METHODS

Molecular docking. The 1.5 Å X-ray structure of Tm0936 was used in the docking calculations (PDB code 1P1M). The active site metal ion was assigned a charge of +1.4, the remaining charge of 0.6 was distributed among the ligating residues, His 55, His 57, His 200 and Asp 279, to keep the correct net charge and to account for charge distribution effects in metal complexes[20,33]. His 228 was protonated according to its assumed function as the base to activate the catalytic water molecule by abstracting a proton; the hydroxide ion itself was removed from the active site, because it is part of each high-energy intermediate structure that we dock[20].

The programs CHEMGRID and DISTMAP were used to compute docking grids for van der Waals potentials and excluded volume, respectively[31]. The electrostatic potential grid was calculated with DELPHI using an internal dielectric of 2 and an external dielectric of 78 (ref. 34). A manually curated set of spheres based on a set calculated by the program SPHGEN was used to orient molecules in the binding site[36].

High-energy intermediates of potential substrates were docked into Tm0936 using the docking program DOCK3.5.54 (ref. 23). Initial ligand orientations were sampled using receptor and ligand bin sizes of 0.5 Å and a ligand and receptor overlap of 0.4 Å. A distance tolerance of 1.5 Å was used for matching receptor and ligand spheres. An average of more than a million poses per molecule were calculated, and those that sterically fit the site were scored for electrostatic and van der Waals complementarity and penalized for ligand desolvation[32]. The best scoring orientation was rigid-body-minimized according to these energies. The details of the preparation of high-energy structures for docking, the molecular docking procedure, and methods for the analysis of the results have been previously described[20].

Enzymatic characterization of Tm0936. All compounds and coupling enzymes were obtained from Sigma or Aldrich, unless otherwise specified. The genomic DNA from *T. maritima* was purchased from the American Type Culture Collection (ATCC). The oligonucleotide synthesis and DNA sequencing reactions were performed by the Gene Technology Laboratory of Texas A&M University. The pET30a(+) expression vector was acquired from Novagen. The T4 DNA ligase and the restriction enzymes, *Nde*I and *Eco*RI, were purchased from New England Biolabs. The Platinum *Pfx* DNA polymerase and the Wizard Plus SV Mini-Prep DNA purification kit were obtained from Invitrogen and Promega, respectively. The glycerol stock of the plasmid encoding Tm0172 was kindly provided by the Joint Center for Structural Genomics.

Cloning of Tm0936. The gene encoding Tm0936 from *Thermotoga maritima* was amplified from the genomic DNA by standard PCR methods stipulated in the manufacturer's instructions using oligonucleotide primers with *Nde*I and *Eco*RI restriction sites at either end (Supplementary Table 3). The PCR products were purified, digested with *Nde*I and *Eco*RI, ligated to the expression vector pET30a(+) using T4 DNA ligase, and then transformed into XL1Blue cells. Individual colonies containing the plasmid were selected on LB plates containing 50 μg ml^{-1} kanamycin and then used to inoculate 5 ml cultures of LB. The entire coding regions of the plasmids containing the *Tm0936* gene were sequenced to confirm the fidelity of the PCR amplification.

Purification of Tm0936. Cells harbouring the plasmid for the expression of Tm0936 were grown overnight and a single colony was used to inoculate 50 ml of LB media containing 50 μM kanamycin, and subsequently used to inoculate 2 l of the same medium. Cell cultures were grown at 37 °C with a rotary shaker until an A_{600} of ~ 0.6 was reached. Induction was initiated by the addition of 1.0 mM isopropyl-thiogalactoside (IPTG), and further incubated overnight at 30 °C. The bacterial cells were isolated by centrifugation at 5,200 × *g* for 15 min at 4 °C. The pellet was re-suspended in 50 mM HEPES buffer, pH 7.5 (buffer A), containing 5 μg ml^{-1} RNase and 0.1 mg ml^{-1} PMSF per gram of wet cells and then disrupted by sonication. The soluble protein was separated from the cell debris by centrifugation at 14,000 × *g* for 15 min and heated at 65 °C for 15 min to precipitate the *Escherichia coli* proteins. The soluble protein was separated from the precipitated protein by centrifugation at 14,000 × *g* for 15 min, loaded onto a 6 ml Resource Q anion ion exchange column (GE Health Care) and eluted with a gradient of NaCl in 20 mM HEPES, pH 8.5 (buffer B). The fractions containing Tm0936 were pooled and re-precipitated by saturation with ammonium sulphate, centrifuged at 14,000 × *g* for 15 min at 4 °C, and resuspended in a minimum amount of buffer A. The final step in the purification was accomplished by chromatography on a High Load 26/60 Superdex 200 prep grade gel filtration column (GE Health Care) and eluted with buffer A. The purity of the protein during the isolation procedure was monitored by SDS–PAGE.

Purification of Tm0172. Cells harbouring the plasmid for the expression of Tm0172 were grown overnight and a single colony was used to inoculate 50 ml of LB media containing 100 μM ampicillin and subsequently used to inoculate 2 l of the same medium. Cell cultures were grown at 37 °C with a rotary shaker until an A_{600} of ~ 0.6 was reached. Induction was initiated by the addition of 1.0 mM arabinose, and further incubated overnight at 37 °C. The bacterial cells were isolated by centrifugation at 5,200 × *g* for 15 min at 4 °C. The pellet was re-suspended in 20 mM Tris-Cl buffer, 5 mM imidazole and 500 mM NaCl at pH 7.5 (buffer A), containing 0.1 mg ml^{-1} phenylmethylsulphonyl fluoride per gram of wet cells and then disrupted by sonication. The soluble protein was separated from the cell debris by centrifugation at 14,000 × *g* for 15 min and heated at 65 °C for 15 min to precipitate the *E. coli* proteins. The soluble protein was separated from the precipitated protein by centrifugation at 14,000 × *g* for 15 min, loaded onto a Chelating Sepharose Fast Flow column for histidine-tagged fusion protein purification and eluted with a gradient of imidazole in buffer A. Fractions containing the desired protein were pooled by catalytic activity and purity. The purity of the protein during the isolation procedure was monitored by SDS–PAGE.

Metal analysis and amino acid sequence verification. The purified Tm0936 was subjected to amino-terminal amino acid sequence analysis by the Protein Chemistry Laboratory at Texas A&M University. The first five amino acids were MIIGN, which agrees with the protein sequence reported for Tm0936. The protein concentration was determined spectrophotometrically at 280 nm using a SPECTRAmax-340 microplate reader (Molecular Devices). An extinction coefficient of 51,020 M^{-1}cm^{-1} was used for Tm0936 on the basis of the protein sequence. The metal content of the purified protein was determined by inductively coupled plasma emission-mass spectrometry (ICP-MS) and found to contain 1.2 equivalents of Zn per subunit.

Determination of SAH deaminase activity. The measurement of the deaminating properties of Tm0936 was conducted by coupling the production of ammonia to the oxidation of NADH with glutamate dehydrogenase. The decrease in the concentration of NADH was followed spectrophotometrically at 340 nm using a SPECTRAmax-340 microplate reader. The standard assay was modified from the report in ref. 36, and contained 100 mM HEPES at pH 8.0, 7.4 mM α-ketoglutarate, 0.4 mM NADH, 6 units of glutamate dehydrogenase, Tm0936 and the appropriate compound in a final volume of 250 μl at 30 °C. Following the initial, purely bioinformatic predictions of cytosine deaminase activity, the following compounds were tested for enzymatic activity at a concentration of 10 mM using this protocol: cytosine, 5-methylcytosine, 5-fluorocytosine, 6-aminouracyl, 4,6-diamino-2-hydroxypyrimidine, 2-deoxycytidine, cytosine-β-D-arabinofuranoside, cytidine, cytidine-5′-diphosphocholine, cytidine-5′-monophosphate, 2′-deoxycytidine-5′-diphosphate, cytidine-5′-diphosphate, cytidine-5′-triphosphate, cytidine-3′-phosphate. Subsequently, we cast a wider net looking for activity on *N*-formimino-L-glutamate, *N*-formimino-L-aspartate, *N*-formimino-L-glycine. It was only with the structure-based docking predictions that we turned to deamination of adenosines, first directly testing the docking predicted metabolites adenosine, adenosine-5′-monophosphate, 5′-methylthioadenosine, *S*-adenosine-5′-homocysteine. Eventually we tested also several other analogues including adenosine-5′-diphosphate, adenosine-5′-triphosphate, *S*-adenosine-5′-methionine, folate, thiamine, pterin, and guanine. Only adenosine, *S*-adenosine-5′-homocysteine, and 5′-methylthioadenosine were found to be substrates.

The products of the reaction catalysed by Tm0936 were confirmed by mass spectroscopy and by changes in the UV spectra. When *S*-adenosyl-5′-homocysteine was incubated with Tm0936, the mass spectral signal for SAH at a [M+H] of 385 *m*/*z* disappeared and was replaced by a new signal at a [M+H] of 386 *m*/*z* that corresponds to the mass expected for *S*-inosyl-5′-homocysteine (SIH). The UV spectrum for SAH has a maximal absorbance at 260 nm and after the addition of Tm0936 the absorbance maximum shifts to 250 nm. These results are consistent with the deamination of the adenine moiety of the substrate and conversion to an inosyl substituent. Similar results were observed for 5′-methylthioadenosine ([M+H] of 298 *m*/*z* and A_{max} of 260 nm) on conversion to 5′-methylthioinosine ([M+H] of 299 *m*/*z* and A_{max} of 250 nm) and adenosine ([M+H] of 268 *m*/*z* and A_{max} of 260 nm) on conversion inosine ([M+H] of 269 *m*/*z* and A_{max} of 250 nm) with an isosbestic point at 251 nm (Supplementary Information and Supplementary Fig. 3).

Determination of SAH hydrolase activity. The homocysteinase activity of Tm0172 was determined by reaction of the free thiol group of the homocysteine product with DTNB. The increase in the absorbance at 412 nm was monitored using an extinction coefficient of 13,600 M^{-1}cm^{-1} (ref. 37). The standard assay contained 100 mM HEPES at pH 8.0, 1.0 mM DTNB, 1.0 mM EDTA, 13 μM Tm0172 and the appropriate substrate in a final volume of 250 μl at 30 °C. The following compounds were tested for catalytic activity at concentrations up to 10 mM: *S*-adenosine-5′-homocysteine, *S*-inosyl-5′-homocysteine, 5′-methylthioadenosine and 5′-methylthioinosine. Activity was obtained only for SAH and SIH.

Data analysis. The kinetic parameters, k_{cat}, and k_{cat}/K_m were determined by fitting the initial velocity data to the equation (1), where v is the initial velocity, E_T is the enzyme concentration, k_{cat} is the turnover number, S is the substrate concentration, and K_m is the Michaelis constant[38]. In the cases where substrate

©2007 Nature Publishing Group

inhibition was observed an extra parameter (K_{is}) was included to calculate the apparent inhibition constant for the substrate, as observed in equation (2) (ref. 39):

$$v / E_T = k_{cat}S / (K_m + S) \quad (1)$$

$$v / E_T = k_{cat}S / [K_m + S + (S^2 / K_{is})] \quad (2)$$

Sequence alignment. A multiple sequence alignment of Tm0936 with the likely orthologues from other organisms is presented in Supplementary Figure 2 of Supplementary Information.

X-ray crystallography. Tm0936 was co-crystallized with *S*-inosyl-homocysteine (SIH) and $ZnCl_2$. The enzyme solution at 12.9 mg ml^{-1} in 20 mM HEPES, pH 8.0 was incubated for 60 min at 4 °C with 10 mM SIH and 0.5 mM $ZnCl_2$. The ternary complex was crystallized by hanging drop vapour diffusion using 1 μl of the protein–ligand solution and 1 μl of a reservoir solution containing 3.5 M Na formate, 0.5 mM $ZnCl_2$, pH 7.0. Crystals appeared in 1–2 days and exhibited diffraction consistent with the space group $P3_221$ ($a = 113.28$ Å, $c = 80.30$ Å, with 1 molecule of the ternary complex per asymmetric unit). X-ray diffraction data to 2.1 Å were collected at the NSLS X4A beamline (Brookhaven National Laboratory) on an ADSC CCD detector. Diffraction data were integrated and scaled using the programs DENZO and SCALEPACK[40]. The final 2.1 Å data set was 93.2% complete with $R_{merge} = 0.097$.

Structure determination and model refinement. The structure of the ternary Tm0936·SIH·Zn complex was solved by molecular replacement with the program PHASER[41], using apo Tm0936 (PDB code 1J6P) as the search model. The solution was subsequently refined with CNS[42]. The bound SIH and Zn were clearly visible in the electron density maps after the first cycle of rigid body refinement of the protein molecule alone. Iterative cycles of manual rebuilding with TOM[43] and refinement with CNS resulted in a model with R_{cryst} and R_{free} of 0.209 and 0.238, respectively. The final structure contains 3,210 protein atoms, 1 inhibitor molecule, 1 Zn atom, and 76 water molecules for one monomer of the complex in the asymmetric unit (Supplementary Information and Supplementary Table 4).

33. Irwin, J. J., Raushel, F. M. & Shoichet, B. K. Virtual screening against metalloenzymes for inhibitors and substrates. *Biochemistry* **44,** 12316–12328 (2005).
34. Gilson, M. K. & Honig, B. H. Calculation of electrostatic potentials in an enzyme active site. *Nature* **330,** 84–86 (1987).
35. Kuntz, I. D. *et al.* A Geometric approach to macromolecule–ligand interactions. *J. Mol. Biol.* **161,** 269–288 (1982).
36. Muszbek, L., Polgar, J. & Fesus, L. Kinetic determination of blood coagulation Factor XIII in plasma. *Clin. Chem.* **31,** 35–40 (1985).
37. Ellman, G. L. A colorimetric method for determining low concentrations of mercaptans. *Arch. Biochem. Biophys.* **74,** 443–450 (1958).
38. Cleland, W. W. Statistical analysis of enzyme kinetic data. *Methods Enzymol.* **63,** 103–138 (1979).
39. Cleland, W. W. Substrate inhibition. *Methods Enzymol.* **63,** 500–513 (1979).
40. Otwinowski, Z. & Minor, W. in *Methods in Enzymology* Vol. 276 (eds Carter, C. W. & Sweet, R. M.) 307–326 (Academic Press, New York, 1997).
41. Storoni, L. C., McCoy, A. J. & Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D* **60,** 432–438 (2004).
42. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
43. Jones, T. A. Diffraction methods for biological macromolecules. Interactive computer graphics: FRODO. *Methods Enzymol.* **115,** 157–171 (1985).

©2007 Nature Publishing Group