Supporting information for:

"Biochemical and structural studies of N^5-carboxyaminoimidazole ribonucleotide mutase (PurE) from the acidophilic bacterium *Acetobacter aceti*"

*Charles Z. Constantine, Courtney M. Starks, Christopher P. Mill, Aaron E. Ransome, Steven J. Karpowicz, Julie A. Francois, Rena A. Goodman, and T. Joseph Kappock**

Department of Chemistry, Washington University in St. Louis, St. Louis, MO 63130-4899

EXPERIMENTAL

PurK activity assays. PurK activity assays were performed at 37°C using a published method (*1*).

Partial purification of A. aceti PurK. For PurK isolations it was necessary to repress "leaky" expression of *Aa*PurE by supplementing pJK174 starter cultures with 1% glucose. The subsequent steps in the *Aa*PurE isolation were performed as described. The supernatant from the 35% ammonium sulfate (AS) precipitation was adjusted to 60% saturation by the addition of solid AS over 30 min. After another 30 min stirring, solids were collected by centrifugation at 37,000 x *g* for 15 min and the supernatant was discarded. The pellet was redissolved in a small volume of 25 mM MOPS pH 6.5 and desalted using a Sepharose G25 column (2.5 x 13 cm) equilibrated in the same buffer. Fractions containing protein were applied to a QAE cellulose FF column (1.5 x 7.5 cm), which was washed with 25 mM MOPS pH 6.5 (5 column volumes, C.V.) and developed in a linear gradient of KCl (0 to 0.4 M, 20 C.V.) in the same buffer. Pooled protein-containing fractions were concentrated and partially desalted by several cycles of dilution in 50 mM Tris•Cl pH 8.0 followed by ultrafiltration using a YM10 membrane. Protein was then applied to a DEAE Sepharose FF column (1.5 x 8 cm), which was washed with 50 mM Tris•Cl pH 8.0 (6 C.V.) and developed in a linear gradient of KCl (0 to 0.5 M, 20 C.V.) in the same buffer. Fractions containing protein were concentrated by ultrafiltration to >5 mg/mL. Aliquots were frozen and stored at -80 °C. Aliquots were thawed only once before discarding.

Isolation of H6AaPurK. BL21(DE3) cells transformed with pJK130 were grown at 37 °C in LB supplemented with 70 μg/mL kanamycin. At A_{600} = 0.6, IPTG was added to 0.1 mM and the cells were grown another 4 hr. Cells were harvested by centrifugation, resuspended in 50 mM Tris•Cl pH 8.0 (5 mL/g cells), and disrupted by sonication (three cycles, 30 s apiece). All subsequent steps were performed at 4 °C in 50 mM Tris•Cl pH 8.0. Debris was removed by centrifugation (3,080 x *g*, 15 min). The supernatant was adjusted to 30% saturation by the addition of solid AS over 30 min. After another 30 min stirring, solids were removed by centrifugation (27,000 x *g*, 25 min) and the supernatant was adjusted to 50% saturation by the addition of solid AS over 30 min. After another 30 min stirring, solids were collected by centrifugation, dissolved in Tris buffer (5 mL/g cell paste), and applied to a Ni^{2+}-charged iminodiacetic acid Sepharose FF column (NiIDA, 2.5 x 10 cm). The column was washed in Tris buffer (5 C.V.) and stripped with Tris buffer containing 50 mM EDTA and 100 mM KCl (10 C.V.). (Using an

imidazole gradient to elute H6*Aa*PurK from the NiIDA column caused slow protein aggregation.) Blue *Aa*PurK-containing fractions were pooled (110 mL), concentrated by ultrafiltration over a YM10 membrane (15 mL), and applied to a DEAE Sepharose FF column (1.5 x 8 cm) equilibrated in Tris buffer. The column was developed in a linear gradient of KCl (0 to 0.5 M, 20 C.V.) in Tris buffer. Fractions containing protein were pooled and concentrated by ultrafiltration over a YM10 membrane to >10 mg/mL. Single-use aliquots were frozen and stored at -80 °C.

Determination of Δε values at other pH values. Extinction coefficents are known at 260 nm at pH 6, 7, and 8 for AIR and CAIR, which allows the computation of Δε values used in PurE activity assays (*2*). CAIR solutions were standardized using endpoint assays at pH 8, in which small aliquots of CAIR (<0.1 mM) were added to a large amount of *Aa*PurE (typically 66 μg, 2.5 U at pH 8), using ΔA_{260} of the resultant burst to compute the concentration of the stock solution. A standardized CAIR solution was subjected to replicate endpoint assays at 30 °C and pHs ranging from 4.5 - 9 to determine Δε values at intermediate pHs (Table S2). The buffer conditions used were the same as in the construction of pH-rate profiles. In all cases, the bursts used for quantitation appeared to be monophasic and reached stable endpoints within 2 min. Assuming some decarboxylation of N^5-CAIR to AIR occurs in this period, especially at the lower pH values, the flat endpoint implies there is no substantial difference in ε_{260} values for N^5-CAIR and AIR over the pH range used.

RESULTS

Recombinant AaPurK isolation and characterization. *Aa*PurK was cleanly separated from *Aa*PurE by a 35% AS precipitation step, but attempts to purify *Aa*PurK from the 35% AS supernatant yielded only ~50% pure protein with specific activity 25 U/mg at 37 °C. A form with a 20 amino acid N-terminal fusion, H6*Aa*PurK, was purified to homogeneity in two steps but with lower specific activity (21 U/mg). On SDS-PAGE, *A. aceti* H6*Aa*PurK migrates with M_r = 42,000 (Figure S1). ESI-MS showed 43,918 ± 3 Da, consistent with the 43,918.3 Da expected for the fusion protein lacking one Met residue, presumably due to removal of Met1. Gel filtration shows a single peak at 88 kDa, showing it is exclusively a dimer. *Ec*PurK has specific activity 55 U/mg at 37 °C and is a dimer of 39.5 kDa subunits (*3*).

SUPPORTING INFORMATION REFERENCES

1. Firestine, S. M., Misialek, S., Toffaletti, D. L., Klem, T. J., Perfect, J. R., and Davisson, V. J. (1998) Biochemical role of the *Cryptococcus neoformans* ADE2 protein in fungal *de nov*o purine biosynthesis, *Arch. Biochem. Biophys. 351*, 123-134.

2. Mueller, E. J., Meyer, E., Rudolph, J., Davisson, V. J., and Stubbe, J. (1994) N^5-carboxyaminoimidazole ribonucleotide: Evidence for a new intermediate and two new enzymatic activities in the *de novo* purine biosynthetic pathway of *Escherichia coli*, *Biochemistry 33*, 2269-2278.

3. Thoden, J. B., Kappock, T. J., Stubbe, J., and Holden, H. M. (1999) Three-dimensional structure of N^5-carboxyaminoimidazole ribonucleotide synthetase: A member of the ATP grasp protein superfamily, *Biochemistry 38*, 15480-15492.

4. Meyer, E., Leonard, N. J., Bhat, B., Stubbe, J., and Smith, J. M. (1992) Purification and characterization of the *pur*E, *pur*K, and *pur*C gene products: identification of a previously unrecognized energy requirement in the purine biosynthetic pathway, *Biochemistry 31*, 5022-5032.

5. Mathews, I. I., Kappock, T. J., Stubbe, J., and Ealick, S. E. (1999) Crystal structure of *Escherichia coli* PurE, an unusual mutase in the purine biosynthetic pathway, *Structure 7*, 1395-1406.

6. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions, *Protein Eng. 8*, 127-134.

Table S1. Oligodeoxynucleotides (ODNs) used in this study.[a]

ODN#	Sequence (5' → 3')
408	GAGCTCGAATTCTTAACTGGCAGATTCTTCAT
410	GATATACATATGACCTCCTTACATTCTCC
614	CCCACCCATATGATGAGCGAAACCGCCC
615	AGGAGGGAATTCTCATTTATCTTCGGTAATG
686	biotin-CTCTTAAAGGCATGGACAGCCTGCTTTCCA
1002	GTGGAGCTGCAAATTTGCCGGGTATG
1003	CATACCCGGCAAATTTGCAGCTCCAC
1011	GTGGAGCTGCAGATCTGCCGGGTATG
1012	CATACCCGGCAGATCTGCAGCTCCAC
1013	GTGGAGCTGCATTTCTGCCGGGTATG
1014	CATACCCGGCAGAAATGCAGCTCCAC
1015	CATTGTTTCAGCATTTCGTACGCCAGAC
1016	GTCTGGCGTACGAAATGCTGAAACAATG
1017	CATTGTTTCTGCAGATCGTACGCCAGAC
1018	GTCTGGCGTACGATCTGCAGAAACAATG
1019	CATTGTTTCAGCAAATCGTACGCCAGAC
1020	GTCTGGCGTACGATTTGCTGAAACAATG
1062	GTGGAGCTGCAGTTCTGCCGGGTATG
1063	CATACCCGGCAGAACTGCAGCTCCAC
1064	GTGGAGCTGCAGGTCTGCCGGGTATG
1065	CATACCCGGCAGACCTGCAGCTCCAC
1066	CATTGTTTCAGCACAACGTACGCCAGAC
1067	GTCTGGCGTACGTTGTGCTGAAACAATG
1068	CATTGTTTCAGCAGCTCGTACGCCAGAC
1069	GTCTGGCGTACGAGCTGCTGAAACAATG
1070	CATTGTTTCAGCAAGTCGTACGCCAGAC
1071	GTCTGGCGTACGACTTGCTGAAACAATG
1078	CATTGCCTGAATTCCTCCGTTACAGAC
1079	GCATTGGGCTTAAGCTTAGAATGTAAG

1161	CGGATGACGTTGAGCTGATTCATTTTCC
1162	GAAGCCGTGGTAAACCCGCGTGGATTGG

[a] ODNs 1002 - 1071 were used for QuikChange mutagenesis of *Aa*PurE. The other ODN functions are described in the text.

Table S2. Carboxylic acids as potential *Aa*PurE ligands at low pH.

Compound	Counterion[c]	Final pH	K_d (mM)[d]	ΔF[e]
citrate[a]	sodium	5.5	0.19 ± 0.02	D
DL-isocitrate[a]	sodium	5.6	0.8 ± 0.2	D
OAA[a]	-	5.41	>3	D
fumaric acid[a]	-	4.90	>5	D
succinic acid[a]	-	4.98	>10	N
α-ketoglutarate[a]	sodium	5.16	>5	N
L-glutamate[a]	sodium	5.43	>30	N
L-lactate[a]	-	5.43	>20	N
cis-aconitate[a]	-	5.45	~1	I
trans-aconitate[a]	-	5.44	~1	D
DL-aspartate[a]	-	5.5	>1.5	D
DL-2-fluorocitrate[b]	potassium	4.16	0.53 ± 0.04	D
"	"	4.78	0.70 ± 0.06	"
"	"	5.39	1.0 ± 0.1	"
DL-malate[b]	-	5.23	2.8 ± 1.0	D
L-malate[b]	-	5.23	3.8 ± 0.6	D

[a] Buffer was 50 mM ammonium acetate, 100 mM potassium chloride

[b] Buffer was 50 mM potassium acetate, 100 mM potassium chloride

[c] Commercially available or prepared salt used to make titrant solution, (-)indicates free acid used.

[d] From fluorescence emission titrations.

[e] Effect on *Aa*PurE fluorescence: D, decrease; I, increase; N, no change/no binding.

Table S3. Additional crystallographic data and refinement statistics.

	H59N (pH 8)	H59F (pH 8)	H89F (pH 8)	H89G (pH 8)	H89N (pH 7)
PDB id (our code)	2fw7	2fw9	2fwb	2fw8	2fwa
Cell dimensions	a = b = 99.3 Å; c = 164.1 Å; α=β=γ=90°	a = b = 99.4 Å; c = 164.7 Å; α=β=γ=90°	a = b = 99.7 Å; c = 163.9 Å; α=β=γ=90°	a = b = 99.6 Å; c = 164.4 Å; α=β=γ=90°	a = b = 99.2 Å; c = 164.2 Å; α=β=γ=90°
Resolution, Å	50 – 1.75 (1.81 – 1.75)[a]	50 – 1.75 (1.81 – 1.75)	40 – 2.0 (2.07 – 2.00)	50 – 1.75 (1.81 – 1.75)	30 – 1.9 (1.97 – 1.90 Å)
Reflections (total/unique)	240,174 / 38,755	239,521 / 40,968	106,542 / 27,131	317,563 / 40,066	153,109 / 31,276
Completeness	92.9% (73.1%)	98.0% (93.7%)	95.8% (97.0%)	95.3% (78.9%)	95.9% (99.0%)
$\langle I/\sigma \rangle$	30.5 (3.4)	30.1 (2.8)	16.6 (2.8)	36.3 (5.4)	17.2 (3.9)
R_{sym}[b]	4.9% (40.5%)	4.7% (43.8%)	7.1% (48.1%)	4.2% (29.0%)	9.1% (43.9%)
R_{cryst}[c]/R_{free}[d]	17.4% / 18.8%	17.2% / 18.8%	17.8% (20.6%)	17.4% / 19.7%	18.8% / 20.7%
No. of protein atoms	2,314	2,305	2,328	2,300	2,314
No. of water molecules	374	397	308	378	276
No. of ligand ions	0	1	0	0	0
R.m.s. deviation, bond lengths (Å)	0.005	0.004	0.005	0.005	0.005
R.m.s. deviation, bond angles	1.3°	1.3°	1.3°	1.3°	1.3°
Average B-factor ($Å^2$)	20.7	20.9	23.8	20.5	19.8

[a]Numbers in parentheses refer to highest resolution shell.

[b]$R_{sym} = \Sigma|I_h - \langle I_h \rangle|/\Sigma I_h$, where $\langle I_h \rangle$ is the average intensity over symmetry.

[c] $R_{cryst} = \Sigma|F_o - \langle F_c \rangle|/\Sigma F_o$, where summation is over the data used for refinement.

[d] R_{free} is defined the same as R_{cryst}, but was calculated using 5% of data excluded from refinement.

Figures.

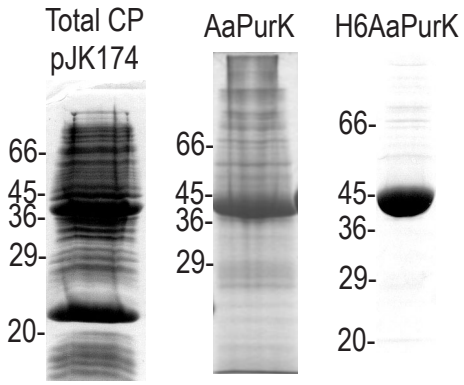


Figure S1. SDS-PAGE analysis of (*left*) total cell protein (CP) for pJK174, which expresses *Aa*PurE at 21 kDa and *Aa*PurK at ~40 kDa; (*center*) partially purified *Aa*PurK (specific activity 25 U/mg); and (*right*) purified H6*Aa*PurK (specific activity 21 U/mg). The positions of size standards (kDa) are indicated for each gel.

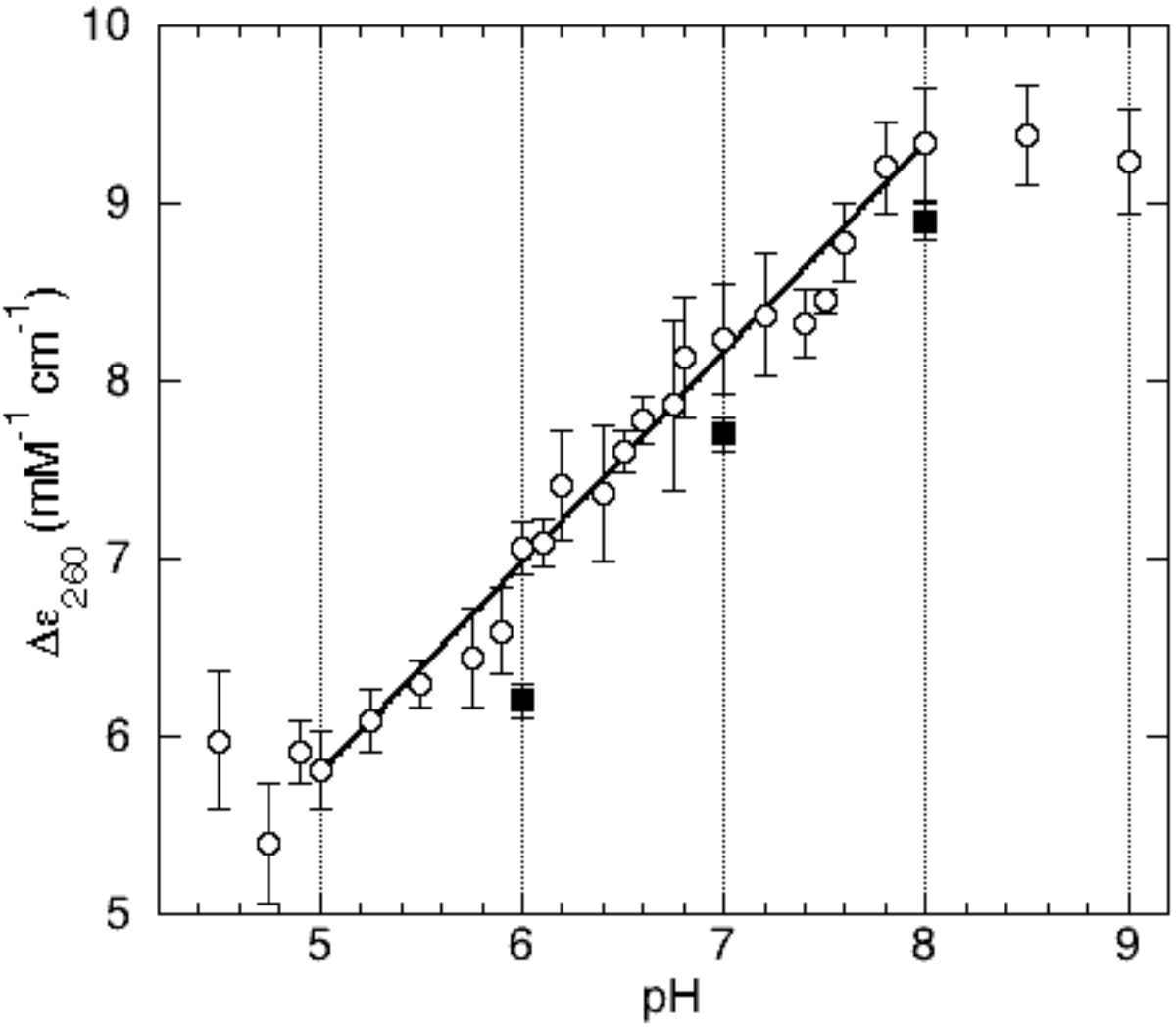
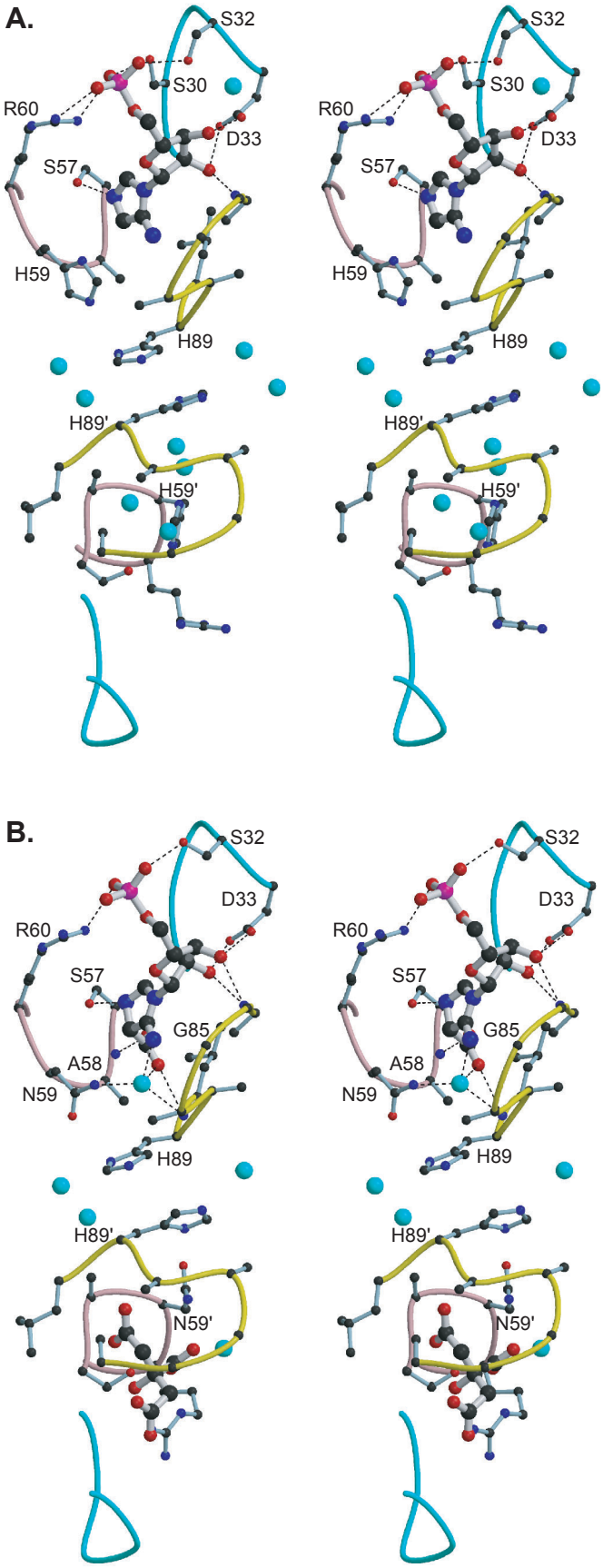


Figure S2. Determination of $\Delta\varepsilon_{260}$ for the conversion CAIR→N^5-CAIR as a function of pH. Mean values for each pH were determined in at least triplicate (open symbols) with the indicated standard deviation. The solid line is a fit of the data in the pH 5.0 - 8.0 interval, which yields the expression $\Delta\varepsilon_{260} = 1.19 \times \text{pH} - 0.15$. Outside this range, individual $\Delta\varepsilon_{260}$ values were determined. Previously reported values for $\Delta\varepsilon_{260}$ and uncertainties are shown for reference (solid squares) (*4*).


A.
S32
S30
R60
D33
S57
H59
H89
H89'
H59'
S32
S30
R60
D33
S57
H59
H89
H89'
H59'
B.
S32
D33
R60
S57
G85
A58
N59
H89
H89'
N59'
S32
D33
R60
S57
G85
A58
N59
H89
H89'
N59'

from previous page:

Figure S3. Divergent stereodiagrams for active sites in chains A (*top*) and D (*bottom*), presented to emphasize the location of His89 and waters located between active sites. Key loops are color coded as for *Ec*PurE (*5*): the P loop is in cyan, the forties loop is in pink, and the seventies loop is in yellow. Selected water O atoms are illustrated in cyan. Dotted lines represent contacts of ≤3.1 Å to the bound nucleotides. *A*, wt *Aa*PurE•AIR (2fwj). *B*, H59N-*Aa*PurE•isoCAIR•citrate (2fwp). Close contacts to water 165 are also shown in dotted lines.

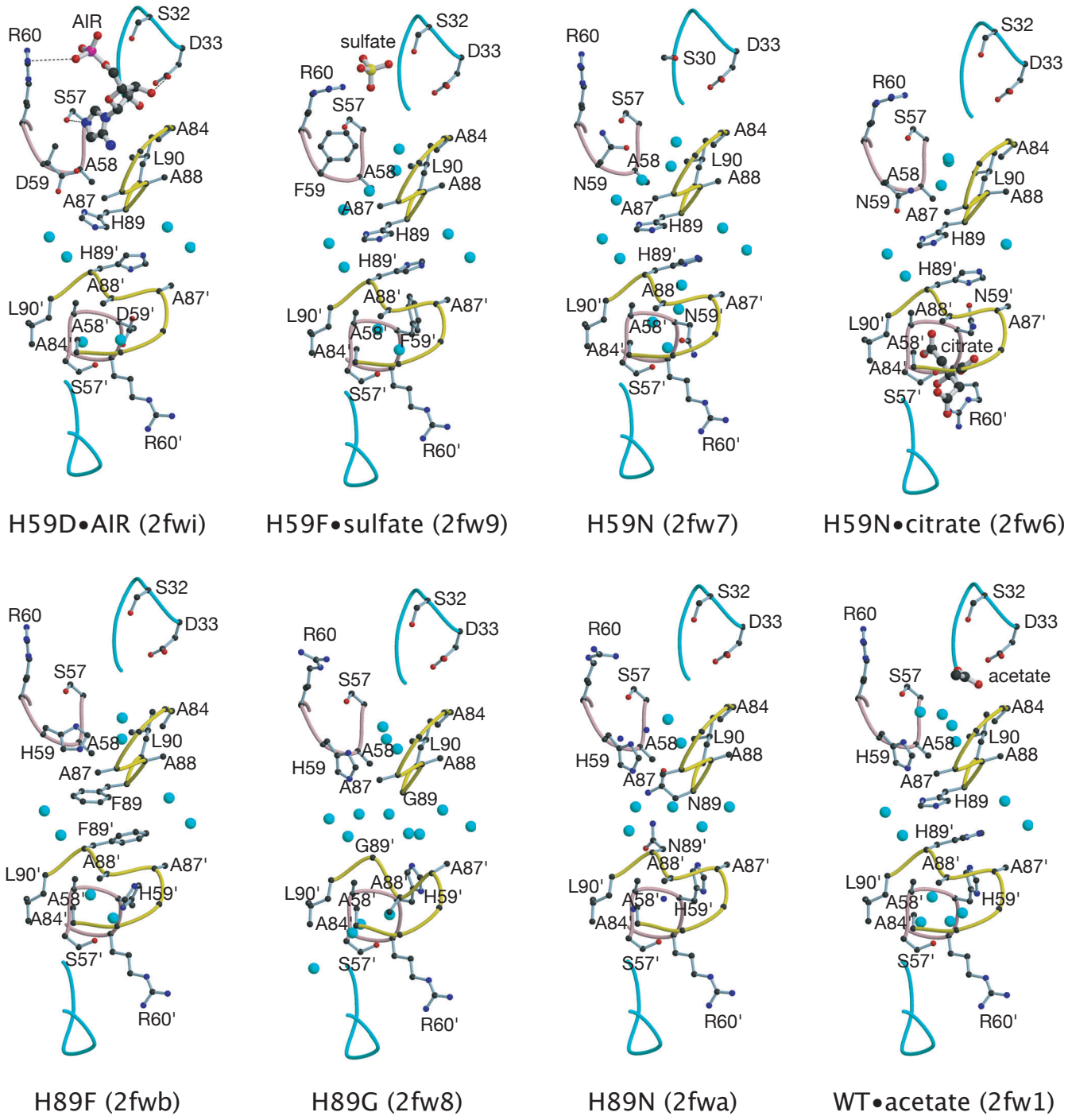


Figure S4. Views of the inter-active site region in the remaining eight *Aa*PurE crystal structures reported here. The viewpoint is the same as the average viewpoint in Figure S3 stereodiagrams. The color scheme is the same as in Figure S3.

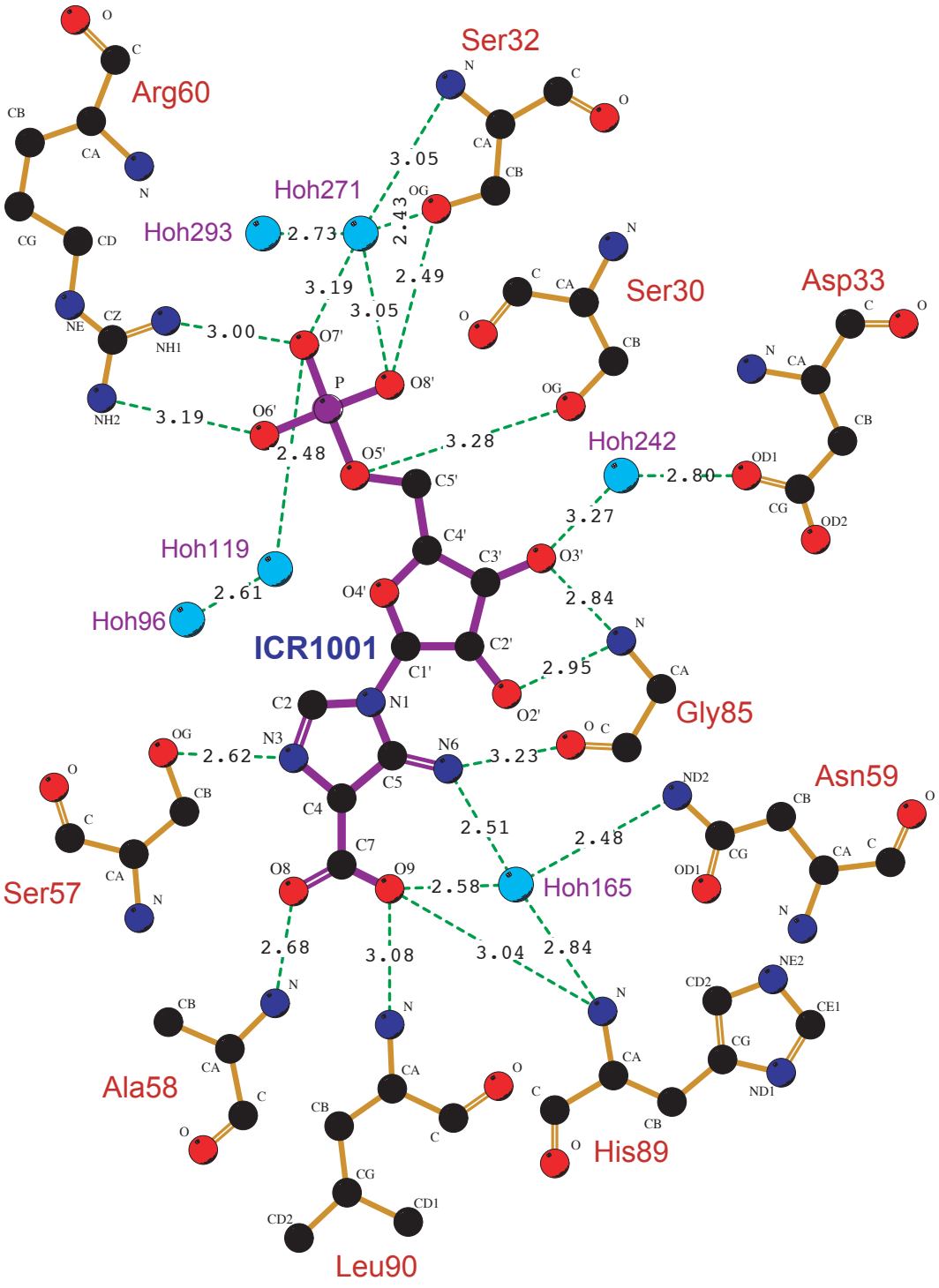


Figure S5. Contacts between isoCAIR (ICR1001), the active site, and waters observed in crystal structure 2fwp (H59N-*Aa*PurE•isoCAIR•citrate). Inferred double bonds are drawn. All illustrated residues are from subunit A. Distances (Å) are given as labels on the green dotted lines. This figure was made with LigPlot (*6*) and LigEd (R. Laskowski, unpublished program).