Supporting Information

Liszczak and Marmorstein 10.1073/pnas.1310365110

SI Materials and Methods

ssNAT Expression and Purification. The archaeal N-amino terminal acetyltransferase NAT from *Sulfolobus solfataricus* (ssNAT) gene (encoding residues 50–216) was cloned out of *S. solfataricus* genomic DNA (American Type Culture Collection) and engineered into a modified pETDUET vector containing a Tobacco Etch Virus protease cleavable 6×-His tag. The resulting plasmid was transformed into Rosetta (DE3)pLysS competent *Escherichia coli* cells, grown to an $OD_{600} = 0.7$ and induced with 0.5 mM isopropyl β-D-thiogalactoside at 16 °C for ∼16 h. Cells were isolated by centrifugation and lysed by sonication in a buffer containing 25 mM Tris at pH 8.0, 1 M NaCl, 10 mM 2-mercaptoethanol, and 0.5 mM phenylmethylsulfonyl fluoride (Sigma-Aldrich). The solution was isolated and passed over Ni-resin (Thermo Scientific), which was subsequently washed with >20 column volumes of lysis buffer supplemented with 25 mM imidazole (Fisher Scientific). The protein was eluted in the same buffer with an imidazole gradient (25–500 mM imidazole), and TEV protease was added to fractions containing the target protein for the duration of a 14-h dialysis into lysis buffer adjusted to 300 mM NaCl. The solution was passed through an additional nickel column to remove TEV protease as well as any uncut ssNAT. The resin was then washed with ∼7 column volumes of dialysis buffer supplemented with 25 mM imidazole, which was pooled with the initial flow-through. This solution was dialyzed into a buffer containing 25 mM Tris at pH 8.5, 50 mM NaCl, and 5 mM β-ME and loaded onto a 5 mL HiTrap Q ion exchange column (GE Healthcare). Although the protein did not stick to this column, a major impurity did interact with the column, and so this step was still essential for obtaining high-purity protein. Peak fractions were concentrated to ∼1–2 mL (10 kDa concentrator; Amicon Ultra, Millipore) and loaded onto an s75prep gel filtration column (GE Healthcare) that was pre-equilibrated in a buffer containing 20 mM Tris at pH 7.5, 150 mM NaCl, and 10% (vol/vol) glycerol. The second of two peaks from the sizing column corresponded to monomeric ssNAT, and so this peak was pooled and concentrated to 10 mg/mL, as measured by $ultraviolet_{280}$ absorbance (extinction coefficient = 25,000), for crystallization trials. All ssNAT mutants were generated using the Stratagene QuikChange protocol, and protein-harboring point mutations were expressed and purified using the protocol described earlier (1). All mutants exhibited gel filtration profiles similar to that of the wild-type enzyme, suggesting they were properly folded.

ssNAT Crystallization and Structure Determination. Before crystallization, ssNAT was mixed with AcCoA (Sigma-Aldrich) at a 1:3 molar ratio. Initial crystallization hits were obtained at 20 °C using hanging drop vapor diffusion with two different well conditions; the first containing 20% PEG 3350 (Hampton Research) and 0.2 M zinc acetate (Hampton Research) and the second containing 20% PEG 3350 and 0.02 M zinc chloride (Hampton Research). The zinc acetate condition showed better initial diffraction, and optimization of this condition resulted in a well containing 14% PEG 3350 and 0.1 M zinc acetate. The best diffracting crystals were from a drop set-up using a protein concentration of 7.5 mg/mL and took 3–7 d to reach maximum dimensions.

Diffraction data were collected on a Rigaku MicroMax-007 HF rotating anode X-ray generator with a Saturn 944+ CCD detector and processed to 1.98-Å resolution using HKL2000 (2). A total of 150° (1°/frame) were collected and indexed in the C2 space group. Molecular replacement was performed using the program Phaser and the apo structure of ssARD1 (Protein Data Bank ID 2x7B) as a search model (3). This solution yielded high-quality electron-density maps, and initial rounds of manual model building and refinement were performed using Coot (4). Subsequent rounds of refinement were performed using Phenix. Refine to refine XYZ coordinates and individual atomic displacement parameters (5). Iterative rounds of phenix refine and manual refinement in Coot were used to reach final R_{work} and R_{free} values of 19.9% and 24.3%, respectively (Table S1). The final map was checked for errors using a composite omit map generated by AutoBuild in the Phenix suite of programs (5).

All 3D alignment rmsds as well as graphics were generated in PyMol (Delano Scientific LLC).

Acetyltransferase Assays. Acetyltransferase assays were performed as previously described (6). All assays were performed in the presence of 500 nM ssNAT enzyme at 37 °C in a buffer containing 25 mM Tris at pH 8.0 and 50 mM NaCl. It is likely that the in vivo k_{cat} values for this enzyme are much higher than we report here because the enzyme is from a thermophilic archaeal species that thrives in environments that reach temperatures much greater than 37 °C. For Ser- amino-terminal peptide assays, reaction times were generally 30 min; however, catalytically deficient mutants were tested for up to 2 h. For Met- amino-terminal peptides assays, reaction times were generally 120 min; however, times ranged from 30 min to 6 h, depending on the effect of the mutant. Mutants that did not show activity levels at least 2-fold greater than background readings were considered inactive. A saturating amount (500 μM) of radiolabeled [^{14}C] acetyl-CoA (4 mCi/mmol; PerkinElmer Life Sciences) was used in all enzymatic reactions, and the substrate peptide concentration was varied to determine steady-state catalytic parameters.

The concentration of the Ser- amino-terminal peptide (NH_2-SASEAGVRWGRPVGRRRRP-COOH; GenScript) ranged from 25 to 1,000 μM. The first seven residues of this peptide correspond to the amino-terminal sequence of a threonyl-tRNA synthetase protein that has been shown to be 100% amino-terminally acetylated in yeast, and the C-terminal 12 residues form a basic tag that is required for compatibility with the filter paper binding assay used here.

The concentration of the Met- amino-terminal peptide (NH_2-MLGPEGGRWGRPVGRRRRP-COOH; GenScript) ranged from 50 to 1,200 μM. The first seven residues of this peptide correspond to the amino-terminal sequence of hnRNP F, which is the only known biologically relevant substrate for the NAA50 enzyme, and the C-terminal 12 residues are identical to that of the Ser- amino-terminal peptide.

In the assay, radiolabeled [^{14}C]AcCoA was mixed with the substrate peptide and allowed to incubate with enzyme in a 25 μL reaction volume. To quench the reaction, mixtures were cooled to 4 °C and spotted (20 μL) on P81 paper discs (Whatman), and the paper discs were immediately placed in wash buffer (10 mM Hepes at pH 7.5). Papers were left submerged in wash buffer under gentle agitation for a total of 15 min to remove any unreacted AcCoA. During this process, buffer was discarded and replaced with fresh wash buffer every 5 min. The papers were then dried with acetone and added to 4 mL scintillation fluid, and the signal was measured using a Packard Tri-Carb 1500 liquid scintillation analyzer. Background control reactions were performed in the absence of enzyme. Reactions were also performed in the absence of the substrate peptide to ensure that any possible signal resulting from autoacetylation was negligible. The background signal was accounted for when calculating catalytic parameters for the wild-type enzyme and mutants. All radioactive

count values were converted to molar units using a standard curve created with known concentrations of radioactive AcCoA added to scintillation fluid. The program GraphPad Prism version 5.01 was used for all data fitting to the Michaelis-Menten equation.

1. Braman J, Papworth C, Greener A (1996) Site-directed mutagenesis using double-stranded plasmid DNA templates. *Methods Mol Biol* 57:31–44.
2. Otwinowski Z, Minor W (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol* 276:307–326.
3. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
4. Emsley P, Cowtan K. (oref) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
5. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
6. Liszczak G, Arnesen T, Marmorstein R (2011) Structure of a ternary Naa50p (NAT5/SAN) N-terminal acetyltransferase complex reveals the molecular basis for substrate-specific acetylation. *J Biol Chem* 286(42):37002–37010.

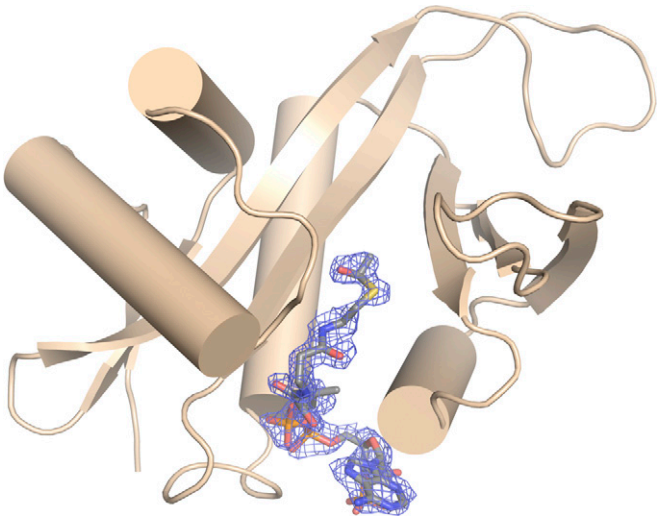

Fig. S1. Electron density (blue mesh) for AcCoA (gray stick) bound to ssNAT (tan cartoon) from a composite omit map (contoured to 1.5 σ).

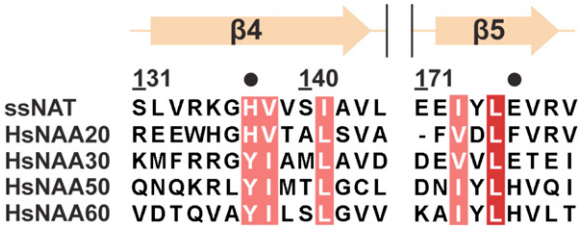

Fig. S2. An alignment of β4 and β5 strands from ssNAT with corresponding regions of all substrate amino-terminal Met- specific human catalytic NAT subunits. Red highlight represents strictly conserved residues and pink highlight represents residues with conserved identity. A black circle has been placed above key catalytic residues used by ssNAT for acetylation of Met- amino-terminal substrates. Residue numbers correspond to ssNAT.

Table S1. Data collection and refinement statistics for ssNAT

Crystal parameters	
Space group	C2
Unit cell dimensions	
a, Å; α, °	100.84, 90.00
b, Å; β, °	35.06, 100.05
c, Å; γ, °	51.22, 90.00
Data collection	
Resolution range, Å	21.05–1.98 (2.05–1.98)
Unique reflections	12,433
Redundancy	3.0 (2.5)
Completeness, %	99.5 (98.0)
R_{sym}, %	13.3 (46.0)
I/σI	7.7 (1.9)
Refinement statistics	
R_{work}/R_{free}, %	19.94/24.17
Number of atoms	
Protein	1289
AcCoA	51
Water	147
Zinc	2
rmsd	
bonds, Å	0.007
Angles, °	1.178
Average B-factor, $Å^2$	
Protein	20.9
AcCoA	21.8
Water	28.2
Zinc	33.9
Ramachandran statistics, %	
Favored	98.7
Allowed	1.3

Values in parentheses are for the highest resolution shell. R_{free} was calculated using 5% of the reflection data chosen randomly and omitted at the start of refinement.

Table S2. The k_{cat} values calculated for all ssNAT variants with the Ser- substrate peptide and the Met- substrate peptide

	Ser- peptide		Met- peptide	
ssNAT variant	min^{-1}	Normalized to WT	min^{-1}	Normalized to WT
WT	3.3 ± 0.4	1.0	0.73 ± 0.07	1.0
L82A	0.16 ± 0.02	0.048 ± 0.006	0.063 ± 0.005	0.087 ± 0.007
P83A	0.16 ± 0.02	0.048 ± 0.006	0.039 ± 0.004	0.054 ± 0.005
E84A	4.4 ± 0.4	1.3 ± 0.1	4.6 ± 0.4	6.4 ± 0.5
E84Q	1.8 ± 0.3	0.55 ± 0.09	1.4 ± 0.1	2.0 ± 0.1
E84V	3.0 ± 0.2	0.91 ± 0.06	2.5 ± 0.4	3.5 ± 0.5
Y86A	0.097 ± 0.010	0.029 ± 0.003	Activity not detected	
Y86F	3.3 ± 0.5	1.0 ± 0.2	0.23 ± 0.06	0.32 ± 0.08
H137A	1.3 ± 0.1	0.39 ± 0.03	Activity not detected	
H137F	1.4 ± 0.2	0.42 ± 0.06	Activity not detected	
H137Q	1.2 ± 0.3	0.36 ± 0.09	Activity not detected	
Y174A	3.4 ± 0.4	1.0 ± 0.1	0.47 ± 0.05	0.64 ± 0.07
Y174S	4.1 ± 0.4	1.2 ± 0.1	0.39 ± 0.04	0.54 ± 0.05
E176A	3.4 ± 0.3	1.0 ± 0.1	Activity not detected	
E176H	2.4 ± 0.2	0.73 ± 0.06	0.043 ± 0.003	0.059 ± 0.004
E176Q	1.4 ± 0.2	0.42 ± 0.06	0.14 ± 0.02	0.19 ± 0.03
R178A	2.0 ± 0.4	0.85 ± 0.12	1.8 ± 0.3	2.5 ± 0.4
Y203A	2.0 ± 0.4	0.70 ± 0.12	0.12 ± 0.01	0.16 ± 0.01
Y203F	3.6 ± 0.3	1.1 ± 0.1	0.14 ± 0.01	0.19 ± 0.01
E84Q/H137A	2.4 ± 0.5	0.73 ± 0.15	0.24 ± 0.04	0.33 ± 0.05
E84Q/E176Q	Activity not detected		Activity not detected	
N125A/I126A	0.83 ± 0.1	0.25 ± 0.03	Activity not detected	
H137A/E176Q	2.2 ± 0.3	0.67 ± 0.09	Activity not detected	
E176Q/R178A	3.0 ± 0.5	0.91 ± 0.15	1.0 ± 0.3	1.4 ± 0.3
E176Q/Y203F	1.5 ± 0.3	0.45 ± 0.09	Activity not detected	

Errors represent SD (n = 3).

Table S3. The K_m values calculated for all ssNAT variants with the Ser- substrate peptide and the Met- substrate peptide

	Ser- peptide		Met- peptide	
ssNAT variant	μM	Normalized to WT	μM	Normalized to WT
WT	54 ± 11	1.0	400 ± 40	1.0
L82A	120 ± 10	2.2 ± 0.2	350 ± 50	0.88 ± 0.13
P83A	47 ± 8	0.87 ± 0.15	300 ± 100	0.75 ± 0.25
E84A	220 ± 23	4.1 ± 0.4	480 ± 50	1.2 ± 0.2
E84Q	100 ± 11	1.9 ± 0.2	510 ± 80	1.2 ± 0.2
E84V	160 ± 11	3.0 ± 0.2	500 ± 50	1.3 ± 0.2
Y86A	70 ± 12	1.3 ± 0.2	Activity not detected	
Y86F	77 ± 21	1.4 ± 0.4	300 ± 30	0.75 ± 0.08
H137A	35 ± 5	0.65 ± 0.09	Activity not detected	
H137F	52 ± 6	0.96 ± 11	Activity not detected	
H137Q	90 ± 15	1.7 ± 0.3	Activity not detected	
Y174A	94 ± 19	1.7 ± 0.4	1,000 ± 120	2.5 ± 0.3
Y174S	100 ± 15	1.9 ± 0.3	750 ± 95	1.9 ± 0.2
E176A	62 ± 5	1.1 ± 0.1	Activity not detected	
E176H	53 ± 8	0.98 ± 0.15	300 ± 60	0.75 ± 0.15
E176Q	39 ± 4	0.72 ± 0.07	490 ± 98	1.2 ± 0.2
R178A	63 ± 10	1.2 ± 0.2	370 ± 40	0.93 ± 0.1
Y203A	54 ± 5	1.0 ± 0.1	403 ± 70	1.0 ± 0.2
Y203F	69 ± 13	1.3 ± 0.2	605 ± 100	1.5 ± 0.3
E84Q/H137A	400 ± 30	7.4 ± 0.6	413 ± 67	1.0 ± 0.2
E84Q/E176Q	Activity not detected		Activity not detected	
N125A/I126A	210 ± 40	3.9 ± 0.7	Activity not detected	
H137A/E176Q	230 ± 30	4.3 ± 0.6	Activity not detected	
E176Q/R178A	230 ± 15	4.3 ± 0.3	770 ± 150	1.9 ± 0.4
E176Q/Y203F	140 ± 25	2.6 ± 0.5	Activity not detected	

Errors represent SD (n = 3).