Crystal structures of TM0549 and NE1324—two orthologs of *E. coli* AHAS isozyme III small regulatory subunit

JANUSZ J. PETKOWSKI,^{1,3} MAKSYMILIAN CHRUSZCZ,^{1,3} MATTHEW D. ZIMMERMAN,^{1,3}
HEPING ZHENG,^{1,3} TATIANA SKARINA,^{2,3} OLENA ONOPRIYENKO,^{2,3}
MARCIN T. CYMBOROWSKI,^{1,3} KATARZYNA D. KOCLEGA,^{1,3}
ALEXEI SAVCHENKO,^{2,3} ALED EDWARDS,^{2,3} AND WLADEK MINOR^{1,3}

¹Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia 22908, USA

²Department of Medicinal Biophysics, University of Toronto, and Ontario Center for Structural Proteomics, Ontario Cancer Institute, Toronto, Ontario M5G 2C4, Canada

³Midwest Center for Structural Genomics

(RECEIVED January 30, 2007; FINAL REVISION April 9, 2007; ACCEPTED April 16, 2007)

Abstract

Crystal structures of two orthologs of the regulatory subunit of acetohydroxyacid synthase III (AHAS, EC 2.2.1.6) from *Thermotoga maritima* (TM0549) and *Nitrosomonas europea* (NE1324) were determined by single-wavelength anomalous diffraction methods with the use of selenomethionine derivatives at 2.3 Å and 2.5 Å, respectively. TM0549 and NE1324 share the same fold, and in both proteins the polypeptide chain contains two separate domains of a similar size. Each protein contains a C-terminal domain with ferredoxin-type fold and an N-terminal ACT domain, of which the latter is characteristic for several proteins involved in amino acid metabolism. The ferredoxin domain is stabilized by a calcium ion in the crystal structure of NE1324 and by a $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ ion in TM0549. Both TM0549 and NE1324 form dimeric assemblies in the crystal lattice.

Keywords: acetohydroxyacid synthase; actolactate synthase; regulatory subunit; ACT domain; AHAS; protein refolding

Acetohydroxyacid synthase (AHAS, EC 2.2.1.6) catalyzes two physiologically significant reactions in the synthesis of isoleucine, valine, and leucine. In the first reaction, which is an initial step in the synthesis of iso-

leucine, 2-aceto-2-hydroxy butyrate is produced by condensation of 2-ketobutyrate with pyruvate. In the second reaction, 2-acetolactate is synthesized from two pyruvate molecules, and this provides substrates for synthesis of valine and leucine (Umbarger 1978, 1987; Chipman et al. 1998). This pathway of branched-chain amino acid biosynthesis is characteristic for bacteria, fungi, algae, and higher plants, but not for animals. Accordingly, AHAS inhibitors have been developed as herbicides which have found broad application (Short and Colborn 1999). The inhibitors of branched-chain amino acids synthesis are also being tested as potential antituberculosis agents (Grandoni et al. 1998; Zohar et al. 2003; Choi et al. 2005).

Three different FAD-dependent AHAS isozymes (I, II, and III) are found in enterobacteria (e.g., *Escherichia*

Reprint requests to: Wladek Minor, Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, VA 22908, USA; e-mail address: wladek@iwonka.med.virginia.edu; fax: (434) 982-1616.

Abbreviations: AHAS, acetohydroxy synthase; CSU, catalytic subunit; IPTG, isopropyl-β-D-thiogalactopyranoside; MCSG, Midwest Center for Structural Genomics; MPD, 2-methyl-1,3-propanediol; NDSB, nondetergent sulfobetain; PEG, poly(ethylene glycol); SAD, single-wavelength anomalous diffraction; SSU, small regulatory subunit; SSM, secondary structure matching; TEV, tobacco etch virus.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.072793807>.

coli), but most other organisms from the bacteria encode only a single AHAS enzyme, similar to isozyme III from *E. coli*. The acetohydroxyacid synthases are multimeric proteins, and most frequently their biological unit is composed of two catalytic subunits (CSU) and two small regulatory subunits (SSU). In the absence of the SSU subunit, the CSU dimer is unstable (Vyazmensky et al. 1996), suggesting that in the holoenzyme the AHAS dimer is stabilized by a SSU dimer. The large catalytic subunits also show very weak activity without their associated regulatory subunits (Weinstock et al. 1992; Pang and Duggleby 1999). The *E. coli* AHAS enzyme can be reconstituted with SSU subunits from other organisms (Porat et al. 2004), and the reconstituted enzymes form stable heterotetramers. The properly associated holoenzymes show full catalytic activity, and are also susceptible to valine inhibition.

Previous mutagenesis studies (Mendel et al. 2001; Kaplun et al. 2006) revealed that the valine binding sites are located at the dimerization interface of the small regulatory subunits. It was determined that valine binds the N-terminal ACT domain (Aravind and Koonin 1999) of the regulatory subunit (Mendel et al. 2001). It was shown that the N-terminal domain of the *E. coli* protein is sufficient for activation of the catalytic subunit of the holoenzyme (Mendel et al. 2003).

Although the AHAS proteins have been studied for a long time, until recently no structure of an SSU has been reported. The first reported crystal structure of an AHAS III regulatory subunit is the SSU from *E. coli* (Kaplun et al. 2006). The *E. coli* regulatory subunit forms a dimer that contains two ferredoxin type domains in each monomer. In this paper, we present the structures of two orthologs of the *E. coli* SSU, from *Thermotoga maritima* (TM0549) and *Nitrosomonas europaea* (NE1324) that were determined by single-wavelength anomalous diffraction (SAD) methods by the application of selenomethionine derivatives.

Results and Discussion

Crystallization in the presence of L-arginine

TM0549 was crystallized in the presence of a high concentration of L-arginine, which was used during refolding. Although L-arginine is quite often used during protein refolding and purification (Tsumoto et al. 2004), its application during crystallization seems to be very limited. The presence of such a high concentration of arginine is quite unusual, and a search of REMARK 280 (crystallization conditions) in the PDB (as of September 2006) with keyword "arginine" yields only 18 matches (PDB codes: 1BG0, 1G00, 1GJV, 1GKX, 1GKZ, 1N13, 1P50, 1P52, 1Q20, 1S9R, 1SD0, 1T5G, 3CEV, 2AAF,

2ABR, 2ACI, 2A9G, and 2FLQ). In 12 of these cases, the structures represented arginine-processing enzymes, and in three other cases, arginine derivatives were used for crystallization. Only in the case of the structures of a branched-chain α -ketoacid dehydrogenase (BCKD) (PDB codes: 1GJV, 1GKX, and 1GKZ), was a relatively high concentration of arginine (300 mM) used during purification and crystallization (Machius et al. 2001), in order to achieve better protein stability. During the crystallization process, the protein has to be stable for a prolonged period of time in highly concentrated solution, which may not be possible for proteins that easily aggregate. Arakawa and Tsumoto (2003) proposed that arginine suppresses protein aggregation during refolding. Arginine could also be preventing excessive nucleation by acting as a "neutral crowder" (Baynes and Trout 2004).

Overall structure

The orthologs of AHAS III small regulatory subunit from *T. maritima* and *N. europaea* have the same overall fold. Two separate domains of similar size can be distinguished within each monomer. Both domains contain antiparallel β -strands and two α -helices (Fig. 1A). The N-terminal domain (ACT domain) is composed of a β -sheet created by strands β_1 , β_2 , β_3 , and β_4 and α -helices α_1 and α_2 , whereas the C-terminal domain contains a β -sheet created by strands β_5 – β_9 and two α -helices, α_3 and α_4 . The N and C domains share significant structural similarity, and both domains are of a ferredoxin fold type (Chipman and Shaanan 2001). Superposition of the domains with SSM (secondary structure matching) (Krissinel and Henrick 2004), as implemented in COOT, gives an RMSD value of 2.8 Å for 64 aligned residues. In the linker regions between the domains, the short fragments of residues 77–79 in NE1324 and 106–108 in TM0549 have a 3_{10} helix conformation. The crystal structure of SSU from *E. coli* is very similar to the structures presented here in terms of both the sequence and the overall fold (Kaplun et al. 2006). The sequence identity between the homologs from *E. coli* and *T. maritima* is 37% (similarity 67%) and between *E. coli* and *N. europaea* 58% (similarity 81%). The sequence identity/similarity between TM0549 and NE1324 is also quite high (45%/69%). The differences in main-chain conformation between all three proteins are small, with all RMSD values below 1.8 Å.

Sequence identities for both structures reported here versus other proteins with structures deposited in PDB (excluding the ortholog from *E. coli*) are rather low. The highest rate of identity among the regulatory subunits (~30%) is observed for the N-terminal regions that contain ACT domains, which play a role in the activation of CSU. Although the sequence identity is low, a ProFunc (Laskowski et al. 2005) search with the SSM program

Figure 1. (A) Stereoview of a ribbon diagram showing one monomer of the AHAS small regulatory subunit ortholog TM0549 from *Thermotoga maritima*. (B) Stereoview of a ribbon diagram of the dimer created by two monomers (chains A and C) of the AHAS small regulatory subunit ortholog NE1324 from *Nitrosomonas europaea*.

reveals that many proteins, not just those involved in amino acid metabolism, share a similar fold with the ACT domains from TM0549 and NE1324. Among these structurally similar proteins, those involved in amino acid metabolism are phosphoglycerate dehydrogenases (PDB codes: 1PSD, 1SC6, 1YBA, 1YGY), transcriptional regulators from the Lpr/AsnC family (PDB codes: 2CG4, 2CFX), and phenylalanine hydroxylase (PDB code: 1PHZ). There are also proteins (or protein fragments) with a very similar overall fold to TM0549 and NE1324, but with completely different functions. This group of proteins includes Atx1 metallochaperone protein from *Saccharomyces cerevisiae* (PDB codes: 1CC7, 1CC8), the copper transport proteins ATOX1 from *E. coli* and *Homo sapiens* (PDB codes: 1FEE, 1TL4, 1TL5), 30S ribosomal proteins from *Thermus thermophilus* (PDB codes: 1FKA, 1G1X), the CheY-binding domain of the chemotaxis

kinase from *E. coli* (PDB code: 1FWP), the C-terminal domain of nickel-responsive transcription factor NikR from *E. coli* (PDB code: 1Q5Y), and a bifunctional DNA primase-polymerase from *Sulfolobus islandicus* (PDB code: 1RNI).

Quaternary structure

Both NE1324 and TM0549 form dimers in their crystal structures (Fig. 1B). An asymmetric unit of NE1324 contains two dimers (four polypeptide chains in all), while only a single protein chain is found in the asymmetric unit of TM0549. The internal twofold symmetry of the dimer of TM0549 coincides with a symmetry element of the crystal lattice, and the probable biological unit can be obtained by transformation of a single chain by the twofold axis ($y, x, -z$). In the structure of NE1324

the biggest difference between chains (in terms of RMS displacement of C_{α} positions) is 0.96 Å for chain B compared to C, which are parts of two different dimers. The smallest RMS displacement value (0.51 Å) is observed between chains A and B, also belonging to different dimers. Analyses with the Protein Quaternary Structure Server (Postingl et al. 2003) show that the loss of solvent-accessible surface area (specifically, the sum of the areas lost by both monomers) upon dimer formation is significant: 2422 Å² for TM0549, and 2515 Å² for NE1324. The same calculations give a value of 2308 Å² for the equivalent homodimer of the *E. coli* protein. The hydrogen bond networks that participate in dimer formation in NE1324 and in the SSU of *E. coli* are similar (Fig. 2A), which is to be expected, given their high sequence identity. This is not the case, however, for TM0549. Despite its significant structural (Fig. 2B) and sequence similarity, TM0549's H-bond network in the dimer interface does not correspond to that of NE1324 or *E. coli* SSU. There is only one dimer interface H-bond that is conserved among the three SSUs of acetolactate synthases currently deposited in the PDB. This is the hydrogen bond between the main-chain carbonyl of Ile113 (in chain A) to NH1 of Arg156 (in chain C, as represented in the structure of NE1324).

Metal binding

NE1324 was crystallized in the presence of $CaCl_2$, and Ca^{2+} ions were localized in the crystal structure. The ions were mainly identified based on geometric criteria and the type of ligands forming the coordination sphere of the ions. Unrestrained refinement of the ion–ligand distances yield values in the 2.4–2.8 Å range. There were three metallic cation types in the crystallization solution, Na^+ , K^+ , and Ca^{2+} . The presence of K^+ ions can be excluded on the basis of the unrestrained cation–ligand distances (Harding 2002). The distances were later restrained using dictionary values for Ca^{2+} , and in the final model (2PC6) the ion–ligand distances are in the 2.3–2.5 Å range. In addition, refinement of the ion positions was also performed with *SHELXL* (Sheldrick 1997) and *CNS* (Brunger et al. 1998), to check the consistency of the resulting geometry around the calcium ions to the results obtained with *REFMAC*, as our analysis of PDB deposits shows a correlation between calcium–oxygen distances and the particular program used for structure refinement.

Finally, a model of the protein with a Na^+ ion instead of the Ca^{2+} ion was refined. The distances obtained for sodium were comparable to those observed in the case of calcium, as the ions are very similar in size (Harding 2002), and the B-factors were similar. Thus, due to the relatively low resolution of the data and the small differences in the B-factors, pure crystallographic data cannot

unambiguously prove that calcium and not sodium (or both types of ions) are present in the structure. There are two arguments that support the choice of Ca^{2+} instead of Na^+ : the presence of calcium ions in similar loops found in higher resolution structures (see below), and the lack of a metal ion in chain D despite the fact that the concentration of Na^+ in the crystallization solution is much higher than that of Ca^{2+} . The Ca^{2+} ions are bound by a loop composed of amino acids Asn65–Val70 (Fig. 2C) in three of the four chains (A, B, and C) in the asymmetric unit. The cations are coordinated by carbonyl oxygen atoms from Asn65, Leu67, Val70, and by water molecules. The corresponding loop in chain D is not completely ordered, so it seems that the presence of a calcium ion is needed in order to stabilize this part of the structure. Chain D is the only one which lacks density in the Asn65–Val70 loop, with no electron density that would correspond to the residue Glu69. In the crystal structure of TM0549, binding of a metal ion is also observed, in this case magnesium hexahydrate. The Mg^{2+} ion occupies a similar position as the Ca^{2+} ion in the structure of NE1324. The magnesium cation is coordinated by water molecules in an octahedral configuration. The $Mg^{2+} \cdots OH_2$ distances range from 1.9 Å to 2.3 Å. In contrast to the Ca^{2+} cations found in the crystal structure of NE1324, the magnesium hexahydrate ion mediates crystal contacts between symmetry-related molecules. The $Mg(H_2O)_6^{2+}$ cation is bound by amino acids that form the loop Tyr92–Val97. The symmetry related molecule ($-x + 1/2, -y + 1/2, z + 1/2$) interacts with $Mg(H_2O)_6^{2+}$ through oxygen atoms O and OE1 from Glu172, while the second symmetry molecule ($-y + 1/2, -x + 1/2, -z + 1/2$) contacts the $Mg(H_2O)_6^{2+}$ through atom OE2 from Glu66. In the crystal structure of *E. coli* SSU, which was crystallized in the presence of a high concentration (0.4–0.6 M) of $MgCl_2$, several magnesium ions were localized. All of them mediate crystal contacts between protein molecules, and none of them is closer than ~10 Å to the position occupied by the $Mg(H_2O)_6^{2+}$ in the structure of TM0549.

Comparison of the sequences of the NE1324 and TM0549 metal-binding loops shows the presence of a motif containing valine, leucine, and a three amino acid linker between them. An extensive search of the PDB shows that this motif is not characteristic for magnesium binding, but rather for calcium binding. Similar loops bind Ca^{2+} cations in different bacterial lipases, such as lipases from *Pseudomonas glumae* (formerly *Chromobacterium viscosum*) (Noble et al. 1993; Lang et al. 1996) and *Burkholderia cepacia* (formerly *Pseudomonas cepacia*) (Shrag et al. 1997; Luić et al. 2001). The lipases are not the only enzymes that contain such a metal binding motif; it is also observed in human DNA polymerase β (Pelletier and Sawaya 1996). The specific role of the

Figure 2. (A) Sequence alignment of NE1324 (NE), SSU from *E. coli* (EC), and TM0549 (TM). Conserved amino acids marked by green squares are involved in H-bond-mediated dimer formation in NE1324 and *E. coli* SSU. Blue squares indicate residues participating in dimer formation that are conserved in all three SSUs. (B) Stereoview of the superposition of mainchain atoms: *T. maritima* (red), *N. europea* (green), and *E. coli* (blue). For the superposition the A chain from each structure was used. (C) Stereoview of the superposition of main-chain atoms of all chains from NE1324 structure (2PC6), with chain A in red, B in blue, C in green, and D in purple. The calcium ion from chain A is marked as a black sphere. (D) Coordination of the calcium ion in chain A of the NE1324 structure. Water molecules are shown as red spheres.

metal ions in the NE1324 and TM0549 structures is not yet known. Most probably metal ions stabilize the proteins by reducing local flexibility thus decreasing their susceptibility to partial unfolding (Gregory et al. 1993). This hypothesis is consistent with the higher mobility of chain D in the NE1324 structure, which lacks the bound Ca^{2+} ion.

Summary

As expected, the acetohydroxy acid synthase small regulatory subunits from *T. maritima* and *N. europea* share the same fold with their ortholog from *E. coli*. All three SSUs form dimers in the crystal structures. The crystal structures of TM0549 and NE1324 reveal a metal binding site common to both proteins. It is not clear if the metal binding site has any physiological significance, because the metal ions that were found in the crystal structures were introduced with crystallization buffers, and the AHAS small regulatory subunit from *E. coli* did not have a metal bound in an equivalent position, despite being crystallized in the presence of magnesium ions.

Materials and Methods

Protein cloning, expression, purification, and crystallization

Nitrosomonas europea

The small regulatory subunit of acetolactate synthase ortholog from *N. europea* (NE1324) was cloned, expressed, and purified as described previously (Zhang et al. 2001). After cleavage of the polyhistidine tag, the purified selenomethionine derivative of the protein was crystallized at room temperature using vapor diffusion in sitting drops. Crystals were obtained by mixing 1:1 10 mg/mL protein solution in 500 mM NaCl, 10 mM HEPES-Na pH = 7.5 with a well solution composed of 2% Tacsimate, 10% PEG400, 0.1 M KCl, 10 mM CaCl_2 , 2 mM L-cysteine and 50 mM HEPES-Na pH 7.0. After 12 h of growth protein crystals had grown to $200 \times 100 \times 100 \mu\text{m}$ in size, and were chosen for data collection.

Thermotoga maritima

The *T. maritima* acetolactate synthase small regulatory subunit ortholog TM0549 was cloned using the same protocol as for NE1324 and similarly was expressed as the selenomethionine derivative in a modified pET-15b vector (Novagen). The vector has been modified to replace the thrombin cleavage site coding sequence at the N terminus by a TEV protease recognition sequence (ENLYFQG). The modified pET-15b vector containing a cloned TM0549 was transformed into *E. coli* B834(DE3)pLysS. TM0549, due to its low solubility, could not be purified using the standard MCSG (Midwest Center for Structural Genomics) protocol. Instead, the protein was purified under denaturing conditions. The cell pellet with expressed selenomethionine-substituted protein was resuspended in a buffer composed of 500 mM NaCl, 50 mM HEPES-Na pH = 7.5, 5% glycerol, 5 mM imidazole and 6 M guanidine hydro-

chloride. Resuspended cells were sonicated, passed through a $0.45\text{-}\mu\text{m}$ filter, and applied onto a column of Ni-NTA resin (Qiagen). The denatured protein was eluted with a buffer containing 500 mM NaCl, 50 mM HEPES-Na pH = 7.5, 5% glycerol, 250 mM imidazole and 6 M guanidine hydrochloride. The purified protein was refolded by diluting it ten times in refolding buffer (500 mM L-arginine pH = 8.5, 50 mM NaCl) and incubating overnight at room temperature. Subsequently the refolded protein was purified on a Superdex 200 column, using an AKTA FPLC (GE Healthcare) system. The purified and concentrated protein was stored in the refolding buffer; the polyhistidine tag was not cleaved.

The protein was crystallized at 20°C using vapor diffusion method in hanging drops. An initial hit was obtained with Hampton Research Crystal Screen I, condition #44 (200 mM magnesium formate). The tracking, analysis, and design of optimizations of crystallization conditions were performed with the *Xtaldb* crystallization expert system (M.D. Zimmerman and W. Minor, unpubl.). Single, well-formed, diffraction-quality crystals ($50 \times 50 \times 50 \mu\text{m}$) appeared after 24 h in a drop containing 1:1 mixture of 7.7 mg/mL of protein solution and solution containing 2% glycerol, 0.4% w/v NDSB201, and 150 mM magnesium formate.

Data collection, structure determination, and refinement

Prior to data collection, crystals from *T. maritima* were transferred to a cryosolution composed of an 8:5 mixture of well solution and 50% MPD, and cooled by plunging into liquid nitrogen. In the case of the crystals from *N. europea*, the cryosolution was made of a 7:3 mixture of well solution and PEG400. Data collection for both proteins was done at beamline 19-ID of the Structural Biology Center (Rosenbaum et al. 2006) at the Advanced Photon Source (APS). Data for both structures were collected at 100 K. Data collection, structure determination, and refinement statistics are summarized in Table 1. Data from Se-Met substituted samples were processed with *HKL-2000* (Otwinowski and Minor 1997). Structures were solved using SAD data, and initial models were built with *HKL-3000* (Minor et al. 2006), which is integrated with *SHELXD* (Schneider and Sheldrick 2002), *SHELXE* (Sheldrick 2002), *MLPHARE* (Otwinowski 1991), *DM* (Cowtan 1994), *O* (Jones et al. 1991), *SOLVE* (Terwilliger and Berendzen 1999), and *RESOLVE* (Terwilliger 2002). Initial models were refined with *REFMAC5* (Murshudov et al. 1997) and *COOT* (Emsley and Cowtan 2004). The TLSMD Web server (Painter and Merritt 2006) was used for generation of multigroup TLS models. *MOLPROBITY* (Lovell et al. 2003) and *PROCHECK* (Laskowski et al. 1993) were used for structure validation. The atomic coordinates for both structures, together with the structure factors, were deposited in the PDB, with accession codes 2FGC and 2PC6 for the proteins from *T. maritima* and *N. europea*, respectively. In the crystal structure of TM0549 (2FGC), the N-terminal polyhistidine tag, the first four, and the last six residues could not be observed in the final electron density map. For NE1324 (2PC6), the electron density for chain D is of lower quality than that of other chains. In this chain residue 69 and many side chains are missing. The high mobility of chain D could be related to the lack of a calcium ion, which is found to stabilize similar moieties in other parts of the structure, or more probably is a consequence of crystal packing. The only residue (Glu55) that is outside the allowed region of the Ramachandran plot also belongs to chain D.

Table 1. Summary of data collection, phasing, and refinement for the AHAS III SSU orthologs from *Thermotoga maritima* and *Nitrosomonas europaea*

Organism PDB code	<i>Thermotoga maritima</i> 2FGC	<i>Nitrosomonas europaea</i> 2PC6
Crystal		
Space group	I422	P4 ₂ 2 ₁ 2
Unit cell	a = b = 104.8 Å c = 78.5 Å α = β = γ = 90°	a = b = 122.1 Å c = 111.6 Å α = β = γ = 90°
Solvent content (%)	49.8	56.7
Matthews coefficient (Å ³ /Da)	2.45	2.84
No. of molecules in asymmetric unit	1	4
Data collection		
Diffraction protocol	SAD	SAD
Wavelength (Å)	0.9791	0.9795
Resolution (Å)	27.3–2.3	35.6–2.5
Highest resolution shell (Å)	2.36–2.30 Å	2.57–2.50 Å
Unique reflections	9510 (723)	28,271 (2161)
Redundancy	7.1 (6.6)	18.9 (19.3)
Completeness (%)	99.5 (100.0)	99.9 (100.0)
I/σ(I)	35.9 (4.6)	60.7 (7.4)
R _{merge} (%)	9.7 (45.8)	8.4 (43.1)
SAD analysis		
No. of Se sites located/theoretical	5/7	17/20
FOM _{MLPHARE}	0.29	0.36
FOM _{DM}	0.85	0.84
Refinement		
R _{work} (%)	16.8 (17.4)	20.5 (26.5)
R _{free} (%)	23.6 (29.1)	27.6 (38.2)
RMSD bond length (Å)	0.022	0.019
RMSD bond angles (°)	1.85	1.84
Number of protein atoms	1232	4918
Number of water molecules	43	200
Average B factor (Å ²)	55.9	47.5
B factor from Wilson plot (Å ²)	49.3	44.2
Ramachandran plot		
Most favored regions (%)	91.7	92.8
Additional allowed regions (%)	8.3	7.0
Disallowed regions (%)	0.0	0.2

Data for the highest resolution shell are given in parentheses. Ramachandran plots were generated with *PROCHECK*. R_{merge} was calculated for merged Friedel pairs.

Acknowledgments

The authors thank Andrzej Joachimiak and the members of the Structural Biology Center at the Advanced Photon Source and the Midwest Center for Structural Genomics for help and discussions. We also thank Alex Wlodawer for reading the manuscripts and making valuable comments. The results shown in this report are derived from work performed at Argonne National Laboratory, at the Structural Biology Center of the Advanced Photon Source. Argonne is operated by University of Chicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract

DE-AC02-06CH11357. The work described in the paper was supported by NIH PSI Grants GM62414 and GM074942.

References

- Arakawa, T. and Tsumoto, K. 2003. The effects of arginine on refolding of aggregated proteins: Not facilitate refolding, but suppress aggregation. *Biochem. Biophys. Res. Commun.* **304**: 148–152.
- Aravind, L. and Koonin, E.V. 1999. Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. *J. Mol. Biol.* **287**: 1023–1040.
- Baynes, B.M. and Trout, B.L. 2004. Rational design of solution additives for the preventing of protein aggregation. *Biophys. J.* **87**: 1631–1639.
- Brunker, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Krukowski, J., Nilges, M., Pannu, N.S., et al. 1998. Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**: 905–921.
- Chipman, D.M. and Shaanan, B. 2001. The ACT domain family. *Curr. Opin. Struct. Biol.* **11**: 694–700.
- Chipman, D., Barak, Z., and Schloss, J.V. 1998. Biosynthesis of 2-aceto-2-hydroxy acids: Acetolactate synthases and acetohydroxyacid synthases. *Biochim. Biophys. Acta* **1385**: 401–419.
- Choi, K.-J., Yu, Y.G., Hahn, H.G., Choi, J.-D., and Yoon, M.-Y. 2005. Characterization of acetohydroxyacid synthase from *Mycobacterium tuberculosis* and the identification of its new inhibitor from screening of a chemical library. *FEBS Lett.* **579**: 4903–4910.
- Cowtan, K. 1994. Joint CCP4 and ESF-EACBM newsletter. *Protein Crystallogr.* **31**: 34–38.
- Emsley, P. and Cowtan, K. 2004. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**: 2126–2132.
- Grandoni, J.A., Marta, P.T., and Schloss, J.V. 1998. Inhibitors of branched-chain amino acid biosynthesis as potential antituberculosis agents. *J. Antimicrob. Chemother.* **42**: 475–482.
- Gregory, D.S., Martin, A.C.R., Cheetham, J.C., and Rees, A.R. 1993. The prediction of metal binding sites in proteins. *Protein Eng.* **6**: 29–35.
- Harding, M.M. 2002. Metal–ligand geometry relevant to proteins and in proteins: Sodium and potassium. *Acta Crystallogr. D Biol. Crystallogr.* **58**: 872–874.
- Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**: 110–119.
- Kaplan, A., Vyazmensky, M., Zherdev, Y., Belenky, I., Slutzker, A., Mendel, S., Barak, Z., Chipman, D.M., and Shaanan, B. 2006. Structure of the regulatory subunit of acetohydroxyacid synthase isozyme III from *Escherichia coli*. *J. Mol. Biol.* **357**: 951–963.
- Krissinel, E. and Henrick, K. 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60**: 2256–2268.
- Lang, D., Hofmann, B., Haalck, L., Hecht, H.-J., Spener, F., Schmid, R.D., and Schomburg, D. 1996. Crystal structure of bacterial lipase from *Chromobacterium viscosum* ATCC 6918 refined at 1.6 Å resolution. *J. Mol. Biol.* **259**: 704–717.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**: 283–291.
- Laskowski, R.A., Watson, J.D., and Thornton, J.M. 2005. ProFunc: A server for predicting protein function from 3D structure. *Nucleic Acids Res.* **33**: W89–W93.
- Lovell, S.C., Davis, I.W., Arendall III, W.B., de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. 2003. Structure validation by Cα geometry: φ, ψ and Cβ deviation. *Proteins* **50**: 437–450.
- Luić, M., Tomić, S., Lešić, I., Ljubović, E., Šepac, D., Šunjić, V., Vitale, L., Saenger, W., and Kojić-Prodić, B. 2001. Complex of *Burkholderia cepacia* lipase with transition state analogue of 1-phenoxy-2-acetoxybutane. *Eur. J. Biochem.* **268**: 3964–3973.
- Machius, M., Chuang, J.L., Wynn, R.M., Tomchick, D.R., and Chuang, D.T. 2001. Structure of rat BCKD kinase: Nucleotide-induced domain communication in a mitochondrial protein kinase. *Proc. Natl. Acad. Sci.* **98**: 11218–11223.
- Mendel, S., Elkayam, T., Sella, C., Vinogradov, V., Vyazmensky, M., Chipman, D.M., and Barak, Z. 2001. Acetohydroxyacid synthase: A proposed structure for regulatory subunits supported by evidence from mutagenesis. *J. Mol. Biol.* **307**: 465–477.

- Mendel, S., Vinogradov, M., Vyazmensky, M., Chipman, D.M., and Barak, Z. 2003. The N-terminal domain of the regulatory subunit is sufficient for complete activation of acetohydroxyacid synthase III from *Escherichia coli*. *J. Mol. Biol.* **325**: 275–284.
- Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. 2006. HKL-3000: The integration of data reduction and structure solution—From diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr.* **62**: 859–866.
- Murshudov, G.N., Vagin, A.A., and Dodson, E.J. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**: 240–255.
- Noble, M.E.M., Cleasby, A., Johnson, L.N., Egmond, M.R., and Frenken, L.G.L. 1993. The crystal structure of triacylglycerol lipase from *Pseudomonas glumae* reveals a partially redundant catalytic aspartate. *FEBS Lett.* **331**: 123–128.
- Otwinowski, Z. 1991. Proceedings of the CCP4 study weekend. isomorphous replacement and anomalous scattering (eds. W. Wolf, P.R. Evans, and A.G.W. Leslie), pp. 80–86. Daresbury Laboratory, Warrington, UK.
- Otwinowski, Z. and Minor, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. In *Methods in enzymology: Macromolecular crystallography, part A*. (eds. Carter, C.W. Jr., and R.M. Sweet). Vol. 276, pp. 307–326. Academic Press, New York.
- Painter, J. and Merritt, E.A. 2006. *TLSMD* web server for the generation of multi-group TLS models. *J. Appl. Crystallogr.* **39**: 109–111.
- Pang, S.S. and Duggleby, R.G. 1999. Expression, purification, characterization, and reconstruction of the large and small subunits of yeast acetohydroxyacid synthase. *Biochemistry* **38**: 5222–5231.
- Pelletier, H. and Sawaya, M.R. 1996. Characterization of the metal ion binding helix-hairpin-helix motifs in human dna polymerase β by X-ray structural analysis. *Biochemistry* **35**: 12778–12787.
- Porat, I., Vinogradov, M., Vyazmensky, M., Lu, C.-D., Chipman, D.M., Abdelal, A.T., and Barak, Z. 2004. Cloning and characterization of acetohydroxyacid synthase from *Bacillus stearothermophilus*. *J. Bacteriol.* **186**: 570–574.
- Postigl, H., Kabir, T., and Thornton, J.M. 2003. Automatic inference of protein quaternary structure from crystals. *J. Appl. Crystallogr.* **36**: 1116–1122.
- Rosenbaum, G., Alkire, R.W., Evans, G., Rotella, F.J., Lazarski, K., Zhang, R.-G., Ginell, S.L., Duke, N., Naday, I., Lazarz, J., et al. 2006. The structural biology center 19ID beamline: Facility, specifications and protein crystallographic results. *J. Synchrotron Radiat.* **13**: 30–45.
- Schneider, T.R. and Sheldrick, G.M. 2002. Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.* **58**: 1772–1779.
- Sheldrick, G.M. 1997. SHELXL-97 program for crystal structure refinement, vol. 4, p. D-3400. Institut für Anorganische Chemie der Universität, Göttingen, Germany.
- Sheldrick, G.M. 2002. Macromolecular phasing with SHELXE. *Z. Kristallogr.* **217**: 644–650.
- Short, P. and Colborn, T. 1999. Pesticide use in the U.S. and policy implications: A focus on herbicides. *Toxicol. Ind. Health* **15**: 240–275.
- Shrag, J.D., Li, Y., Cygler, M., Lang, D., Burgdorf, T., Hecht, H.-J., Schmidt, R., Schomburg, D., Rydel, T.J., Oliver, J.O., et al. 1997. The open conformation of *Pseudomonas* lipase. *Structure* **5**: 187–202.
- Terwilliger, T.C. 2002. Automated structure solution, density modification and model building. *Acta Crystallogr. D Biol. Crystallogr.* **58**: 1937–1940.
- Terwilliger, T.C. and Berendzen, J. 1999. Automated MAD and MIR structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **55**: 849–861.
- Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Philo, J.S., and Arakawa, T. 2004. Role of arginine in protein refolding, solubilization, and purification. *Biotechnol. Prog.* **20**: 1301–1308.
- Umbarger, H.E. 1978. Amino acid biosynthesis and its regulation. *Annu. Rev. Biochem.* **47**: 533–606.
- Umbarger, H.E. 1987. Biosynthesis of branched chain aminoacids. In *Escherichia coli and Salmonella typhimurium: Cellular and molecular biology* (eds. F.C. Neidhardt et al.), pp. 352–367. American Society for Microbiology, Washington, DC.
- Vyazmensky, M., Sella, C., Barak, Z., and Chipman, D.M. 1996. Isolation and characterization of subunits of acetohydroxy acid synthase isozyme III and reconstitution of the holoenzyme. *Biochemistry* **35**: 10339–10346.
- Weinstock, O., Sella, C., Chipman, D.M., and Barak, Z. 1992. Properties of subcloned subunits of bacterial acetohydroxy acid synthases. *J. Bacteriol.* **174**: 5560–5566.
- Zhang, R.G., Skarina, T., Katz, J.E., Beasley, S., Khachatryan, A., Vyas, S., Arrowsmith, C.H., Clarke, S., Edwards, A., Joachimiak, A., et al. 2001. Structure of *Thermotoga maritima* stationary phase survival protein SurE: A novel acid phosphatase. *Structure* **9**: 1095–1106.
- Zohar, Y., Einav, M., Chipman, D.M., and Barak, Z. 2003. Acetohydroxyacid synthase from *Mycobacterium avium* and its inhibition by sulfonylureas and imidazolinones. *Biochim. Biophys. Acta* **1649**: 97–105.