Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Acidophilic adaptations in the structure of *Acetobacter aceti* N^5-carboxyaminoimidazole ribonucleotide mutase (PurE)

Ethan C. Settembre,[a] Johnathan R. Chittuluru,[a] Christopher P. Mill,[b] T. Joseph Kappock[b] and Steven E. Ealick[a]*

[a]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853, USA, and [b]Department of Chemistry, Washington University, St Louis, MO 63130, USA

Correspondence e-mail: see3@cornell.edu

Received 13 July 2004
Accepted 28 July 2004

PDB Reference: PurE, 1u11, r1u11sf.

The crystal structure of *Acetobacter aceti* PurE was determined to a resolution of 1.55 Å and is compared with the known structures of the class I PurEs from a mesophile, *Escherichia coli*, and a thermophile, *Thermotoga maritima*. Analyses of the general factors that increase protein stability are examined as potential explanations for the acid stability of *A. aceti* PurE. Increased inter-subunit hydrogen bonding and an increased number of arginine-containing salt bridges appear to account for the bulk of the increased acid stability. A chain of histidines linking two active sites is discussed in the context of the proton transfers catalyzed by the enzyme.

1. Introduction

Acetobacter aceti is an aerobic Gram-negative bacterium that has been used for millenia to convert wine into vinegar, a concentrated solution of acetic acid that is toxic to many microbes (Asai, 1968). Acetic acid readily permeates cell membranes and dissociates at the higher internal pH (pH_{in}) of susceptible cells, poisoning and de-energizing them (Foster, 1999). In contrast, *A. aceti* tolerates the constant influx of acetic acid by allowing pH_{in} to fall in register with the external pH and is viable with pH_{in} = 4 (Menzel & Gottschalk, 1985). *A. aceti* enzymes are thus excellent models for understanding how proteins adapt to function at low pH, which generally destabilizes proteins. Acid-stable proteins such as rusticyanin (Botuyan *et al.*, 1996; Walter *et al.*, 1996) seem to have characteristic structural differences, including a comparatively low number of surface charges and a higher number of internal hydrogen-bonding interactions. By analogy, thermostable proteins have various stabilizing structural features (Sterner & Liebl, 2001). There is a much greater amount of structural data available for thermostable enzymes than for acid-stable enzymes.

N^5-Carboxyaminoimidazole ribonucleotide (N^5-CAIR) mutase (PurE) catalyzes the only C—C-bond-forming reaction in *de novo* purine biosynthesis, the formation of 4-carboxy-5-aminoimidazole ribonucleotide (CAIR). Crystal structures of *Escherichia coli* PurE (*Ec*PurE) reveal a homo-octameric protein with active sites at subunit interfaces (Mathews *et al.*, 1999). PurE isoforms functionally segregate into two families: class I enzymes reversibly convert N^5-CAIR to CAIR, while class II enzymes (AIR carboxylases) convert AIR and CO_2 into CAIR (Firestine *et al.*, 1994; Mueller *et al.*, 1994). Each class has a characteristic set of conserved residues in the '70s loop' (*Ec*PurE sequence numbering): AG*X*(A/S)A*X*L in class I PurEs and A(G/A)RSN(G/A)L in class II PurEs (Mathews *et al.*, 1999). This biochemical divergence in the middle of the ancient purine biosynthetic pathway has interesting evolutionary implications (Kappock *et*

al., 2000). Most bacteria contain class I PurE and N^5-CAIR synthetase (PurK), the preceding enzyme in the purine pathway, as individual proteins (Meyer *et al.*, 1992). Yeasts and plants produce a bifunctional PurE-PurK fusion protein (Firestine *et al.*, 1998). Completely sequenced animal genomes contain a class II PurE gene but lack a recognizable gene for PurK, consistent with biochemical evidence that AIR is a substrate for class II PurE (Firestine *et al.*, 1994). Vertebrate PurE is a C-terminal fusion to 4-(*N*-succinocarboxamide)-5-aminoimidazole synthetase (PurC; Chen *et al.*, 1990). Sequence analyses suggest the existence of more than two PurE subfamilies, but as yet there is no biochemical evidence for further differentiation.

*Aa*PurE also contends with acid-mediated decomposition of its substrate. N^5-CAIR readily decarboxylates to 5-aminoimidazole ribonucleotide (AIR) with a half-life of 0.9 min (pH 7.8, 303 K; Mueller *et al.*, 1994). In the cytoplasmic environment of *A. aceti* or in thermophiles, non-enzymatic N^5-CAIR decomposition would be rapid and must be avoided in order to maintain efficient purine biosynthesis. The structure of PurE from *Thermotoga maritima* (*Tm*PurE), which closely resembles *Ec*PurE, reveals nothing about how N^5-CAIR might be stabilized at the organism's growth temperature, 368 K (Schwarzenbacher *et al.*, 2004).

Here, we describe the crystal structure of *Aa*PurE determined at 1.55 Å. Through structural comparison of *Aa*PurE to both *Ec*PurE and *Tm*PurE, we identify structural features that may confer acid stability or thermostability, respectively.

2. Experimental procedures

2.1. *Aa*PurE isolation

The contiguous genes for *Aa*PurE and *Aa*PurK were cloned from *A. aceti* strain 1023, which is highly resistant to acetic acid (Ohmori *et al.*, 1980), and placed in pET23a to yield plasmid pJK174. An extra N-terminal Met is encoded by the pJK174 gene for *Aa*PurE. Recombinant *Aa*PurE was isolated from BL21(DE3)/pJK174 cells using a procedure adapted from the purification of *Ec*PurE (Meyer *et al.*, 1992). *Aa*PurE cloning, purification, characterization and experiments demonstrating its acid stability will be reported elsewhere. Pure *Aa*PurE was concentrated to 33 mg ml^{-1} in 10 m*M* Tris pH 8.0 and stored at 193 K until needed. Edman degradation (Midwest Analytical, St Louis, MO, USA) showed an ∼3:1 ratio of MMSETAP to SETAPLP, exactly matching the expected N-terminus.

2.2. Crystallization of *Aa*PurE

*Aa*PurE was crystallized using the hanging-drop method with 1 µl protein to 1 µl well solution. The crystals grew best at 295 K with the optimized mother-liquor conditions of 21–23% PEG 4K, 0.19 *M* ammonium acetate and 90 m*M* citrate pH 5.25–5.5. They grew over two weeks to maximum dimensions of 0.4 × 0.4 × 0.1 mm. Crystals were either frozen in liquid nitrogen or in a nitrogen stream following a quick dunk in the above solution with an additional 2% PEG 4K and 15% ethylene glycol. Preliminary crystallographic analysis indicated the crystals belonged to space group *I*422, with unit-cell parameters a = 99.25, c = 164.81 Å. This corresponded to a solvent content of ∼47% with two molecules per asymmetric unit.

Table 1
Data-collection and refinement statistics.

Values for the outer resolution shell are given in parentheses. Data were collected at Cornell High Energy Synchrotron Source (CHESS) beamline F2.

Data collection	
Wavelength (Å)	0.9760
Resolution (Å)	50–1.55 (1.62–1.55)
No. reflections	427878
No. unique reflections	59415
Redundancy	7.2 (6.9)
Completeness	99.7 (99.9)
R_{sym}† (%)	6.9 (36.4)
$I/\sigma(I)$	24 (4.5)
Refinement	
Resolution (Å)	50-1.55
Total No. non-H atoms	2698
No. protein atoms	2410
No. water atoms	262
No. ligand atoms	26
No. reflections in refinement	58257
No. reflections in test set	5846
R factor‡ (%)	18.9
R_{free}§ (%)	19.9
R.m.s. deviation from ideal geometry	
Bonds (Å)	0.006
Angles (°)	1.27
Ramachandran plot	
Most favored region (%)	94.8
Additional allowed region (%)	5.2
Average B factors (Å^2)	
Protein	15.0
Water	27.7
Ligand	34.2

† $R_{sym} = \sum\sum_i |I_i - \langle I\rangle| / \sum\langle I\rangle$, where $\langle I\rangle$ is the mean intensity of the N reflections with intensities I_i and common indices hkl. ‡ R factor = $\sum_{hkl} \big||F_{obs}| - k|F_{cal}|\big| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{cal} are observed and calculated structure factors, respectively. § For R_{free}, the sum is extended over a subset of reflections excluded from all stages of refinement, ∼10% for native data.

2.3. Data collection and processing

Single-wavelength X-ray data were collected at the Cornell High Energy Synchrotron Source (CHESS) on beamline F2 at a wavelength of 0.976 Å using a Quantum 210 detector (Area Detector System Corp). Data were collected over 100° using 10 s for each 0.5° oscillation with a crystal-to-detector distance of 135 mm. Integration and scaling were performed with *HKL*2000 (Otwinowski & Minor, 1997). Data-collection and processing statistics are summarized in Table 1.

2.4. Structure determination, model building and refinement

The current structure was determined by molecular replacement using the program *CNS* (Brünger *et al.*, 1998). The previously reported *Ec*PurE monomer (PDB code 1qcz) was used as the search model (Mathews *et al.*, 1999). Residues that were not identical between the two species were mutated to alanine. The refinement procedure included successive rounds of rigid-body refinement, simulated-annealing refinement, temperature-factor refinement and model building. A composite omit map was calculated and all side chains that

differed between the two structures were mutated to the proper residue type. All model building was performed with the program *O* (Jones *et al.*, 1991) and all model refinement was performed with *CNS* (Brünger *et al.*, 1998). The unexplained active-site density was modeled as citrate owing to the high content of citrate in the crystallization conditions and the general shape of the density. The contents of the active site were manually fitted and followed by a few cycles of refinement. The model was examined using *CNS* and *PROCHECK*. The final refinement statistics are summarized in Table 1.

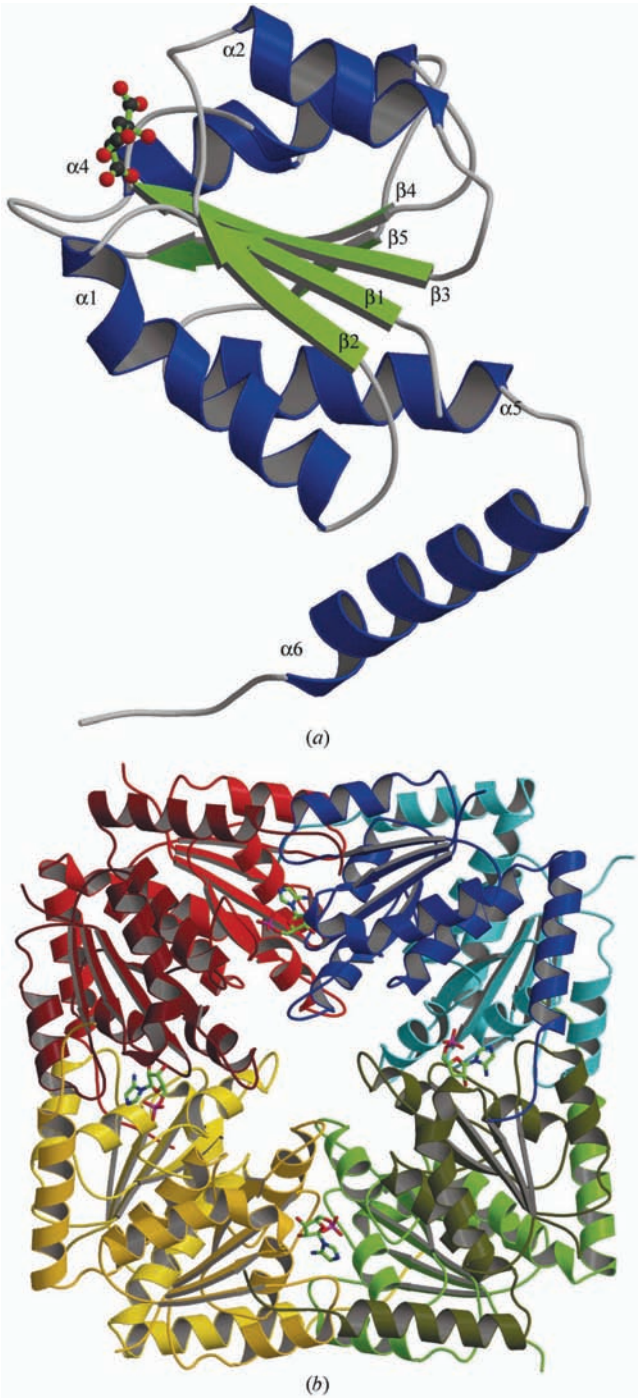

Figure 1
Structure of *Aa*PurE. (*a*) Subunit structure of *Aa*PurE with α-helices colored blue and β-strands colored green. A molecule of citrate is shown bound in the putative active site in ball-and-stick representation. (*b*) View of *Aa*PurE octamer down the molecular fourfold axis. Subunits on the back face are colored in darker colors and subunits on the front face are colored in lighter colors. A model of CAIR based on the *Ec*PurE structure is shown in stick representation modeled into four active sites on the upper face of the octamer. This figure was prepared with *MOLSCRIPT* (Kraulis, 1991) and *RASTER3D* (Merritt & Bacon, 1997).

3. Results

3.1. Quality of the model

The final model contains 2410 protein atoms, 262 water molecules and 22 citrate atoms. Residues 1–19 and 179–182 of monomer *A* and residues 1–17 and 181–182 of monomer *B* are disordered in the density and were not built into the model. Overall, there were 11 residues with alternate conformations. The final model has an *R* factor of 18.9 and an $R_{\rm free}$ of 19.9. The Ramachandran plot shows 94.8% of the residues in the most favorable region, with the remaining 5.2% in the additional allowed area.

3.2. Monomer fold

The *Aa*PurE subunit is a single-domain protein that consists of a central core containing a five-stranded parallel β-sheet flanked by two α-helices on one side and by three α-helices on the other (Fig. 1*a*). The single C-terminal α-helix protrudes from this core and is involved in contacts with neighboring subunits. Superposition of the C^{α} positions of *Aa*PurE with those of *Ec*PurE or *Tm*PurE indicated that their folds were nearly identical. *Ec*PurE showed an r.m.s. difference of 0.6 Å for 158 of 161 amino acids with 58% identity, while *Tm*PurE showed an r.m.s. difference of 1.1 Å for 155 of 169 residues with 57% identity.

3.3. Quaternary structure

PurE is an octamer with 422 point symmetry (Fig. 1*b*). The overall octamer is box-like with dimensions of ∼75 Å along two edges and ∼40 Å in thickness. The top and bottom faces each contain four monomers that are related by fourfold symmetry. The monomers in each face are related to the other face by the twofold axes. At the center of the octamer, there is a cavity that is ∼19 Å in its largest dimension. This cavity has two openings, with diameters of 6.5 and 8.5 Å. The outer rims of the openings are lined by the positively charged residues Lys113 and Arg110, while the interior of the cavity is composed of more hydrophobic residues, including Met115, Leu119 and Leu112. Owing to the fourfold and twofold symmetry axes, the cavity is bounded by eight symmetry-related groups of the preceding residues.

Each monomer has contacts with four other monomers. The buried surface area of a single monomer is ∼4220 $Å^2$, which is approximately half of the possible monomer surface area of ∼8458 $Å^2$. The main subunit interactions occur between subunits on one face of the octamer and subunits on the other face. In each monomer, there are 89 pairwise inter-subunit interactions (defined as any pair of residues having at least one atom within 3.5 Å radius of an atom in the interacting

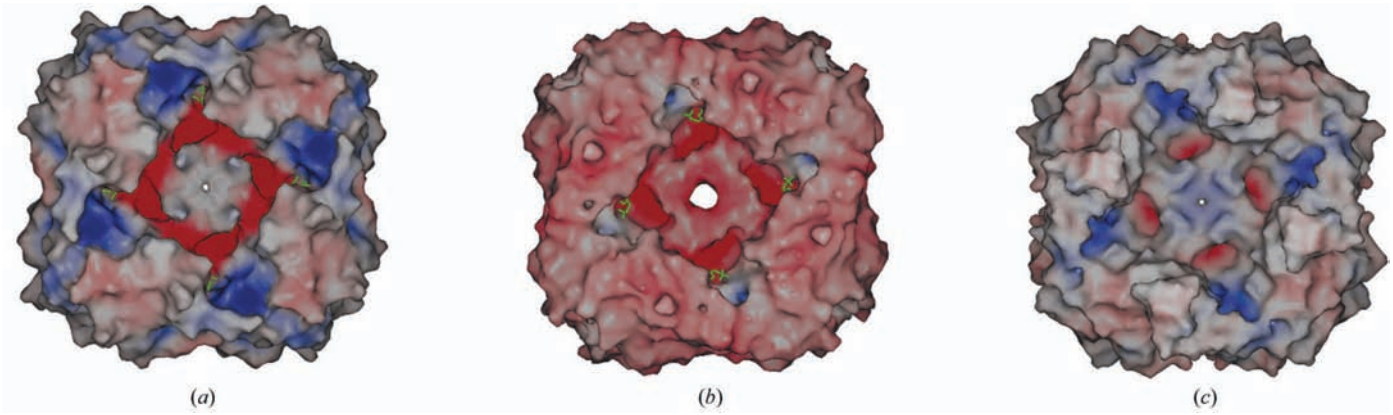

Figure 2
Electrostatic surface representation of *Aa*PurE, *Ec*PurE and *Tm*PurE calculated with *SPOCK* (Christopher, 1998). Only protein atoms were considered in the calculation. The surface is colored blue for positively charged residues and red for negatively charged residues. The saturation of the color is proportional to the degree of electrostatic charge from $-25kT$ to $25kT$. (*a*) *Aa*PurE is shown with citrate bound at the active sites. Citrate is colored green and drawn in ball-and-stick representation. (*b*) *Ec*PurE is shown with AIR bound at the active site. AIR is colored green and drawn in ball-and-stick representation. (*c*) *Tm*PurE in the same orientation as the other two PurEs.

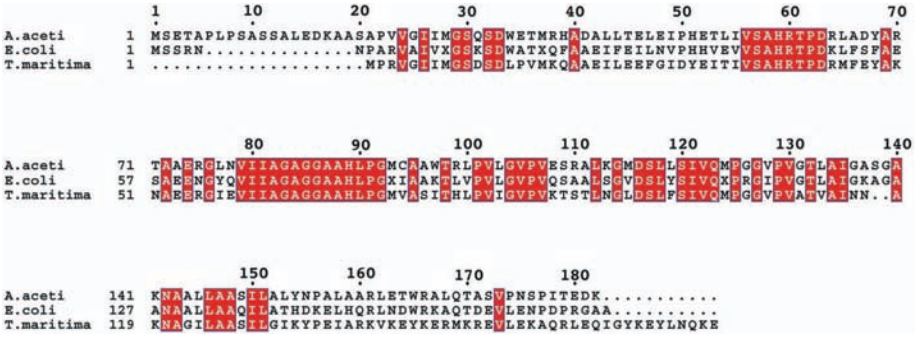

Figure 3
Multiple sequence alignment of *Aa*PurE, *Ec*PurE and *Tm*PurE. The *Aa*PurE sequence is labeled for reference and identical residues are highlighted in red. This figure was prepared with *ESPript* (Gouet *et al.*, 1999).

residue). Of these interactions, 40 are hydrogen-bonding interactions and 59 are hydrophobic interactions. These values are in contrast to the 52 pairwise interactions determined for *Ec*PurE. In this alternate case, 20 interactions were hydrogen bonds and 32 were hydrophobic. In turn, *Tm*PurE has a total of 78 pairwise interactions, with 34 hydrogen-bonding and 44 hydrophobic interactions.

The calculated pI values from *ExPASy* (Wilkins *et al.*, 1999) for *Aa*PurE, *Ec*PurE and *Tm*PurE are 5.5, 6.0 and 6.3, respectively. However, surface-charge calculations performed at pH 7.0 with *SPOCK* (Christopher, 1998) show that *Aa*PurE has a more positively charged surface than either *Ec*PurE or *Tm*PurE (Fig. 2). However, *Aa*PurE contains fewer charged residues than *Ec*PurE and has a lower number of positively charged residues. *Tm*PurE has a higher percentage of both positively and negatively charged residues than either *Aa*PurE or *Ec*PurE.

3.4. Active site

The location of the *Aa*PurE active site was determined by comparison with a structure of *Ec*PurE bound to mononucleotide(s) derived from CAIR (given as AIR in PDB entry 1d7a). The residues in this region of the structure are highly conserved between the two species (Fig. 3). An advantitiously bound citrate molecule was present in the two active sites in the asymmetric unit. *Aa*PurE Asp33 interacts with the pro-(*R*) acetyl arm carboxylate in citrate, while the comparable residue in *Ec*PurE (Asp19) primarily contacts the 3′-OH of the bound mononucleotide. *Aa*PurE Ser29 interacts with the pro-(*S*) acetyl arm of citrate, while the comparable residue in *Ec*PurE (Ser16) contacts the mononucleotide phosphoryl group. *Aa*PurE His59 N^{δ} is involved in hydrogen bonding with the central carboxylate in citrate. In both monomers, this hydrogen bond appears to be ∼2.6 Å. Citrate has hydrogen-bonding interactions with two active-site waters that are absent from the *Ec*PurE–mononucleotide structure but are present in the unliganded *Ec*PurE structure.

The *Aa*PurE active site is at the interface of three subunits (Fig. 1*b*). Residues from the $\beta1$–$\alpha1$ loop and $\beta2$–$\alpha2$ loop compose two sides of the active site. The hydrophobic base of the active site is formed by the $\beta3$–$\alpha3$ loop. The remaining side of the active site is formed by the $\alpha4$–$\beta5$ loop from the neighboring subunit. Some potentially important interactions that help stabilize the upper portion of the active site involve hydrogen bonds between residues Pro174 and Arg64, Asn175 and Ile55, and Ile178 and Gln31.

4. Discussion

4.1. Comparison of *Aa*PurE with *Ec*PurE and *Tm*PurE

*Aa*PurE, *Ec*PurE and *Tm*PurE all crystallized in space group *I*422; however, *Aa*PurE had two monomers per asymmetric unit, while *Ec*PurE and *Tm*PurE both have one monomer per asymmetric unit. Unit-cell parameters were as

follows: a = 113.04, c = 49.41 Å for *Ec*PurE, a = 103.25, c = 65.45 Å for *Tm*PurE (Mathews *et al.*, 1999; Schwarzenbacher *et al.*, 2004) and a = 99.25, c = 164.81 Å for *Aa*PurE. *Ec*PurE and *Tm*PurE pack in the crystal in a very similar manner. *Aa*PurE packs slightly differently such that everywhere that there was a single octomer in the previous cases there are two octamers. This leads to twice the number of molecules in the asymmetric unit and a longer c axis. The overall structures of *Aa*PurE and *Ec*PurE are nearly identical as shown by the structural alignment (Fig. 4). The greatest differences between the two structures occur at the ends of the protein.

*Aa*PurE has 14 more N-terminal residues than *Ec*PurE and 20 more than *Tm*PurE. Protein sequencing confirmed that the N-terminus is present in the *Aa*PurE batch used for crystallizations. However, all three models are missing residues at the N-terminus (19, six and one, respectively, not counting a 12-residue affinity fusion in *Tm*PurE). The disordered N-termini of *Aa*PurE and *Ec*PurE contain many polar residues.

In *Aa*PurE, the C-terminal α-helix has one turn fewer than the structurally similar α-helix in *Ec*PurE. This final helix is terminated by Pro174 in *Aa*PurE and Pro163 in *Ec*PurE. A few residues away from the helix-terminating proline is a second proline (*Aa*PurE Pro177) that is buried in a hydrophobic pocket that includes an absolutely conserved Met14 and a partially conserved Trp20. This proline may increase the stability of α6 by binding tightly to the core of the protein. The shorter C-terminal helix in *Aa*PurE allows increased interaction between the residues Asp175, Pro177 and Ile178 from the loop following α6 and residues Gln31 and Ile65 from α1 and β2, respectively. Additional interactions that stabilize β2 may also stabilize the β2–α2 loop which is directly involved in substrate binding. Since the final four amino acids in *Aa*PurE are disordered, it is unclear whether the shorter C-terminus of *Aa*PurE might increase the overall solvent exposure of the active-site region. (The C-terminus of *Ec*PurE is completely ordered.)

*Tm*PurE, on the other hand, has an extra 1.5 turns in the structurally similar C-terminal helix. Leu158 of *Tm*PurE is located in the conserved hydrophobic pocket where Pro177 in *Aa*PurE is buried. There is also an additional two-turn C-terminal helix in *Tm*PurE that is missing from the other two structures.

4.2. Active-site comparisons

The active sites of *Aa*PurE and *Tm*PurE are nearly identical to that of *Ec*PurE. The three loops primarily implicated in substrate binding in *Ec*PurE are present in nearly identical conformations in *Aa*PurE: the P loop (*Ec*PurE Asp19 and Ser16 correspond to *Aa*PurE Asp33 and Ser30), the '40s loop' (SAHR; *Ec*PurE residues 43–46, *Aa*PurE residues 57–60) and the '70s loop' (AGGAAHLP; *Ec*PurE residues 70–77, *Aa*PurE residues 84–91). A small hydrophobic pocket within the active site that has been proposed to facilitate CO_2 transfer (Mathews *et al.*, 1999; Meyer *et al.*, 1999) is defined by Ala84 and Leu90 from the 70s loop and Ala58 from the 40s loop (*Ec*PurE residues 70, 76 and 44, respectively).

Of the ∼19 residues lining the active site, only two positions are different. *Aa*PurE Gly126 and Lys141′ (the prime indicates donation from a neighboring subunit; the same amino-acid pair is present in *Tm*PurE) replace *Ec*PurE Arg112–Ala127′. The side chain of Lys141′ occupies the same position as *Ec*PurE Arg112 and is similarly charged. This change might alter the orientation of the phosphoryl group in a bound mononucleotide.

The second residue change is within the P loop. *Ec*PurE Lys17 and the corresponding *Aa*PurE residue Gln31 face away from the active site. However, *Tm*PurE has Asp11 at this position, which is involved in an apparently stabilizing salt-bridge network with C-terminal helix residues Lys154 and Lys170.

4.3. Proton transfers in PurE

The current structure was solved at a slightly acidic pH that is well within the normal *A. aceti* pH_{in} range. While the three structurally characterized class I PurEs are from organisms that represent a wide range of temperature and pH optima, the similarity of the active-site regions indicates a common enzymatic mechanism.

The intramolecular carboxylate transfer catalyzed by PurE is thought to involve a transient PurE–AIR–CO_2 intermediate, a complex that could be formed in both class I and class II PurEs (Meyer *et al.*, 1999). Consistent with this idea is the observation of a small hydrophobic pocket within the active site that might serve to sequester nascent CO_2

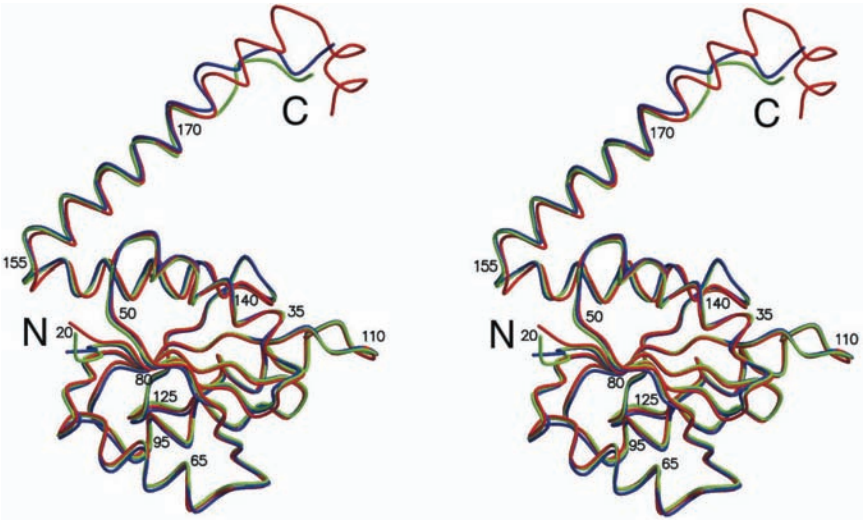

Figure 4
Stereoview of three monomers of PurE superimposed in coil representation. Each is colored differently: *Aa*PurE in green, *Ec*PurE blue and *Tm*PurE red. *Aa*PurE residues are labelled for reference. This figure was prepared with *MOLSCRIPT* (Kraulis, 1991) and *RASTER3D* (Merritt & Bacon, 1997).

(Mathews *et al.*, 1999). This 'decarboxylation–recarboxylation' mechanism for class I PurE requires accompanying protonation of N5 and deprotonation of C4 in N^5-CAIR by an unknown active-site acid/base. The reverse reaction requires protonation of C4 and deprotonation of N5 in CAIR. In neutral solutions at 303 K, the proton on the C4 in AIR (or AIR riboside) exchanges in deuterium oxide with a half-life of 8 min, consistent with its proposed role in nucleophilic attack on CO_2 (Mathews *et al.*, 1999; Schendel, 1986). In the forward reaction, class II PurE binds AIR and CO_2, which carry at least one more proton than N^5-CAIR. It is plausible the mechanism of proton transfer and the active site differ between class I and class II PurEs. The strict differential conservation of the 70s loop highlights it as the most likely region for a functional difference.

The chemical instability of N^5-CAIR has hampered efforts to obtain a structure of a PurE–substrate complex, which would illuminate the mechanism of carboxylate and proton transfer. Cocrystallization of *Ec*PurE with CAIR yielded a complex that may contain at least one AIR-derived mononucleotide (N^5-CAIR, CAIR and/or AIR $\pm$ CO_2). The mononucleotides were only present in four active sites on one face of the *Ec*PurE octamer (Mathews *et al.*, 1999). In this complex, N^{δ} of *Ec*PurE His45 (*Aa*PurE His59) was 3.2 Å from C4 in AIR, suggesting a possible role in proton transfer.

In the current structure of *Aa*PurE bound to citrate, a hydrogen bond links a citrate carboxylate and His59, which is likely to be positively charged in at least one of the two active sites found in the asymmetric unit. This is the first evidence for ionization of the universally conserved His59 by an active-site ligand. In the 70s loop of class I PurEs, there is an almost entirely conserved histidine (*Aa*PurE His89, *Ec*PurE His75) that is in direct contact with the active-site His59. In addition, His89 lies near a molecular twofold axis such that His89 from one subunit contacts His89 from the neighboring subunit (Fig. 5). Hence, there is the potential for communication between active sites in class I PurE, relayed through four histidine residues. These histidine residues are surrounded by highly conserved hydrophobic groups, including the absolutely conserved Met124 and Pro125. Arg60 and a bound mononucleotide would complete the isolation of the histidine chain from solvent. (The 70s loop of class II PurEs lacks a His89 equivalent, but it does contain three strictly conserved polar residues.)

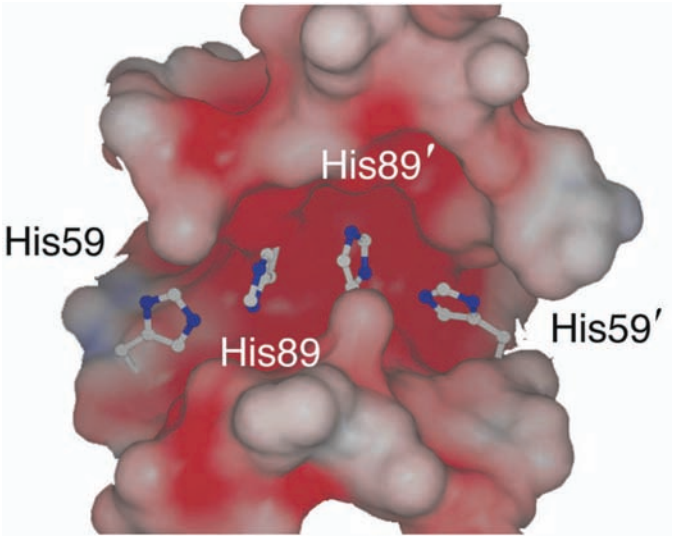

Figure 5
The chain of histidines linking two active sites are drawn in ball-and-stick representation and labeled with a prime to indicate donation from a neighboring subunit. A surface representation of the surrounding residues is shown with the opening created by cutting off one side of the channel. The coloring is based on charge from $-25kT$ to $25kT$. This figure was prepared with *SPOCK* (Christopher, 1998).

Table 2
Comparison of *Aa*PurE, *Tm*PurE and *Ec*PurE.

	*Aa*PurE	*Tm*PurE	*Ec*PurE
Polar accessible surface area (%)	33	44	46
Hydrogen bonds			
Inter-subunit	138	152	124
Intra-subunit	40	34	23
Salt bridges			
Inter-subunit	6	8	4
Intra-subunit	4	8	6
Salt bridges containing arginine (%)	100	37.5	0

One possible function of the histidine chain is to support a polar contact with the carboxylate moiety of N^5-CAIR or CAIR. This is inconsistent with a mechanistic proposal in which unfavorable non-polar interactions between the carboxylate moiety and the PurE hydrophobic pocket favor formation of a PurE–AIR–CO_2 intermediate (Meyer *et al.*, 1999). Another possibility consistent with the mechanistic proposal is that the chain of histidines serves as a proton shuttle and/or to signal between two active sites. When only one active site is occupied, His89 in the other subunit would be exposed to solvent at the other end of the buried histidine chain. This might allow proton transfers in and out of the active site without exposing the bound nucleotide base and/or CO_2 to solvent. Alternately, His59 ionization might trigger a conformational change in the other active site. Transfer of anything larger than a proton (*e.g.* CO_2) between active sites would appear to require a substantial reorganization of the protein.

4.4. Comparison of *Aa*PurE with other acid-stable and thermostable proteins

A. aceti can maintain a suprisingly low pH_{in} and its cytoplasmic enzymes are adapted to tolerate low pH (T. J. Kappock and coworkers, unpublished observations). *Aa*PurE is the first structurally characterized enzyme from this unusual acidophile. Structurally characterized acid-stable enzymes tend to be secreted proteins, such as xylanase C from *Aspergillus kawachi*. Relative to a comparable *Bacillus circulans* xylanase, it has a lower pH optimum and more negatively charged surface residues, which are localized to a single area (Fushinobu *et al.*, 1998).

A useful analogy to protein acid stability is protein thermostability, for which a large amount of structural data is available. Optimized polar/electrostatic interactions, such as

increased polar surface area and a greater number of hydrogen-bonding interactions, are most frequently proposed to confer thermostability (Kumar & Nussinov, 2004; Sterner & Liebl, 2001; Vieille & Zeikus, 2001; Vogt & Argos, 1997). Since *A. aceti* and *E. coli* prefer similar growth temperatures, structural differences in *Aa*PurE and *Ec*PurE can be attributed to acid-specific stabilization, where differences in *Ec*PurE and *Tm*PurE reveal temperature-specific stabilization (Schwarzenbacher *et al.*, 2004). As expected from the high degree of sequence identity, the active-site regions of all three PurEs are the same and it is mainly the amino acids on surfaces that are not involved in the active site that differ.

Polar surface residues contribute energetically favorable hydrogen bonds to water. The *Aa*PurE octamer has ∼33% solvent-accessible polar regions, while *Ec*PurE and *Tm*PurE have 44 and 46%, respectively. *Ec*PurE has the greatest proportion of negatively charged surface residues (Fig. 2). Thus, it appears that *Aa*PurE is not stabilized by additional solvent interactions.

Increasing the number of hydrogen bonds and/or salt bridges can 'stitch together' and thereby stabilize a protein. *Aa*PurE has more intra-subunit hydrogen bonds than *Ec*PurE, but fewer than *Tm*PurE (138, 124 and 152, respectively). However, *Aa*PurE has more inter-subunit hydrogen bonds than *Ec*PurE or *Tm*PurE (40, 23 and 34, respectively). These extra hydrogen bonds may increase the stability of *Aa*PurE and *Tm*PurE (Table 2).

There are 58 salt bridges in both the *Ec*PurE and *Aa*PurE octamers, but the location and the types of salt bridges differ. Each monomer of *Ec*PurE has six intra-subunit and four inter-subunit salt bridges. In addition, eight of these salt bridges include a lysine or a histidine. The situation differs in the monomer of *Aa*PurE, in which there are four intra-subunit salt bridges and six inter-subunit salt bridges, all of which involve an arginine. Most of the *Aa*PurE salt bridges involve residues that have hydrogen bonds that are both shorter and have near-ideal geometry compared with those from *Ec*PurE. *Tm*PurE contains the largest number of salt bridges (96 in total), with eight intra-subunit interactions and eight inter-subunit interactions. Six interactions include an arginine and ten include either a lysine or histidine.

Overall, the acid-stability of *Aa*PurE appears to relate to increased hydrogen bonding and improved salt-bridge formation and more of these stabilizing forces are present between subunits. Interestingly, these are several of the structural features ascribed to thermostable proteins. For instance, a gain of 11.7 hydrogen bonds accounts for a 10 K rise in thermostability (Vogt & Argos, 1997). Perhaps there is a common set of general rules for increasing the stability of a protein in hot or acidic environments.

5. Conclusions

PurE from the acidophile *A. aceti* shows a greater number of intra-subunit and inter-subunit hydrogen bonds than its mesophilic homolog from *E. coli*. In addition, the salt bridges from *Aa*PurE have better geometry and exclusively involve arginines, whereas those from *Ec*PurE almost entirely involve lysines or histidines. The acid stability of *Aa*PurE may relate to the general increase in intra-subunit stability afforded by both increased hydrogen bonds and stronger salt bridges. The additional thermostability of *T. maritima* PurE may relate to increased hydrogen bonding, a greater number of salt bridges and increased polar surface area.

This work was supported by grant RR-15301 from the National Institutes of Health (SEE) and the Herman Frasch Foundation (TJK). SEE is indebted to the W. M. Keck Foundation and the Lucille P. Markey Charitable Trust. Rena Goodman and Julie Francois are thanked for help in the early stages of this project. We thank Leslie Kinsland for assistance in the preparation of this manuscript.

References

Asai, T. (1968). *Acetic Acid Bacteria. Classification and Biochemical Activities.* University of Tokyo Press.
Botuyan, M. V., Toy-Palmer, A., Chung, J., Blake, R. C. II, Beroza, P., Case, D. A. & Dyson, H. J. (1996). *J. Mol. Biol.* **263**, 752–767.
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. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Chen, Z. D., Dixon, J. E. & Zalkin, H. (1990). *Proc. Natl Acad. Sci. USA*, **87**, 3097–3101.
Christopher, J. A. (1998). Texas A&M University, College Station, Texas, USA.
Firestine, S. M., Misialek, S., Toffaletti, D. L., Klem, T. J., Perfect, J. R. & Davisson, V. J. (1998). *Arch. Biochem. Biophys.* **351**, 123–134.
Firestine, S. M., Poon, S. W., Mueller, E. J., Stubbe, J. & Davisson, V. J. (1994). *Biochemistry*, **33**, 11927–11934.
Foster, J. W. (1999). *Curr. Opin. Microbiol.* **2**, 170–174.
Fushinobu, S., Ito, K., Konno, M., Wakagi, T. & Matsuzawa, H. (1998). *Protein Eng.* **11**, 1121–1128.
Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). *Bioinformatics*, **15**, 305–308.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kappock, T. J., Ealick, S. E. & Stubbe, J. (2000). *Curr. Opin. Chem. Biol.* **4**, 567–572.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Kumar, S. & Nussinov, R. (2004). *Chembiochem*, **5**, 280–290.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Mathews, I. I., Kappock, T. J., Stubbe, J. & Ealick, S. E. (1999). *Structure Fold. Des.* **7**, 1395–1406.
Menzel, U. & Gottschalk, G. (1985). *Arch. Microbiol.* **143**, 47–51.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Meyer, E., Kappock, T. J., Osuji, C. & Stubbe, J. (1999). *Biochemistry*, **38**, 3012–3018.
Meyer, E., Leonard, N. J., Bhat, B., Stubbe, J. & Smith, J. M. (1992). *Biochemistry*, **31**, 5022–5032.
Mueller, E. J., Meyer, E., Rudolph, J., Davisson, V. J. & Stubbe, J. (1994). *Biochemistry*, **33**, 2269–2278.
Ohmori, S., Masai, H., Arima, K. & Beppu, T. (1980). *Agric. Biol. Chem.* **44**, 2901–2906.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Schendel, F. J. (1986). PhD thesis, University of Wisconsin, Madison, Wisconsin, USA.

Schwarzenbacher, R. *et al.* (2004). *Proteins*, **55**, 474–478.
Sterner, R. & Liebl, W. (2001). *Crit. Rev. Biochem. Mol. Biol.* **36**, 39–106.
Vieille, C. & Zeikus, G. J. (2001). *Microbiol. Mol. Biol. Rev.* **65**, 1–43.
Vogt, G. & Argos, P. (1997). *Fold. Des.* **2**, S40–S46.
Walter, R. L., Ealick, S. E., Friedman, A. M., Blake, R. C. II, Proctor, P. & Shoham, M. (1996). *J. Mol. Biol.* **263**, 730–751.
Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D. & Hochstrasser, D. F. (1999). *Methods Mol. Biol.* **112**, 531–552.