

Supporting Information

Enzymatic Deamination of the Epigenetic Base *N*-6-Methyladenine

Siddhesh S. Kamat, Hao Fan, J. Michael Sauder, Stephen K. Burley, Brian K. Shoichet, Andrej Sali, and Frank M. Raushel

Cloning of Bh0637 from B. halodurans (C-125). The DNA sequence for Bh0637 from *B. halodurans C-125* was cloned (gil15613200). The gene for Bh0637 was amplified utilizing the primer pair 5'-AAAGAGAATCATATGTGTGAACAAAAGTATCGCTGGACGAAAAAGC-3' and 5'-AGAAGAAAGCTTTTAACGCATAATTGAAGGAAAGAGAACTGTTTTCTTATGAAC-3'. *NdeI* and *HindIII* restriction sites were introduced into the forward and the reverse primers, respectively. The PCR product was purified with a PCR cleanup system (Qiagen), digested with *NdeI* and *HindIII*, and ligated into a pET30a(+) vector which was previously digested with *NdeI* and *HindIII*. Colony PCR was used to verify whether the colony had the gene-insert prior to sequencing. The cloned gene fragment was sequenced to verify the fidelity of the PCR amplification.

Protein Expression and Purification. The recombinant plasmid bearing the gene for Bh0637 was transformed into *E. coli* BL21 (DE3) competent cells by electroporation. A single colony was grown overnight at 37 °C in 5 mL of LB medium containing 50 µg/mL kanamycin. Five mL aliquots were used to inoculate 6 L of the same medium. The cell cultures were grown at 37 °C and induced with 0.5 mM isopropyl-β-thiogalactoside (IPTG) when the A₆₀₀ reached ~0.6 in the presence of 1.0 mM MnCl₂. The cells were centrifuged and then resuspended in 50 mM HEPES, pH 7.5, containing 0.1 mg/mL phenylmethylsulfonyl fluoride and lysed by sonication. The soluble proteins were separated from the cell debris by centrifugation at 12000 x

g for 15 minutes at 4 °C. The nucleic acids were removed by dropwise addition of 2% w/v protamine sulfate. After centrifugation, solid ammonium sulfate was added to 60% saturation to the supernatant solution. The precipitated protein was dissolved in buffer and then applied to a High Load 26/60 Superdex 200 prep grade gel filtration column (GE HealthCare). The active fractions were pooled and loaded onto a ResourceQ column (6 mL) and eluted with a gradient of NaCl in 20 mM Hepes, pH 7.5. The protein purity was confirmed by SDS-PAGE.

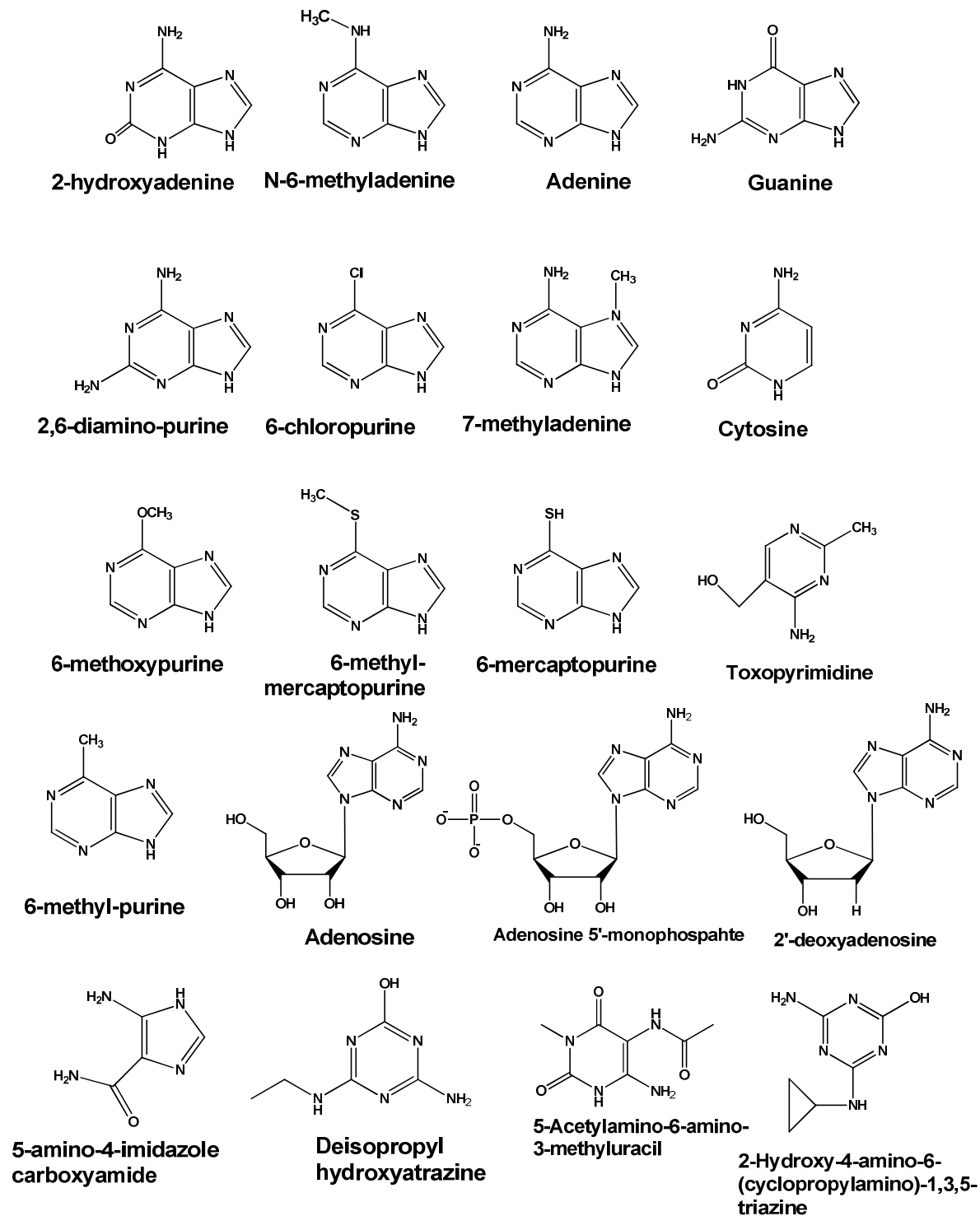
Iron-Free Expression. The iron content of our LB medium was determined to be approximately 36 μM by inductively coupled plasma mass spectrometry (ICPMS). The iron-specific chelator 2,2'-dipyridyl was used to remove this metal during protein expression. A single colony was grown overnight at 37 °C in 5 mL of LB medium containing 50 $\mu\text{g}/\text{mL}$ kanamycin and then added to 6 L of the same medium. When the A_{600} reached 0.15 – 0.20, 50 μM 2,2'-dipyridyl was added to sequester the iron, followed by the addition of 0.5 mM IPTG and 1.0 mM MnCl_2 when the A_{600} was ~ 0.6 .

Table S1: Proteins in subgroup 2 in cog 1001

Organism	Locus tag	Gi number
<i>Bacillus halodurans</i> C-125	Bh0637	15613200
<i>Bacillus clausii</i> KSM-K16	ABC1075	56962849
<i>Geobacillus kaustophilus</i> HTA426	GK0271	56418806
<i>Anoxybacillus flavithermus</i> WK1	Aflv_0242	212638091
<i>Geobacillus thermodenitrificans</i> NG80-2	GTNG_0250	138893926
<i>Bacillus cereus</i> Q1	BCQ_2842	222096502
<i>Lysinibacillus sphaericus</i> C3-41	Bsph_0221	169825824
<i>Bacillus licheniformis</i> ATCC 14580	BL01492	52079144
<i>Bacillus amyliquefaciens</i> FZB42	RBAM_006960	154685151
<i>Bacillus subtilis</i> str. 168	BSU06560	16077724
<i>Bacillus pumilus</i> SAFR-032	BPUM_0617	157691407
<i>Bacillus cereus</i> ATCC 14579	BC3012	30021127
<i>Bacillus cereus</i> B4264	BCB_4264_A3029	218233799
<i>Exiguobacterium sibiricum</i> 255-15	Exig_0459	172056498
<i>Bacillus cereus</i> G9842	BCG9842_B2218	218898088
<i>Bacillus thuringiensis</i> str. Al Hakam	BALH_2712	118478348
<i>Bacillus cereus</i> 03BB102	BCA_3097	225864990
<i>Bacillus cereus</i> E33L	BCZK2753	52142489
<i>Bacillus thuringiensis</i> serovar konkukian str. 97-27	BT9727_2766	49480044
<i>Bacillus cereus</i> AH187	BACH_A3083	217960451

<i>Bacillus anthracis str. 'Ames Ancestor'</i>	GBAA3032	47528323
<i>Bacillus anthracis str. Ames</i>	BA_3032	30262986
<i>Bacillus anthracis str. Sterne</i>	BAS2818	49185824
<i>Bacillus cereus AH820</i>	BCAH820_3026	218904142
<i>Bacillus weihenstephanensis KBAB4</i>	BcerKBAB4_2823	163940762
<i>Oceanobacillus iheyensis HTE831</i>	OB0751	23098206
<i>Picrophilus torridus DSM 9790</i>	PTO1085	48478157
<i>Brevibacillus brevis NBRC 100599</i>	BBR47_06120	226310199
<i>Rubrobacter xylanophilus DSM 9941</i>	Rxyl_1744	108804579
<i>Geobacillus thermodenitrificans NG80-2</i>	GTNG_1889	138895539
<i>Geobacillus kaustophilus HTA426</i>	GK1989	56420524
<i>Geobacillus metallireducens</i>	GS-15	78223406

Chart S1: Compounds tested as substrates for Bh0637 and Bsu06560



Scheme S1: Sequence Alignment between Bh0637 and Atu4426

Sequence alignment between Bh0637 and Atu4426 (PDB code 3NQB, A chain). Residues that are within 4 Å of the binuclear metal center are highlighted.

```

Bh0637  9 TKKQIR-----QLAVVRGEMAPT LVLKNATY LNSVRGKWL DANIWI 50
3nqbA  9 EPADLNDDTLRARAVAAARGDQRFDVLI TGGTLVDVVTGELRPADIGI 56

Bh0637  51 YQDRIVYV--GQDMPAKL DDETEVVD CGQQVIVPGYIEHFAH- PFQL 94
3nqbA  57 VGALIASVHEPAS-----RRDAAQVIDAGGAYVSPGLIDT HMHIES SM 99

Bh0637  95 YNPHSFANYAAAMGTTTTLINDNLMFFLALEK K KALS MIESLDEL P SSM 142
3nqbA  100 ITPAAYAAAVVARGVTTI VWD PHEFGNVHGV D GVRWAAKAIENLPLRA 147

Bh0637  143 YWWCYRYPQTE--MNDEEGHFLNSKIKEWLEHPLVVQGGELT SWPKVI 188
3nqbA  148 ILLAPSCVPSAPGLERGGADFDAAILADLLSWPEIGGIAEIMNMRGVI 195

Bh0637  189 TGDDGILHWMQETRRLRKP IEGH FPGASEKTLTQMSLLGVTSDHEAMT 236
3nqbA  196 ERDPRMSGIVQAGLAAEKLVCGHARGLKNADLNAFMAAGVSSDHELVS 243

Bh0637  237 GEEVIRRLDLGYMTSLRHSSIRSDLAKILREMKELGIDDFSRCLTTD 284
3nqbA  244 GEDLMAKLRAGLTIELR-GSHDHLLEPFVAAALNTLGHLPQVTTLCTD 289

Bh0637  285 GSPPSFYEQ-GIMDRLIKIALDEGIPPKDAYGMATYYVARYYGLDYEL 331
3nqbA  290 DVFPDDLQGGGLDDVVRRLVRYGLKPEWALRAATLNAAQRLGRSD-L 336

Bh0637  332 GMIA PGRI AHLNFDNVFNPVPTSVLAKGQWVVRDQGRYGS DSVFPWE 379
3nqbA  337 GLIAAGRRA DIVVFEDLNGFSARHV LASGRAVAEGGRMLVDIPTCDT- 383

Bh0637  380 DFGMKRLTIDWDL SVDELHFSMPMG-----IELV-NSVILKPYQVSV 420
3nqbA  384 TVLKGSMKL-PLRMANDFLVKS---QGAKVRLATIDRPRFTQWGETEA 427

Bh0637  421 EASRDTLAEDHDE--CFFLLLDKHG--KWKIP TMIKGF AKKV SGLAS 463
3nqbA  428 DVKDG FVVPPE--GATMISVTHRHGMAEPTTKTGFLTGWGRWNGAFAT 473

Bh0637  464 SFSTTG-DIILIGKCIQDMI VAFNALKQQGGGIVLVENGEV ISNIPLE 510
3nqbA  474 TVSHDSHNLTVFGGNAGDMALAA NAVIGTGGGM AVASEGKVTAI LPLP 521

Bh0637  511 IMGLLSSKPMEEVMEE EKKFVKALRE-RGYEHD-DPIYSLLFFSSTHL 556
3nqbA  522 LSGLVSDAPLEEVARAFEDLREAVGKVVEWQPPYL VFKACFGATLACN 569

Bh0637  557 PYIRVTQRGIYDVHKKTVLFP SIMR 581
3nqbA  570 IGPHQTD MGIADVLTGKVMESPIE 594

```