20.2 (9814)[c]	21.6 (9833)[c]
solvent	30.1 (849)[c]	23.1 (377)[c]
ligands	30.0 (148)[c]	30.0 (114)[c]
rms deviation for bond lengths (Å) and angles (deg)	0.005, 1.3	0.007, 1.2

[a] Final refinement statistics for the unacetylated Acs model against synchrotron data to 1.75 Å, or the acetylated form with data to 2.3 Å. The rms deviation for all Cα atoms between the two structures is 0.2 Å, and no differences in protein structure are observed. [b] The highest-resolution shell is from 1.86 to 1.75 Å for the unacetylated protein and from 2.44 to 2.30 Å for the acetylated protein. [c] Total number of atoms used in the calculation.

The three-dimensional structure of Acs shows that the molecule is composed of two domains (Figure 2). The large N-terminal domain is composed of the first 517 residues. This globular domain is ∼40 Å × 50 Å × 60 Å. Like PheA and firefly luciferase, the N-terminal domain is formed primarily by two mostly parallel eight-stranded β-sheets (Figure 3). Adopting the nomenclature used for PheA, we refer to these two strands as sheets A and B; sheet A contains one more strand (strand A4a) than sheet A of PheA (Figure 2). The N-terminal domain also contains a distorted four-stranded antiparallel β-sheet (sheet C), which continues to Asp^{517} of the A8 motif. This residue forms the hinge for the conformational change described below, containing ϕ and ψ angles of −103° and −169° in Acs and −64° and −32° at the homologous residue (Asp^{430}) in PheA, respectively (Figure 2B).

The final ∼130 residues of Acs form the C-terminal domain. This domain begins with a short loop containing two antiparallel strands at residues 519−522 and 525−529. The turn between these two strands contains the conserved GXR sequence of the A8 motif and is an important feature of the Acs active site. The remainder of the C-terminal domain forms a three-stranded β-sheet (green in Figure 2A) that is surrounded on both sides by two α-helices. A comparison of the C-terminal domain of Acs to the C-terminal domain seen in the PheA structures demonstrates

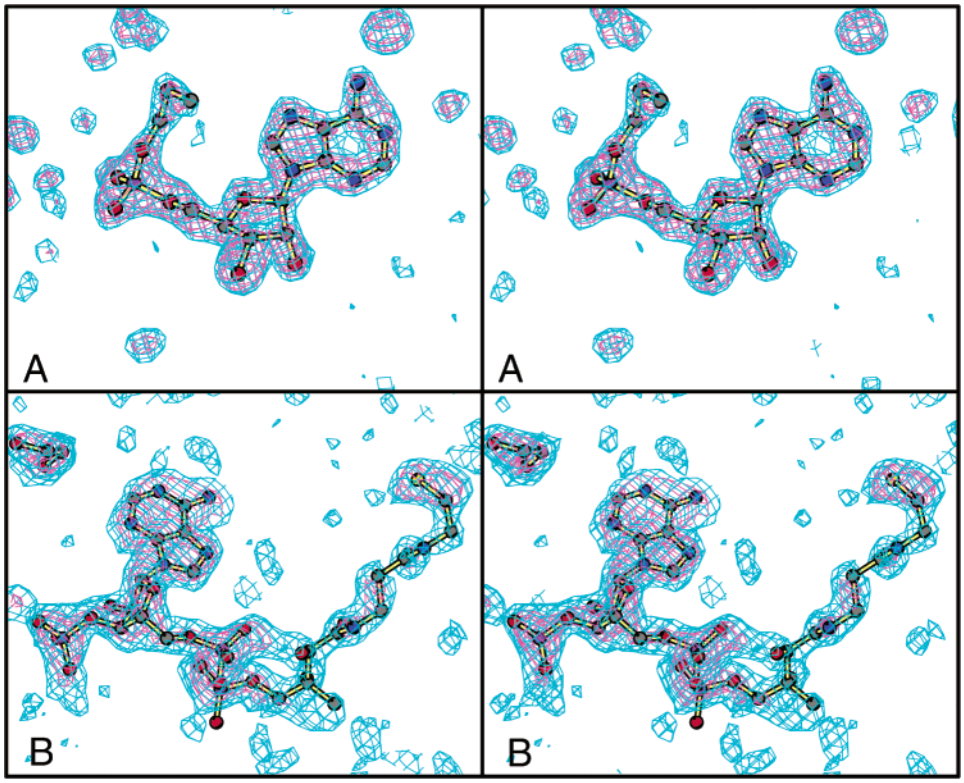


FIGURE 1: Electron density of the ligands for the structure of unacetylated Acs. (A) Electron density of the adenosine-5′-propylphosphate calculated with coefficients of the form $F_o - F_c$ and phases determined prior to inclusion of the inhibitor in the structure, plotted at 3σ (blue) and 4σ (magenta). (B) Electron density for the CoA molecule calculated in a similar manner and plotted at 2.4σ (blue) and 4σ (magenta) prior to the addition of CoA to the refinement model. The density indicates that the inhibitor and the nucleotide portion of CoA were extremely well ordered and the pantetheinate group, while less well ordered, was still of sufficient quality to be included in the final model. Panel B also contains a molecule of the cryoprotectant ethylene glycol that was present near the CoA molecule. This figure was produced with MOLSCRIPT, BOBSCRIPT, and RASTER3D (*40*−*42*).

that while the secondary structural elements of the two proteins are nearly identical, the domain in Acs is rotated by ~140° relative to the orientation seen in PheA (Figure 2C). As discussed below, we believe this conformational difference is a consequence of the binding of CoA and that the conformation observed in our structure is responsible for catalysis of the second half-reaction.

Active Site of Acs. Crystals of Acs were grown in the presence of adenosine-5′-propylphosphate and CoA. Adenosine-5′-propylphosphate is a competitive inhibitor against ATP or propionyl-AMP for the reaction catalyzed by PrpE, a propionyl-CoA synthetase, and acts as a mimic for the acyl-adenylate intermediate (*6*, *22*). The inhibitor differs from the true Acs intermediate by the absence of the carbonyl oxygen and an increase in the length of the acyl chain by one methylene group (Figure 4). In vivo, Acs can catalyze the formation of propionyl-CoA from propionate, ATP, and CoA, demonstrating that the active site can accommodate the larger acyl substrate (*34*). Additionally, Acs can react with propionate with a specific activity that is approximately 5-fold lower than the specific activity with acetate (data not shown).

Electron density was apparent for the adenosine-5′-propylphosphate in the experimental map, although the inhibitor was not included in the refinement until later stages. The location of the CoA binding site became apparent as refinement progressed against the high-resolution data set (Figure 1B). The structure clearly identifies two distinct binding sites for the adenine groups of AMP and CoA.

The binding site for adenosine-5′-propylphosphate is almost completely buried (Figure 4). The position of the AMP moiety is very similar to that observed in the structure of PheA bound to AMP and phenylalanine. The adenine ring lies in a hydrophobic pocket generated by Trp^{413} and Ile^{512}. The ring also stacks against the main chain peptide bonds of Gly^{387}, Glu^{388}, and Pro^{389}. The 2′-hydroxyl of the AMP moiety interacts with Asp^{500}, while the 3′-hydroxyl interacts with Gln^{415} and Arg^{515}. These interactions with the 3′-hydroxyl were not observed in the PheA structure as the homologue of Arg^{515} is positioned further from the 3′-hydroxyl. In Acs, the side chain of Arg^{515} is positioned by Asp^{517}. One oxygen of the inhibitor phosphate interacts with Arg^{526} and with the Thr^{416} side chain hydroxyl, while the remaining oxygen interacts with solvent molecules. Arg^{526} also forms a salt bridge across the active site with Glu^{417}, a residue belonging to the conserved A5 motif (*9*). Mutations in the A5 Glu residue in 4-chlorobenzoyl-CoA ligase result in a 50-fold reduction in both overall activity and activity of the adenylation half-reaction (*3*). Interestingly, this enzyme does not contain a homologue of Arg^{526}, containing an Asn residue at this position, demonstrating that the interaction between Arg^{526} and Glu^{417} seen in the Acs structure is not essential for this family of enzymes.

The bridging oxygen to the propyl group of the inhibitor makes no contacts with protein. The propyl group is bound in roughly the same position as the phenylalanine substrate in the PheA structure; the pocket is formed by Val^{310}, Val^{386}, and Trp^{414}. That the smaller substrate contacts only three

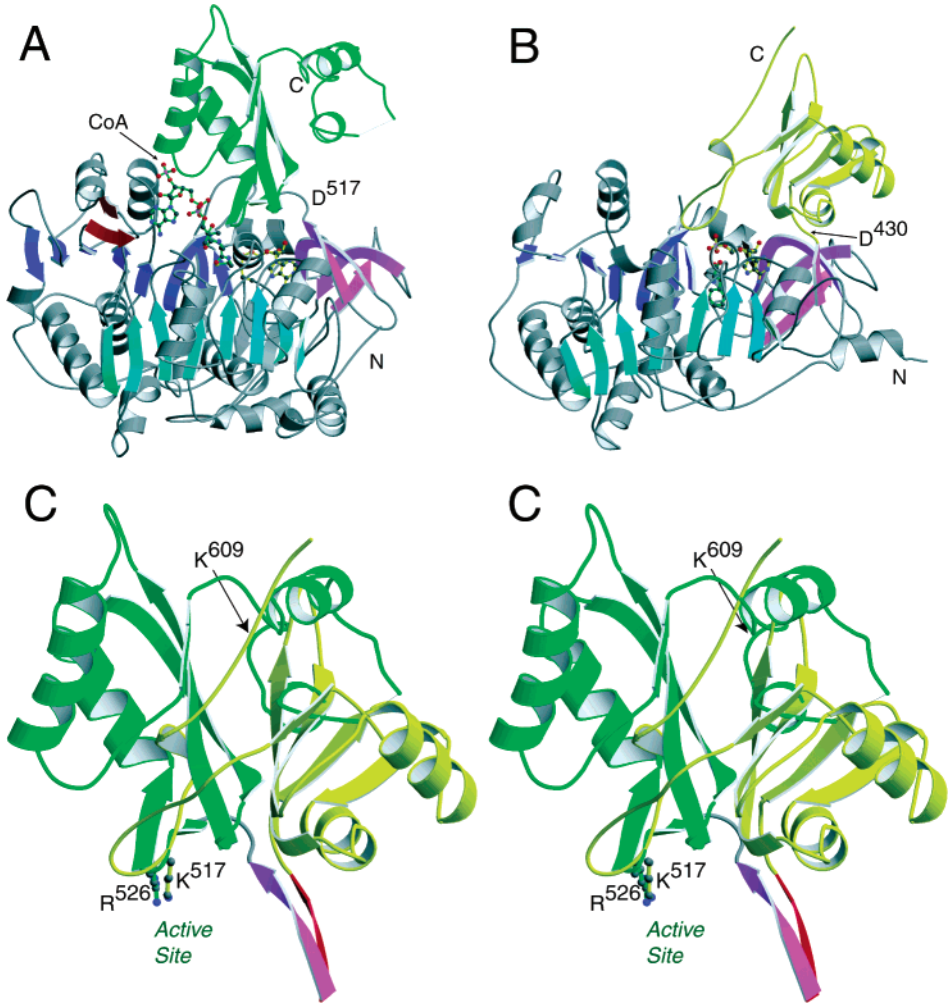


FIGURE 2: Comparison of Acs and the PheA NRPS adenylation domain. (A) Ribbon representation of Acs bound to CoA and adenosine-5′-propylphosphate. The β-sheets that form the N-terminal domain are shown in blue, cyan, and magenta for sheets A−C, respectively. The binding sites for the two ligands are indicated by the ball-and-stick figures. The C-terminal domain is shown in green, while the two strands that form the 11-residue insertion that contains Arg^{191} and Arg^{194} are shown in red. (B) Ribbon representation of the PheA adenylation domain (*4*) bound to phenylalanine and AMP. Features similar to those in panel A are indicated; however, the C-terminal domain in the adenylate-forming conformation is shown in yellow. (C) Models of Acs and PheA were superimposed using the N-terminal domains as the overlapping regions demonstrating the relative orientation of the C-terminal domain. The viewpoint is the same as that for panels A and B. The two strands in magenta (Acs) and red (PheA) illustrate the final strands of the N-terminal domains. The C-terminal domains contain the same overall fold but are rotated by ~140° relative to each other. The side chains of Arg^{526} of Acs and Lys^{517} of PheA are represented with ball-and-stick diagrams and occupy very similar positions relative to the N-terminal domain and the active site. The position of Acs residue Lys^{609}, the site of acetylation and the homologue of PheA Lys^{517}, is indicated with an arrow. This figure was produced with MOLSCRIPT, BOBSCRIPT, and RASTER3D (*40*−*42*).

residues from the binding site may explain why smaller NRPS substrates such as alanine have less well-conserved binding pockets than larger substrates which would require more interactions with the enzyme (*20*).

The nucleotide portion of CoA is bound on the surface of the protein, and the phosphopantetheine moiety is directed into the active site (Figure 4). The adenine ring of CoA fits into a binding pocket where it stacks against the guanidinyl group of Arg^{191} and the peptide bond formed between Gly^{164} and Gly^{165}. This hydrophobic pocket is also partly formed by Phe^{163} and Ile^{196}. While there are no direct interactions between the ring nitrogens and the protein, there is a water-mediated interaction between the C6 amino group and the side chain of Asp^{306}. The 5′-diphosphate of CoA interacts with the side chains of Arg^{191}. The 3′-phosphate of CoA interacts with the side chain of Arg^{584} from the C-terminal domain and with Arg^{194}. The side chain of Arg^{194} adopts two conformations and may interact with the 3′-phosphate from the CoA bound to a symmetry-related Acs molecule. Although two bacterial Acs enzymes have been reported to be dimers (*35*, *36*), *S. enterica* Acs elutes as a monomer during gel filtration chromatography in the absence of ligands or in the presence of 1 mM acetate, AMP, and CoA. Human Acs has also been reported to be a monomer (*37*), and we believe the Arg^{194} interactions seen in the crystal structure of Acs are a result of crystal packing and do not represent dimerization of the protein. Arg^{191} and Arg^{194} are located on an 11-residue insertion in the Acs structure relative to PheA that form a two-stranded antiparallel turn (Figure 2A).

The pantetheine moiety passes through a channel between the N- and C-terminal domains toward the AMP binding site and makes few direct interactions with the protein. This portion of the coenzyme molecule is less well-ordered than the nucleotide moiety; however, the electron density is of sufficient quality to include it in the final model (Figure 1). The two amines of the pantetheinyl group form a distorted

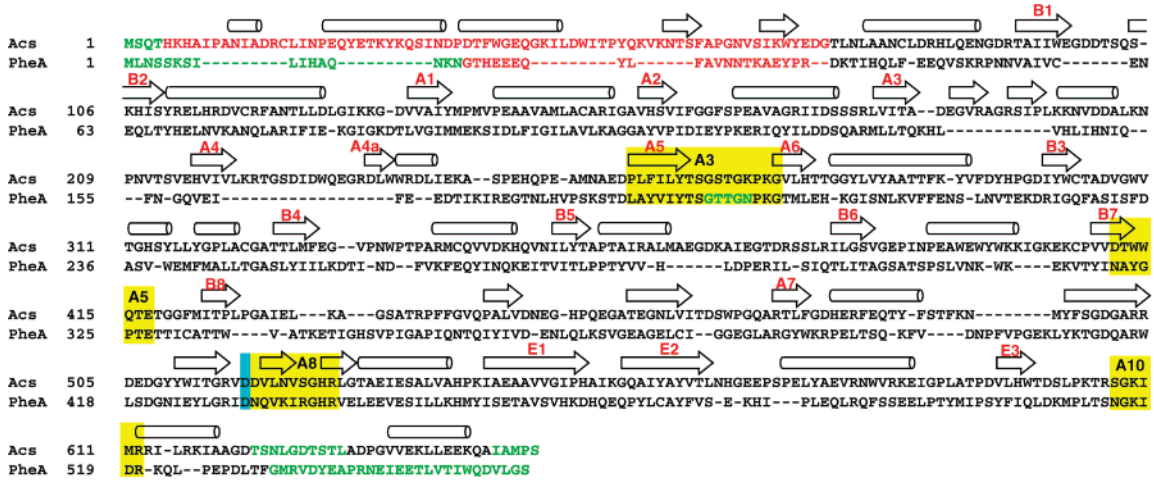


FIGURE 3: Amino acid alignment of Acs and PheA. The sequences of the N- and C-terminal domains of Acs and PheA were aligned using the structure-based alignment algorithm of ALIGN (*43*). For the N-terminal domain, residues 5−517 of Acs were aligned with residues 17−430 of PheA. The C-terminal alignment used residues 518−647 of Acs and 431−530 of PheA. Residues that are disordered are shown in green. The N-terminal sequence shown in red represents the structurally divergent N-termini which do not contain similar secondary structural elements; these residues were aligned using the algorithm of Meyer and Miller (*44*). The conserved motifs discussed in the text are highlighted with the yellow boxes. Asp^{517} of Acs which serves as the hinge for the C-terminal domain is highlighted in blue. The secondary structural elements of Acs are represented with arrows and tubes for β-sheets and α-helices, respectively. The strands that compose the two main structural sheets, A and B, and the C-terminal sheet, E, are labeled.

β-sheet interaction with the backbone carbonyls of Ser^{523} and Gly^{524}, the glycine residue of the A8 motif which is completely conserved in the AMP-forming enzymes.

The sulfur atom of CoA is 5.4 Å from C1 of the propyl group of the inhibitor. The sulfur is oriented in a nonproductive position, most likely because of the presence of the inhibitor instead of the true adenylate intermediate (Figure 4). A simple torsional rotation of the penultimate bond of the CoA pantetheine group that moves only the sulfur atom will reduce this interatomic distance to 3.0 Å.

Crystal Structure of the Acetylated Form of Acs. As reported recently, Acs activity is controlled by its acetylation state (*23*, *24*). Data show that acetylation of Lys^{609} renders the Acs enzyme inactive. To determine the structural effects of the acetylation, we determined the structure of the acetylated protein (Table 2). No evidence for any conformational changes induced by the acetylation of residue Lys^{609} was revealed by this structure. The rms deviation for Cα positions of all residues of the acetylated and unacetylated structures is 0.2 Å. The pantetheine moiety of CoA is less well-ordered in the structure of the acetylated enzyme, and CoA has been modeled only through the nucleotide and phosphates. This is likely due to the data for this structure being lower-resolution data than those for the unacetylated form. In the crystal structures of both the unacetylated and acetylated enzymes, the side chain of Lys^{609}, the site of acetylation, is disordered and the residue is modeled as an alanine. Examination of the electron density of all ordered lysines in the structure of the acetylated form of Acs reveals no evidence for acetylation at any other site.

DISCUSSION

We describe here the results of our structural investigation of Acs, a 652-residue enzyme that catalyzes the two-step synthesis of acetyl-CoA. Our results identify the overall fold of the enzyme, which is similar to other members of the adenylate-forming family of enzymes, and the binding site for CoA. While the secondary structural elements of the N- and C-terminal domains of this enzyme are very similar to PheA and DhbE (*4*, *18*), the orientation of the C-terminal domain seen in Acs is rotated by ∼140° relative to the conformation seen in the structures of PheA and DhbE. We propose that members of the adenylate-forming family of enzymes are likely to adopt two different conformations to catalyze the two half-reactions. The conformation seen in the crystal structures of PheA and DhbE is used to catalyze the adenylation half-reaction, while the conformation of Acs presented here bound to adenosine-5′-propylphosphate and CoA is used to catalyze the thioester-forming half-reaction.

The evidence that supports this hypothesis is as follows. (1) The structures of Acs, PheA, and DhbE are all very similar at the level of secondary structure. The enzymes catalyze very similar two-step reactions, and despite moderate sequence homology throughout the entire protein sequence, numerous well-conserved signature sequences exist for members of this family (*19*, *20*). Interestingly, two of these conserved regions, A8 and A10, are located on opposite sides of the ∼110-residue C-terminal domain. Our model provides an explanation for the conservation of two regions of the enzymes that are located 29 Å from each other. (2) In the structure of PheA, the A10 motif lysine residue, Lys^{517}, is part of the active site, contacting both the AMP ribose ring oxygen and a carboxylate oxygen of the phenylalanine substrate (*4*). The equivalent lysine in Acs, Lys^{609}, is located ∼27 Å from the 4′-oxygen of adenosine-5′-propylphosphate. Mutation of the A10 lysine residue in firefly luciferase (*14*), or PrpE (*6*), a propionyl-CoA synthetase whose sequence is 37.4% homologous with that of Acs, dramatically reduces the catalytic activity for the adenylation half-reaction (by 4 orders of magnitude in the case of PrpE) while having very little effect on activity for the second half-reaction. (3) A similar effect is seen in Acs where acetylation of the Lys^{609} residue of Acs also blocks the ability of Acs to catalyze the first half-reaction while having no effect on the catalysis of the thioester-forming reaction (*23*). (4) Despite numerous attempts, we have been unable to grow the current mono-

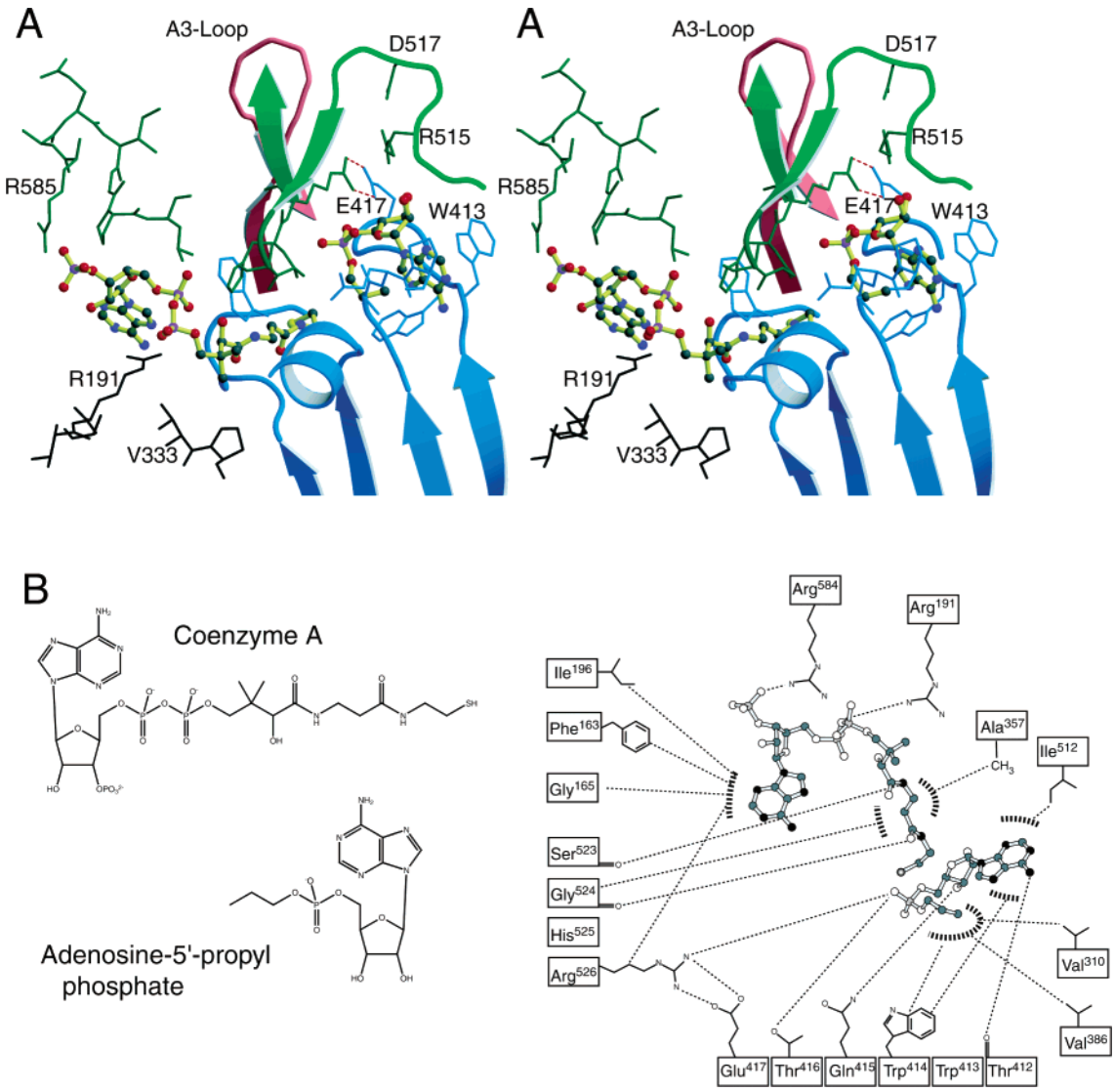


FIGURE 4: Active site interactions of Acs with CoA and adenosine-5′-propylphosphate. (A) Stereorepresentation of the binding pocket for CoA and the inhibitor demonstrating the surface binding of the nucleotide portion of CoA. The orientation shown is similar to the view in Figure 2A with four strands from sheet B shown in blue. The two ligands are represented with ball-and-stick models with yellow bonds. The phosphopantetheine moiety of CoA is directed into the adenylate binding pocket. The salt bridge located between Arg^{526} and Thr^{418} is shown as dashed red lines. The glycine-rich A3 motif is shown in pink. The Arg^{526} side chain lies between the inhibitor phosphate and this loop. (B) Interactions between protein and substrates are shown schematically. The two ligand molecules are shown as ball-and-stick models in the appropriate stereoconfigurations and orientation. The protein residues that interact with the two ligands are shown, and interactions between different functional groups are indicated with the dashed lines. Hydrophobic interactions are represented by the broad hashed curves. This figure was produced with MOLSCRIPT, BOBSCRIPT, and RASTER3D (*40*−*42*).

clinic crystal form of Acs in the absence of CoA, suggesting that CoA is required for Acs to adopt the conformation that is observed in the current crystal structure.

These experiments suggest a role of the A10 lysine residue in the catalysis of the first half-reaction only, an observation that is consistent with our model for members of the adenylate-forming family. We believe that the acyl-adenylate intermediate is formed in Acs by a conformation very similar to that seen in the crystal structures of PheA and DhbE. Lys^{609} plays an important role in this step of the reaction, perhaps properly aligning the acetate by an interaction similar to that seen in the PheA structure. After formation of the adenylate, the binding of CoA would then cause a rotation of the C-terminal domain of the enzyme by ~140° to form the orientation observed in the crystal structure of Acs. This structure appropriately positions the CoA thiol for nucleophilic attack on the acetyl group. This conformation also positions Lys^{609} more than 25 Å from the active site and explains why acetylation of this residue in Acs, or mutation of this residue in homologous enzymes (*6*), has no effect on catalysis of the second half-reaction. While absolute proof of this hypothesis awaits the determination of multiple structures of a single enzyme in the different conformational states, we believe that the cumulative results of the structural and kinetic experiments on several very closely related enzymes provide ample support to this model.

Our results are consistent with kinetic data for this family of enzymes that suggest these enzymes undergo a Bi Uni Uni Bi ping-pong reaction mechanism in which ATP and acetate react to displace PP_i and form the acyl-adenylate intermediate (*6*, *22*, *38*). CoA then binds to the enzyme to allow the formation of the thioester in the second half-reation. Our structure represents the conformation of the enzyme after the binding of CoA.

One of the most highly conserved regions of the AMP-forming enzymes is a glycine-rich A3 loop (*9*). This region has a consensus sequence of YTSG(S/T)TGxPKG (Figure 3). Residues in this loop have been studied in many family members by mutagenesis and kinetic experiments that have suggested this loop may play a role in positioning the triphosphate portion of ATP (*3*, *6*, *16*). This loop is present in both molecules of Acs as a loop connecting strands A5 and A6 from β-sheet A (Figure 4). This loop is appropriately positioned to bind to the displaced pyrophosphate in the adenylation reaction. In the adenylation half-reaction, the in-line attack of the acetate on the α-phosphate would position the leaving group composed of the β- and γ-phosphates very close to this A3 motif. In the current structure, the side chain of Arg^{526} is positioned directly between the phosphate of adenosine-5′-propylphosphate and this conserved loop, emphasizing that the conformation we observe is likely to occur after the completion of the adenylation half-reaction. Although the A3 loop is disordered in both the luciferase (*17*) and PheA structures (*4*), it is present in the DhbE structure (*18*). The loop in Acs is 4−7 Å further from the adenylate phosphate than the loop seen in DhbE. This appears to result from the presence of Arg^{526}.

The model proposed here suggests that it is likely that the adenylation domains of the NRPS enzymes will adopt multiple conformations in the activation of the amino acid substrates to form thioesters with the covalently bound pantetheine cofactors. This is supported by the changes in proteolytic susceptibility seen in the adenylation domain of tyrocidin synthetase I upon addition of substrates (*39*). As some NRPS adenylation domains are located within a large multimodular protein and contain additional protein domains bound to both their N- and C-termini, the question of how these enzymes are able to allow these conformational changes to an internal domain remains.

ACKNOWLEDGMENT

We thank Dr. Walter Pangborn for advice and helpful suggestions during X-ray data collection. A.M.G. also thanks Dr. Ivan Rayment for generous advice and encouragement.

REFERENCES

1. Metzler, D. E. (2001) *Biochemistry: The Chemical Reactions of Living Cells*, 2nd ed., Vol. 1, Harcourt/Academic Press, New York.
2. Babbitt, P. C., Kenyon, G. L., Martin, B. M., Charest, H., Slyvestre, M., Scholten, J. D., Chang, K. H., Liang, P. H., and Dunaway-Mariano, D. (1992) *Biochemistry 31*, 5594−5604.
3. Chang, K. H., Xiang, H., and Dunaway-Mariano, D. (1997) *Biochemistry 36*, 15650−15659.
4. Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997) *EMBO J. 16*, 4174−4183.
5. Dieckmann, R., Pavela-Vrancic, M., Pfeifer, E., Von Dhören, H., and Kleinkauf, H. (1997) *Eur. J. Biochem. 247*, 1074−1082.
6. Horswill, A. R., and Escalante-Semerena, J. C. (2002) *Biochemistry 41*, 2379−2387.
7. Keating, T. A., and Walsh, C. T. (1999) *Curr. Opin. Chem. Biol. 3*, 598−606.
8. Kleinkauf, H., and Von Dohren, H. (1996) *Eur. J. Biochem. 236*, 335−351.
9. Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997) *Chem. Rev. 97*, 2651−2674.
10. Martin, J. F. (2000) *J. Antibiot. 53*, 1008−1021.
11. Shen, B., Du, L., Sanchez, C., Edwards, D. J., Chen, M., and Murrell, J. M. (2001) *J. Ind. Microbiol. Biotechnol. 27*, 378−385.
12. Smith, D. J., Earl, A. J., and Turner, G. (1990) *EMBO J. 9*, 2743−2750.
13. Vater, J., Stein, T., Vollenbroich, D., Kruft, V., Wittmann-Liebold, B., Franke, P., Liu, L., and Zuber, P. (1997) *J. Protein Chem. 16*, 557−564.
14. Branchini, B. R., Murtiashaw, M. H., Magyar, R. A., and Anderson, S. M. (2000) *Biochemistry 39*, 5433−5440.
15. Gocht, M., and Marahiel, M. A. (1994) *J. Bacteriol. 176*, 2654−2662.
16. Stuible, H., Buttner, D., Ehlting, J., Hahlbrock, K., and Kombrink, E. (2000) *FEBS Lett. 467*, 117−122.
17. Conti, E., Franks, N. P., and Brick, P. (1996) *Structure 4*, 287−298.
18. May, J. J., Kessler, N., Marahiel, M. A., and Stubbs, M. T. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 12120−12125.
19. Challis, G. L., Ravel, J., and Townsend, C. A. (2000) *Chem. Biol. 7*, 211−224.
20. Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999) *Chem. Biol. 6*, 493−505.
21. Hamoen, L. W., Eshuis, H., Jongbloed, J., Venema, G., and van Sinderen, D. (1995) *Mol. Microbiol. 15*, 55−63.
22. Grayson, N. A., and Westkaemper, R. B. (1988) *Life Sci. 43*, 437−444.
23. Starai, V. J., Celic, I., Cole, R. N., Boeke, J. D., and Escalante-Semerena, J. C. (2002) *Science 298*, 2390−2392.
24. Starai, V. J., Takahashi, H., Boeke, J. D., and Escalante-Semerena, J. C. (2002) *Genetics* (in press).
25. Otwinowski, Z. M., and Minor W. (1997) *Methods Enzymol. 276*, 307−326.
26. Terwilliger, T. C., and Berendzen, J. (1996) *Acta Crystallogr. D52*, 749−757.
27. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. D55*, 849−861.
28. Collaborative Computational Project 4 (1994) *Acta Crystallogr. D50*, 760−763.
29. Cowtan, K. D., and Main, P. (1993) *Acta Crystallogr. D49*, 148−157.
30. Terwilliger, T. C. (2000) *Acta Crystallogr. D56*, 965−972.
31. Terwilliger, T. C. (2001) *Acta Crystallogr. D57*, 1755−1762.
32. Brunger, 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.
33. McClelland, M., Sanderson, K. E., Spieth, J., Clifton, S. W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., and Wilson, R. K. (2001) *Nature 413*, 852−856.
34. Horswill, A. R., and Escalante-Semerena, J. C. (1999) *Microbiology 145* (Part 6), 1381−1388.
35. Jetten, M. S., Stams, A. J., and Zehnder, A. J. (1989) *J. Bacteriol. 171*, 5430−5435.
36. Martinez-Blanco, H., Reglero, A., Fernandez-Valverde, M., Ferrero, M. A., Moreno, M. A., Penalva, M. A., and Luengo, J. M. (1992) *J. Biol. Chem. 267*, 5474−5481.
37. Luong, A., Hannah, V. C., Brown, M. S., and Goldstein, J. L. (2000) *J. Biol. Chem. 275*, 26458−26466.
38. An, J. H., Lee, G. Y., Jung, J. W., Lee, W., and Kim, Y. S. (1999) *Biochem. J. 344* (Part 1), 159−166.
39. Dieckmann, R., Pavela-Vrancic, M., von Dohren, H., and Kleinkauf, H. (1999) *J. Mol. Biol. 288*, 129−140.
40. Esnouf, R. M. (1999) *Acta Crystallogr. D55*, 938−940.
41. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
42. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
43. Cohen, G. H. (1997) *J. Appl. Crystallogr. 30*, 1160−1161.
44. Myers, E. W., and Miller, W. (1988) *Comput. Appl. Biosci. 4*, 11−17.

BI0271603