Supplemental Data

Title: Molecular factors of hypochlorite tolerance in the hypersaline archaeon Haloferax volcanii

Authors: Miguel Gomez¹⁺, Whinkie Leung¹⁺, Swathi Dantuluri¹, Alexander Pillai¹, Zyan Gani¹, Sungmin Hwang¹, Lana J. McMillan^{1,2}, Saija Kiljunen³, Harri Savilahti⁴ and Julie A. Maupin-Furlow^{1,2*}

Affiliations: ¹Department of Microbiology and Cell Science, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL 32611, USA; ²Genetics Institute, University of Florida, Gainesville, FL 32611, USA; ³Department of Bacteriology and Immunology, Immunobiology Research Program, University of Helsinki, Helsinki, Finland; ⁴Division of Genetics and Physiology, Department of Biology, University of Turku, Turku, Finland.

*Corresponding Author: Department of Microbiology and Cell Science, University of Florida, Gainesville, FL 32611, USA. Phone: (352) 392-4095; Facsimile (352) 395-5922; Email: jmaupin@ufl.edu

*Primary authors: contributing equally to this work.

Contents

Supplemental Table S1. Strains, plasmids and oligonucleotide primers used in this study.

Supplemental Table S2. Cycling conditions for the inverse-nested two-step PCR (INT-PCR) and semirandom two-step PCR (ST-PCR).

Supplemental Figure S1. Transposon insertion sites with respect to the promoter region of *hvo_2469* (SNF family transporter) and *hvo_1957* (AAA-type ATPase proteasome associated nucleotidase PanB2).

Supplemental Figure S2. Twin arginine translocation (Tat) signal peptide and lipoprotein motifs of select *H. volcanii* proteins.

Supplemental Figure S3. *H. volcanii* PstS1 (HVO_2375) is structurally related to bacterial phosphate (Pi) solute binding proteins.

Supplemental Figure S4. PhoU domain family proteins of *H. volcanii*.

Supplemental Figure S5. Relationship of H. volcanii PhoU2 (HVO_2374) to Thermotoga maritima Tm1734.

Supplemental Figure S6. Comparison of *H. volcanii* TsgA6 (HVO_A0494) to the *E. coli* maltose binding protein (MalE).

Supplemental Figure S7. H. volcanii HVO_B0012 (BetT) and its predicted function in choline transport.

Supplemental Figure S8. H. volcanii SNF family transporter HVO_2469.

Supplemental Figure S9. *H. volcanii* SpeE homologs are related to the polyamine aminopropyltransfease of *Thermus thermophilus*.

Supplemental Figure S10. Multiple amino acid sequence alignment of the active site region of haloarchaeal members of the SpeE family.

Supplemental Figure S11. *H. volcanii* HVO_0823 is a cytochrome *c*-oxidase (EC: 1.9.3.1) type helical bundle protein.

Supplemental Figure. S12. HVO_2504 is a putative oxidoreductase of the short-chain dehydrogenase/reductase (SDR, IPR002347) family.

Supplemental Figure S13. *H. volcanii* HVO_2770 is a DUF2342 protein of the zinicin-like metallopeptidase type 2 (IPR018766) and F420 biosynthesis associated (IPR022454) families.

Supplemental Figure S14. *H. volcanii* HVO_1041 is a metallo-β-lactamase superfamily protein.

A)

HVO_2468	TGA CC	CGACTCCACC							
_		▼CTCTCTGATT	CGTGGCGACA	TTCGCCACGA	TAAGGGATGA	CATTCGTTTC			
		GGTGCGCACG	AATTTTTACA	CGGTACGCAA	GTGATTCACG	AACACGACAC			
		TCGTTAGAAA	AATGAGGCTA	TTCGGGGGGC	GAATTCGGAC	GAACGAACCA			
		TATATGATAG	TTGGTTAGAG	AATTGTGTCG	AATTTTGCTC	CGATATGCGA			
		CGGCATAC VCG	C GTCA TTTAT	CACACCCACA	TCGGCAACAG	GGTGATAGAC	ATG	HVO_	2469
		CR	nAAT TTTAW	A					
D)									
D)									
HVO_1956	TGA CG	ACCGCCGACC	GCGACTCTCG	GAAGCGGTCA	CTCGCCGGAT	ACTACCGGAT			
		GAATACCGAC	ATCCCGATTT	ACGGGACGTT	CTCGGCGATG	ACGAATTCCT			
		TTCAAGTCGA	TGGGCGTCGT	ATCCCTTCTC	GCGGAGGGCA	TGTCGGGCCG			
		AAACGAGGGG	CGACCTGTAC	CCACCGAACC	TCCGCAACTT	CCTCCTTCCA			
		CAGCTTCGTC	TTCGTTGCTT	CGCGCGTTCC	TGCGAACCGC	TCGTGCCCGT			
		GAGCCGCGGC	GTCGCGCTGC	CGGCCTGCGA	TTCGCCGCCG	CGAGCGAGCG			
		CCACGGGACC	GCCGCCACCT	AAAGCCGCTT	CGAGCGCGGC	AGTCGGGTTC			
		TCTGACCC AA A	AG G▼CA CA TTT	ACGCGCATAC	CTAAATGAGG	CCAATCACTG	ATG	HVO_	1957
		CRnAA 7	г ттт	AWA					
			CR nAAT						

Figure S1. Transposon insertion sites with respect to the promoter region of hvo_2469 and hvo_1957 . Indicated are transposon insertion sites ($\mathbf{\nabla}$), start/stop codons (red/green), and conserved promoter elements including the TATA box (TTTAWA, blue) and B recognition element (BRE) (CRnAAT, purple) (where W is A or T, R is A or G, and n is any nucleotide base). Panel A, the transposon insertion sites of isolates 7 and 35A identified 250 bp and 42 bp upstream of the start codon of hvo_2469 , respectively. Panel B, transposon insertion site of isolate 36A identified 37 bp upstream of the start codon of hvo_1957 .



Figure S2. Twin arginine translocation (Tat) signal peptide and lipoprotein motifs of select *H. volcanii* proteins. Transposon insertions in the coding sequences of HVO_A0464 (TsgA6, solute binding protein), HVO_2145 (HycF, halocyanin) and HVO_2375 (PstS1, solute binding protein of Pi transport) render cells hypertolerant to hypochlorite stress (top three proteins listed). All three polypeptides are predicted to be translocated via the cell membrane by the Tat system and cleaved at the P2-P1-P1' site of Ala-Gly-Cys consensus to generate a N-terminal cysteine residue (in red) that is lipid modified based on analogy to halocyanin [56]. Predicted site of cleavage is indicated by a down arrow. Other proteins discussed in the text are included for comparison. Blue, positively charged N-domain; yellow, hydrophobic H-domain; green, C-domain of the Tat lipoprotein and Tat signal peptides. The SRRXFLK, [LVI][ASTVI][GAS] and AXA consensus sequences are underlined in the

protein sequence (where X represents any amino acid).



Figure S3. *H. volcanii* PstS1 (HVO_2375) is structurally related to bacterial phosphate (Pi) solute binding proteins. A-B) Phyre2-based structural homology model of HVO_2375 and its relationship at an RMSD of 1.057 Å to the X-ray crystal structure of the Pi solute binding protein *Mycobacterium tuberbulosis* MtPtsS1 (PDB: 1PC3). A) MtPtsS1 residues that coordinate Pi are conserved in HVO_2375 as represented by stick on ribbon diagram with the bound Pi indicated by an arrow and C) as red boxes/asterisk in a multiple amino acid sequence alignment. B) Surface representation of HVO_2375 and MtPtsS1 reveals both proteins completely engulf the bound Pi, but differ in surface charge. By contrast to MtPtsS1, HVO_2375 has a highly acidic shell which likely promotes the activity/stability of this protein in the hypersaline environments. The basic/neutral patch of HVO_2375 may promote interaction with partner proteins. Coulombic surface coloring is used to represent the electrostatic potential with units in the range of -10 (red), 0 (white), and 10 (blue) kcal/mol*e using Chimera v 1.7. The N-terminal Tat motif predicted for HVO_2375 is not included in the models or the multiple amino acid sequence alignment.



Figure S4. PhoU domain family proteins of *H. volcanii*. A) Protein domain architecture including the SpoVT-AbrB DNA binding domain, TrkA-C regulator of K⁺ conductance (RCK) domain and PhoU domain. Scale bar, 50 aa. B) Genomic neighborhood of *phoU* related genes. Down arrowheads (dark blue) represent the site of transposon insertion with the strain number indicated adjacent to the symbol. Arrows (green or white) represent ORFs deduced from the DNA sequence of the *H. volcanii* genome. HVO_ locus tag number and gene names are indicated below the ORF. Genes encoding proteins with one or more PhoU domains are indicated in green. Scale bar, 2 kb.



Figure S5. *H. volcanii* PhoU2 (HVO_2374) and its relationship to *Thermotoga maritima* Tm1734. A) Ribbon diagram of a Phyre2-based structural homology model of HVO_2374 compared to the X-ray crystal structure of Tm1734 (PDB: 1SUM) at RMSD of 0.724 Å. The tandem PhoU domains are highlighted in light and dark blue. Fe and Ni are represented by orange and green balls. B) Protein domain architecture with the PhoU domains colored in light and dark blue as in panel A. The N-terminal SpoVT-AbrB superfamily (IPR037914) DNA binding domain is indicated in green. Scale bar, 50 aa. C) Multiple amino acid sequence alignment of the PhoU domains colored as in panels A-B. The red boxes highlight the two multinuclear iron clusters coordinated by a conserved E(D)XXXD motif pair in the Tm1734 structure. Only the first of the two motifs is conserved in HVO_2374.



Figure S6. Comparison of *H. volcanii* TsgA6 (HVO_A0494) to the *E. coli* maltose binding protein (MalE). The structural homology model of HVO_A0494 (blue ribbon) is based on Phyre2-comparison to the PDB database and is related to the X-ray crystal structure of the *E. coli* MalE (PDB: 3OSQ, tan ribbon) at an RMSD of 0.371 Å. Residues in the active site region of MalE are indicated by tan stick diagram and in the table inset which highlights the single residue of HVO_A0494 that is conserved in this region (colored in green). Maltose is in orange stick diagram.



Figure S7. *H. volcanii* HVO_B0012 (BetT) and its predicted function in choline transport. The betaine/choline metabolic pathway of *H. volcanii* is based on comparison to bacteria and plants (KEGG pathway) [73]. Genes, represented as arrows, are colored in agreement with the encoded enzymes represented in the pathway. Abbreviations: CMO, choline monooxygenase [EC 1.14.15.7]; BetB, betaine aldehyde dehydrogenase [EC 1.2.1.8]; DMDH, dimethylglycine dehydrogenase [EC 1.5.8.4]; SoxA1, sarcosine oxidase [EC 1.5.3.1]; GlyA1/2, glycine hydroxymethyltransferase [EC 2.1.2.1]; IlvA2, L-serine/L-threonine ammonia-lyase [EC 4.3.1.17, 4.3.1.19]; ox, oxidized; red, reduced; L-Met, L-methionine; L-hCys, L-homocysteine; THF, tetrahydrofolate; CH₂-THF, N5,N10-methylenetetrahydrofolate; THF, tetrahydrofolate; FNR, ferredoxin-NADP+ reductase [EC 1.18.1.2] (F420H₂: NAD(P)H oxidoreductase HVO_0433 and ferredoxin reductase-like HVO_0889); HVO_B0006, disordered regions include 105-132 aa and 180-215 aa (MobiDB_lite); HVO_B0003, homolog of HbpS regulator up-regulated in response to haemin- and peroxide-based oxidative stress [74]; HVO_B0008, structural homolog of the C-terminus of minichromosome2 maintenance (MCM) protein (PDB: 2MA3); Fd, ferredoxin (HVO_0211 FerB1, HVO_2729 FerB2, HVO_1246 FerA1, HVO_1482 FerA2, HVO_1721 FerA3, HVO_1831 FerA4, and HVO_2995 FerA5); ETF, electron transferring flavoprotein (HVO_0304 EtfA1, HVO_0305 EtfB1, HVO_2731 EtfA2, HVO_2730 EtfB2, HVO_2732 FixC).



Figure S8. *H. volcanii* SNF family transporter homolog HVO_2469. Panel A) 2D transmembrane diagram of HVO_2469 generated using Protter [75]. Cell membrane is represented in orange. Panel B) 3D homology model of HVO_2469 (blue) compared to the X-ray crystal structure of the bacterial Na⁺-dependent SNF family transporter MhsT (PDB: 4US3, tan) at a RMSD of 0.262 Å. MhsT residues that plug the cytoplasmic side and coordinate the two Na⁺ ions (purple balls) and L-tryptophan (orange stick) are listed in the inset, with the conserved residues of HVO_2469 indicted in green.



Figure S9. *H. volcanii* SpeE homologs related to the polyamine aminopropyltransferase of *Thermus thermophilus*. Panel A) 3D homology models of the soluble domain of the *H. volcanii* SpeE homologs HVO_0255 (blue) and HVO_B0357 (pink) compared to the X-ray crystal structure of *T. thermophilus* triamine/agmatine aminopropyltransferase (PDB: 1UIR, tan) at a RMSD of 0.905 and 0.804 Å, respectively. Panel B) Zoom-in of the conserved motifs described in panel D. Panel C) C-terminal extension of HVO_0255 and 1UIR, which is not apparent in HVO_B0357. Panel D) Zoom-in of the conserved GGG(D/E)G and (E/D)(I/V)D motifs important for binding polyamine and the amino donor S-adenosylmethioninamine (red boxes) and the proton acceptor of the reaction (green box). Residues are numbered according to 1UIR.

	MONE89	DSVNRVIV	IGGAGI	SIPKRFLTE	SNVT	VDVV	ELD	PEVVNAAKTH	NV-TNS	PR	LNIYTQD	ERRY	RNINHI	YDAIVI	D-AYQKNNVPFH	TTID	MQLVSER	DERCVFIQ
	W0K500	DDVDRVIF	GGGG	SGPKRFAKE	-DVT	VDVV	EID	PGVVHAAREY	EV-SEG	PO	FNIHVED	errf	EEUNHT	YDLVVI	-AYORDAVPFH	TTVP	MELAASK	DDDCALVA
	ANA1H3MD		CCCC	TCPKREVEE	ייעם-ע	WDWW	ETD	PEVIDVAKBV	DV-FFS	-EO	T.R THND	ARRY	BNUNET	VDT.TVT	-AYKKDKVPFO	ттко	MELSHOR	TSEDCMVEA
	ACATHOME	DDIDRVF	GGGGI	I GFICKE VEE						<u>50</u>			KIN INE I		AIRIORVEEQ	11100		JE CHVIA
	AUAIH61C	DDIDRVIF.	GGGGI	GPKREVEE	Y-DAT	VDVV	EID	PEVIDVAKQY	DV-EES	EQ	LRIHNDD	erry	RNONET	NDLIVI	–AYKKDKVPFQ	TTKD	MELSNDR	SDDEMVFA
	A0A0W1R8	DDVDRVLFI	GGGGI	FIGPKRFAED	Y-NVT	VDVA	EID	PEVIDVSKEY	RV-EES	EN	LNIYNTD	CRQF	QEINRI	ADTIAT	-AYQKDKVPFE	TTED	MQLADSR	USEDGILFA
	MODF09	DDVDRVLFV	/GGGGI	FIGPKRFVRD	YPNVT	VDVV	EID	PEVVSTAKRY	SV-EES	ER	LNVYTRG	ERQY	QEINQI	YDLIVI	-AYRKDKVPFE	TTVD	MELANDR	LDEDCVLFA
	A0A1I2ST	DDVDRVIF	GCCGI	GPKRFAEE	Y-DVE	VDVA	EID	PEVVDVAKRY	EV-EES	EN	LTIHTTG	ROF	RENDET	YDLIVI	-AYKKDKVPFO	TTVD	MELASDR	TEDCILFA
	3031U9NU	DDTDRUTES	CCCCT	TOPNET			TT TO	DEUTIDUARE V	WV-FFC		TNUTHTM	ADOV			AVKKDKUDEO	mms7.5	MOLASDR	
	ACATHONI	DDIDRVF		IGERVELDI						EK					AIRIORVEEQ	110	MQLASDA	AD C VIIIA
	AUAIH/KW	DDIDRVIE	GGGGE	GPRVFLEK	YPNVT	VDVV	EID	PEVVSVAKEY	DV-EES	DR	LNIYTMG	erqy	QEANQI	ADTIAN	-AYRKDRVPFE	TTVD	MRLTSER	ODEDICILFA
	MOHW98	DDIDRVLFV	/GGGGI	TGPRVFLEK	YPNMT	VDVV	EID	PEVVSVAKEY	DV-EES	DR	LNVYTMG	GRQY	I QE I NQI	YDLIVI	D-AYRKDRVPFE	TTVD	EMRLTSER	IDEDGILFA
	A0A0W1ST	DDIDRVLFV	GGGGI	GPRVFNEM	YPNVT	VDVV	EID	PEVVSVSKQY	GV-EES	DR	LNIYTMD	ERQY	REINQI	YDLIVI	D-AYRKDKVPFE	TTVP	MRLTSER	LDEDGILFA
	A0A0D6JT	DDIDRVIFY	Reeg	TGPRVFLEK	YPNVT	VDVV	ELD	PEVVSVAEEY	GV-EES	PR	LNVHTMD	ROY	REUNRI	YDLIVI	-AYRKDKVPFO	1TTRD	MOLTSDR	IDEDEVLFA
	3030W10H	DDTDRVTF		TOPRVETEO			TT	DEVALE	CV-FFS	.DD	TNUTHTM	POV	DEUNOT		AVEKDEUPEO			DEDETLEA
	ACAUMINI	DDIDRVF		I GFRVF LEQ				EVVSVALE I	GV EES				RE NO I		AIRIDRVFFQ	1120	монтори	DEDCIDIA
	MOFD40	DDIDRVIE	GGGGE	GPRVFLEK	YPNVT	VDVV	ELD	PEVVDVAEEY	RV-EES	PR	LNVHTMD	erqy	REANRI	ADTIAN	-AYQKDKVPFQ	TTQD	MQLTSDR	ODDEVLFA
	M0FVQ5	DDIDRVLFV	/ <mark>GGG</mark> GI	FIGPRVFLEK	YPNVT	VDVV	ELD	PEVVDVAEEY	RV-EES	PR	LNVHTMD	ERQY	REINRI	YDLIVI	D-AYQKDKVPFQ	TTQD	MQLTSDR	UDDDCVLFA
	MOFZ29	DDIDRVIF	GGGGI	TGPRVFLEK	YPNVT	VDVV	ELD	PEVVDVAEEY	RV-EES	PR	LNVHTMD	ROY	REINRI	YDLIVI	-AYQKDKVPFQ	TTOP	MOLTSDR	LDDDGVLFA
	D4GZK0	DDTDRVTFY	RCCC	COPRVELEK	YPNVT	WDVW	ELD	PEVVDTAEEY	BV-EES	PR	T.NVHTMD	EROY	REUNRT	YDT.TVT	-AYOKDKVPFO	TTOP		TDDDCVLFA
÷=	T OTTEGS	DDTDRUTES	CCC	TOPNETER			TT D	DEVAD TARE V	DV-FFC			POV	DEUNDT	VDT.TVT	AVOKDKUDEO	TTTO 1		
Ħ	190105	DDIDRVIF								PR DD			REINKI		AIQIORVFFQ		MQLISDA	DDDCVIIFA
<u>o</u>	MUIIMU	DDIDRO	GGGGE	GPRVFLER	MPNVT	VDVV	ELD.	PEVVDVAEEY	RV-EES	PR	LNVHTMD	ergr	REANET	MULIVI	-AIQKDKVPFQ	TTQD	MQLTSDR	DDDEVLFA
3	l5nvt5	DDIDRVIF	GGGGI	TGPRVFLEK	YPNVT	VDVV	ELD	PEVVDAAEEY	RV-EES	PR	LKVHTMD	GRQF	REINRI	YDLIVI	-AYQKDKVPFQ	TTQP	MQLTSDR	DDDGVLFA
-	M0GF43	DDIDRVLFV	/GGGGI	TGPRVFLEK	YPNVT	VDVV	ELD	PEVVDAAEEY	RV-EES	PR	LNVHTMD	CRQF	REUNRI	YDLIVI	-AYQKDKVPFQ	TTQD	MOLTSDR	LDDDCVLFA
\sim	M0GR19	DDIDRVIF	GGGGG	TGPRVFLEK	PNVT	VDVV	ELD	PEVVSVAEEY	GV-EES	PR	LNVHTMD	ROY	REUNRI	YDLIVI	-AYRKDKVPFO	TTOP	MOLTADR	DEDCVLFA
<	AOAOK1TO	DDTDBWEFY	GCCGT	TGPRVFLEO	YPNVT	VDVV	ELD	PEVVSVAEEY	GV-EES	PR	T.NVHTMD	eroy	REUNET	YDT.TVT	-AYRKDKVPFO	TTRID	MOLTADR	DEDEVLEA
ц	MOTITINA	DDTDDUUT													AVDKDKUDEO			
\sim	MUHHA4	DDIDRVIE	GGGGI	TGPRVFLEQ	PNVT	VD VV	2.1.1	PEVVSVALE1	GV-LES	PR	LNVHTMD	GRQI	REDINKI		-AIRKDKVPPQ	TTRJ	MQLTADR	ODEDEVLFA
G	M01218	DDIDRVLF	GeeeGI	TGPRVFLEQ	YPNVT	VDVV	ELD	PEVVSVAEEY	GV-EES	PR	LNVHTMD	eRQY	REINRI	MDLIVI	-AYRKDKVPFQ	TTQD	MQLTSDR	UDEDEVLFA
\sim	M0JD57	DDIDRVLFV	/GGGGI	FGPRVFLEK	YPNVT	VDVV	ELD	PEVVSVAEEY	GV-EES	PR	LNVHTMD	ERQY	REINRI	YDLIVI	D-AYRKDKVPFQ	TTQD	MQLTSDR	LDEDGVLFA
4	I3R191	DDIDRVIF	GGGGI	TGPRVFLET	YPNVT	VDVV	ELD	PEVVAVSKEY	GV-EES	DR	LNIYTMD	eroy	REINET	YDLII	-AYRKDKVPFE	TTVD	MOLTSDR	DDNGILFA
G	MOTOT1		Recent	TCPRVFT.FT	VDNIVT	WDWW	ETD	DEMNAWSKEY	OV-ETS	RU	LNTYTM	ROY	BEUNTT		-AVEKDKVDEE	77772		
\sim	30317000	ERITED WERE									T C A UMP	Ĩ.			AVOKDOUDTU			
Ċ.	AUAIIUPB	LEVEDUTE	GGGG	GPRDFERK		VD VV	<u></u>	PEVIQAAREH	KL-LLS		LSAHIE	CRIF	QD DKI	V 1 V	-AIQKDQVPIH	1003		ALDEVILA
(J	LOJI36	DEVDDVI F.	GGGG	GPKDFERR	⊻-DVD	VDVV	ELD	PEVTQAAKEY	RL-EEG	EN	MTTHTED	ERIF	QENDET	ADTIAN	-AYQKDQVPIH	TQLS	BMELVEDR	SDDEVFLA
-	MOBNF1	DEVDDVLF.	GGG G 3	GPKDFERR	Y-DVD	VDVV	ELD	PEVTQAAKEY	RL-EEG	EN	MTAHTED	GRIF	IQETDET	YDLIVI	D-AYQKDQVPIH	TQLD	MELVEDR	ISDDCVFLA
	A0A1H8ZE	DEVENVLFI	GGGG	GPKDFERK	Y-DVN	VDVA	ELD	PDVTRAAKDH	RL-EES	EN	LTVHTAD	eKQF	RDUETR	YDVIVI	-AYQKDQVPIH	TQLD	MELVEQR	TDDCVFLA
	L9Y3S8	DEVEDVIF	GGGGS	GPKDFERK	Y-DVD	VDVA	ELD	PEVSOAAKDY	DL-EEG	EN	MTVHTED	ROF	RDUNER	YDVIVI	-AYOKDOVPIH	TOLG	MELAEDR	TEDEVELA
	TOYAWO	DEVEDVIE	CCCC	TCDODEODH			TT T	DEVTADAEAV	CL-FHC		MTCHAED		ONADEE	VDVTVT	AVERDOUDER	TTTT	MOLAADB	SDDCVTTA
	197440	DEVEDVIE								511	MISHAL		ONADEE				MQUAADA	SDDC VIIIA
	AUAIN/CI	DDVDRVI FI	GGGG	VGPQDFEDH	Y-DAT	VDVV	EID	PGVTTAAEDY	GL-EES	ET	MTAHTND	erqf	QNADET	MDVIII	-AYKQDQVPFH	TTVD	MDLVSDR	OTDDEVFYA
	A0A1H6G2	DDVDDVIF	I GGGG3	TGPQDFEEH	Y-DAD	VDVV	EID	PEVTDAAEDY	GL-EHG	EN	MTSHAED	GRQF	I QD I DQ I	YDVIVI	-AYKQDQVPFH	TTED	MELVSDR	TADDEVLHA
	M0A504	DDVDRVLFV	GGGG3	GPQDFAEQ	Y-DAT	VDVV	EID	PYVTDAAEEY	TLEDHD	KRDD	LNVYATD	CRQF	QNTDEF	YDVIVI	D-AYKKDQVPFH	TTVP	MELASDR	LADDCILHA
	D3T068	DDVERVEF	GGGGS	GPTDFAEO	-DAT	VDVV	EID	PAVTDAAEEY	TLEDHD	ERDD	LNVYATD	ROF	OSIDET	YDVIVI	-AYKKDOVPFH	TTVO	MELASDR	TADDCILHA
	MOAGVE		Reect		ייעם-ע	WDWW	ETD	PSVTDAAFFV	TTEDHE	חחפח	T.NVYATO	- ROF	ONUDET	VDVTVT	-AYKKDOVPEH	TTT72	MELASDR	
	2021 74757	DDIDIU								PDDD								
	AUAIJ4LW	DDVDRVIE	GGGGI	TGPTDFALQ	DAT		5 I P	PSVIDAALEI	TLEDHD	ERDD	LNVIATO	GROF	QNIDE1	T V I V	-AIKKDQVPFH	TTVJ	MELASDR	ADDEILHA
	AOA1H1FR	DDVDRVLFV	Geee G 7	TGPQDFAEQ	Y-DAT	VDVV	EID	PEVTDAAETY	GL	DRGA	VNIHTED	eRQF	QG DE1	NDAIAI	-AYKKDQVPFH	TTVD	MELLSDR	UTEDEVVHA
	MOLSJ6	DDIDRVLFV	GGGG7	GPQDFEQR	Y-DAH	VDVV	EID	PEVTDAAEEY	GL-DRS	DDVD	LEAHVAD	GRQF	LESTDET	YDVIVI	D-AYKKDQVPFH	TTVD	EMELVSDR	LAENCVLHA
	A0A111EN	DDVDRVLFI	GGGGY	GPODFEER	Y-DVT	VDVV	EID	PDVTDAASEY	GLEHETN	DD	LEVHTGD	ROY	ORTDET	YDVIVI	-AYKQEQVPFH	TTVD	MELVSDR	LADDCIFHA
	M01.V31	DDVDKWEF	GGGGY	TGPODFERR	Y-DVT	VDVV	ЕΤD	PDVTDAASEY	GLEHETN		LEVHTGD	EROY	OBUDET	YDVTVT	-AYKOEOVPEH	TTVD	MELVSDR	TADDCTLHA
	AOA1C7HM	I CDDDWI I	CCCDT					DEFORVERDUDE	ODVUDDAVDV	CN					I DCARCODI I H	vema	THENDOUT	
	AUAIG/HW	LSDDRVLL	GGGDV	TA-VNILKD	H-GAI		DP	REFORITORPE	QKIHDDAIKI	GN	LIVIRO		RE DRF		LPGARSDDLLH	1315	IKQVKQH	TDUCLAVI
	AUAIH8DV	LSDDRVLL	GGGDV	VLA-ANYLRD	H-GAT	VDMV	DPD	REFORVARENPE	RRIHDDAIRL	EN	LIVHROD	CIAI	READER	NDFTPI	LPGARSDDLLH	YSTS	TROVROH	TUDELAVT
	L9ZG76	SSL-NVL	GGGD3	IA-VDHLRQ	Y-NAS	VDQV	DID	GEFLELAREREF	RQHNHDAYEY	DR	LNTTTED	AITY	QE DET	YDLVLI	IPGARSDETLS	YSTID	TLLRQH	ITTDDGLVVS
	M0B814	SSVD-VLL	/GGGD3	IA-VDHLRQ	Y-NAS	VDQV	DID	GEFLEMAREREF	RQYNDEAYEY	DR	LNTTTED	AFTY	EAUDET	YDLILI	IPGVRSDETLS	YSTD	YSLLRDH	SDECLVVS
	U1PHF4	TDTD-VIVI	GeeD	A ANNLRE	H-NVS	VDLV	DVD	SAFMOOAKTDDL	LERYHEDAYEY	EH	LTVYOOD	AFEF	LETTTDI	YDLILI	LPGATDDDLLH	YSTD	YSLLTER	SPDCVVGT
			RCCDI		u_9779		ртр	CEEMOKA DTODY	POVUNDAVEV	9 M		TVDV	LOOSOTTE			Venin	VCTINT	TARDOAVAT
	MOROAA			ADRIAN				GEFIDIARIDET								VOID	TSHINTI	THE DEAVAL
<u> </u>	MOK044	TDT-RVIL.	GGGDV	AA-MDRLQR	n-Gvs	VDHV		SEPMORARTDPP.	LSQIHDGAIRI	EN	KTTRVQD	AVQI	LOSERF	TT VILL	IPGATDDDLLP.	ISVJ	ISQLKQH	TADEVIAT
÷	A0A0W1RJ	PDT-KVIV	GGGDV	IA-IDHLRE	Y-GVT	VDHV	DLD	GEFMRDAKTDPF	GRWHDDAYEY	DR	LNTTVAD	GYRY	QEINET	YDLVLI	VPGATDDDLLT	YSTID	YRSVRTH	ILSEDGVAVT
0	M0FPY4	PDT-KVLVV	/GGGDV	IA-VDHLRE	Y-GVT	VDHV	DLD	GEFMRDAKTDPF	ARWHDDAYEY	DR	LNTTVSD	GYRY	QKINET	YDLVLI	DIPGATDDDLLT	YSTD	FYRSVRTH	ILSEDGVAVT
F	M0FT97	PDT-KVIV		A-VDHLRE	Y-GVT	VDHV	DLD	GEFMRDAKTDPF	ARWHDDAYEY	DR	LNTTVS	eyry	OKINET	YDLVLI	IPGATDDDLLT	YSTD	YRSVRTH	ILSEDGVAVT
<u> </u>	MOGH73	PDT-KWIV		TA-VDHLRE	Y-GVT	VDHV	DT.D	GEFMEDAKTOPE	ARWHDDAYEY	DR	LNTTVSD	eyry	OKUNET	YDT.VT.T	TPGATDDDT.T.T	YSTD	VRSVRTH	TSEDEVAVT
	D4C004	DDW-WHAT	CCCDT				DTD	CEEMDD & KUDDE	ADVUDDAVEN			VDV	OFUNE			vema	VD CT/DUL	TO SED GVA V
2					GVI			SEPARDARIDEE							TPORTDODDU	1.51.0	IKSVKII	
~	TAATTAA	PDT-KVIIV	GGGDV	ALA-VDHLRE	-GVT	VDHV	ידי	55 FMRDAKTDPF	AKWHDDAYEY	DR	LINTTVAD	SIKI	QKIINET	LVL	TEGATODOLLT	18113	IKSVKTH	SEDEVAVT
<	L5NV96	PDT-KVIV	CCCDV	IA-VDHLRE	Y-GVT	VDHV	DLD	GEFMRDAKTDPF	ARWHDDAYEY	DR	LNTTVAD	eYRY	QKINET	ADTAT	I PGATDDDLLT	YSTD	YRSVRTH	DISEDEVAVT
<u> </u>		PDT-KVIV	GGGDV	IA-VDHLRE	Y-GVT	VDHV	DLD	GEFMRDAKTDPF	ARWHDDAYEY	DR	LNTTVAD	GYRY	QKINET	YDLVLI	IPGATDDDLLT	YSTD	YRSVRTH	ILSEDGVAVT
	MOH3A3				-				ARWHDDAYEY	DR	LNTTVAD	EYRY	OKINET	YDT.VT.T	IPGATDDDLLT	YSTO	NTD CT/D III	U SEDEVAX7T
()	MOH3A3 MOIIJ3	PDT-KVIV	GGGDV	IA-VDHLRE	Y-GVT	VDINV	DT'	GEFFIRDARTOPE							-		TROVELL	OLOC VAVI
\sim	MOHJAJ MOIIJJ TJR9D2	PDT-KVLVV		IA-VDHLRE	Y−GVT			GEFMRDARTDPF GEFMOOTKTDPF	RKWHDDAVEV	DR	TNTTTA	eypv	OESNET	YDLVL	TPGATTOTTT	YSA	YSSTRNU	ILSDE CVAVT
	MOH3A3 MOIIJ3 I3R9D2	PDT-KVLVV PET-KVLVV	/GGGDV /GGGDV	VIA-VDHLRE	Y-GVT Y-NVT			GEFMQQTKTDPF	RKWHDDAYEY	DR	LNTTTAD	GYRY	QESNET	YDLVLI	DIPGATDDDLLT	YSAD	EYSSIRNH	ILSDEGVVVT
ũ	MOHJAJ MOIIJJ IJR9D2 MOH8W1	PDT-KVLVV PET-KVLVV PET-KVLVV	/GGGD¥ /GGGD¥ /GGGD¥	VIA-VDHLRE VIA-IDHLRK VIA-IDHLRA	Y-GVT Y-NVT Y-NVT	VDHV VDHV VDHV	DLD	GEFMRDARTDPF GEFMQQTKTDPF AEFMQEAKTDPY	RKWHDDAYEY	DR	LNTTTAD LNTTTAD	GYRY GYHY	QESNET	YDLVLI YDLVLI	DIPGATDDDLLT DVPGATDDDLLT	YSAD YSAD	EYSSIRNH EYRSVRNH	ILSDEGVVVT ILSDEGVVVT
99	MOH3A3 MOIIJ3 I3R9D2 MOH8W1 AOA1H7T9	PDT-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV	/ GGGDV / GGGDV / GGGDV / GGGDV	VIA-VDHLRE VIA-IDHLRK VIA-IDHLRA VIA-IDHLRA	Y-GVT Y-NVT Y-NVT Y-NVT	VDHV VDHV VDHV VDHV	DLD	GEFMRDARTDPF GEFMQQTKTDPF AEFMQEAKTDPY DEFMQEAKTDPY	FRKWHDDAYEY FRQWHDDAYKY FRQWHDDAYKY	DR DR DR	LNTTTAD LNTTTAD LNTTTAD	GYRY GYHY GYHY	lõesnet Lõetnet Lõetnet	YDLVLI YDLVLI YDLVLI	DIPGATDDDLLT DVPGATDDDLLT DVPGATDDDLLT	YSAE YSAE YSAE	EYSSIRNH EYRSVRNH EYRSVRNH	ILSDEGVVVT ILSDDGVVVT ILSDDGVVVT
99	MOH3A3 MOIIJ3 I3R9D2 MOH8W1 AOA1H7T9 MOH3P0	PDT-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV	/ GGGD¥ / GGGD¥ / GGGD¥ / GGGD¥ / GGGD¥	VIA-VDHLRE VIA-IDHLRK VIA-IDHLRA VIA-IDHLRA VIA-IDHLRA	Y-GVT Y-NVT Y-NVT Y-NVT Y-NVT	VDHV VDHV VDHV VDHV IDHV	DLD	GEFMRDAKTDPF GEFMQQTKTDPF AEFMQEAKTDPY DEFMQEAKTDPY DEFMQEAKTDPY	FRKWHDDAYEY FRQWHDDAYKY FRQWHDDAYKY FRQWHDDAYTY	DR DR DR DR	LNTTTAD LNTTTAD LNTTTAD LNTTTAD	GYRY GYHY GYHY GYHY	l qe snet L qe t net L qe t net L qe t net	YDLVLI YDLVLI YDLVLI	DIPGATDDDLLT DVPGATDDDLLT DVPGATDDDLLT DVPGATDDDLLT	YSAE YSAE YSAE YSAE	YSSIRNH YRSVRNH YRSVRNH YRSVRNH	LSDEGVVVT LSDDGVVVT LSDDGVVVT
99	M0H3A3 M0IIJ3 I3R9D2 M0H8W1 A0A1H7T9 M0H3P0 E7QVN9	PDT-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV	/ GGGDW / GGGDW / GGGDW / GGGDW / GGGDW	VIA-VDHLRE VIA-IDHLRK VIA-IDHLRA VIA-IDHLRA VIA-IDHLRA VA-IDHLRK	Y-GVT Y-NVT Y-NVT Y-NVT Y-NVT Y-DVT	VDHV VDHV VDHV VDHV IDHV VDQV		GEFMRDARTDFF GEFMQQTKTDFF AEFMQEAKTDFY DEFMQEAKTDFY DEFMQEAKTDFY AEFMNRTKNDFF	FRKWHDDAYEY FRQWHDDAYKY FRQWHDDAYKY FRQWHDDAYTY FRQWHDDAYTY FSRWHHDSYEY	DR DR DR DR TR	LNTTTAD LNTTTAD LNTTTAD LNTTTAD LNTTIAD	GYRY GYHY GYHY GYHY GYQY	LQESNET LQETNET LQETNET LQETNET LRKTDEC	ADTALI ADTALI ADTALI ADTALI	DIPGATDDDLLT DVPGATDDDLLT DVPGATDDDLLT DVPGATDDDLLT DIPGATNDDLLK	YSAE YSAE YSAE YSAE YSAE	FYSSIRNH FYRSVRNH FYRSVRNH FYRSVRTH FYGSVRRH	USDEGVVVT USDDGVVVT USDDGVVVT USDDGVVVT USDDGVVVT
99	M0H3A3 M0IIJ3 I3R9D2 M0H8W1 A0A1H7T9 M0H3P0 E7QVN9 A0A110IN	PDT-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV	/GGGDV /GGGDV /GGGDV /GGGDV /GGGDV /GGGDV	VIA-VDHLRE VIA-IDHLRK VIA-IDHLRA VIA-IDHLRA VIA-IDHLRA VVA-IDHLRK VIA-IDHLRO	Y-GVT Y-NVT Y-NVT Y-NVT Y-NVT Y-DVT Y-NVT	VDHV VDHV VDHV VDHV VDHV VDQV VDHV		SEFMRDARTDEF GEFMQQTKTDPF AEFMQEAKTDPY DEFMQEAKTDPY AEFMNRTKNDPF AQFMNMTKHDPF	FRKWHDDAYEY FRQWHDDAYKY FRQWHDDAYKY FRQWHDDAYTY FRQWHDDAYTY FRWHHDSYEY	DR DR DR DR TR TR	LNTTTAD LNTTTAD LNTTTAD LNTTTAD LNTTIAD LQTTATD	GYRY GYHY GYHY GYHY GYQY GYRY	LQESNET LQETNET LQETNET LQETNET LQETNET LQETNET LQOTNET	YDLVII YDLVII YDLVII YDLVII YDLVII	DIPGATDDDLLT DVPGATDDDLLT DVPGATDDDLLT DVPGATDDDLLT DIPGATNDDLLK DVPGATDDDLLK	YSAD YSAD YSAD YSAD YSKD	FYSSIRNH FYRSVRNH FYRSVRNH FYRSVRTH FYGSVRRH FYGSVRRH	LSDEGVVVT LSDDGVVVT LSDDGVVVT LSDDGVVVT LSDDGVVVT LTDDGMMVT
99	M0H3A3 M0IIJ3 I3R9D2 M0H8W1 A0A1H7T9 M0H3P0 E7QVN9 A0A1I0IN A0A1H1C8	PDT-KVEVV PET-KVEVV PET-KVEVV PET-KVEVV PET-KVEVV PET-KVEVV PET-KVEV	/ GGG Di / GGG Di / GGG Di / GGG Di / GGG Di / GGG Di / GGG Di	VIA-VDHLRE VIA-IDHLRK VIA-IDHLRA VIA-IDHLRA VIA-IDHLRA VVA-IDHLRK VIA-IDHLRQ	Y-GVT Y-NVT Y-NVT Y-NVT Y-DVT Y-DVT Y-NVT	VDHV VDHV VDHV VDHV IDHV VDQV VDQV		SEFMRDAATDF GEFMQQTKTDFF AEFMQEAKTDFY DEFMQEAKTDFY DEFMQEAKTDFF AQFMNMTKHDFF AQFMNMTKHDFF	RKWHDDAYEY RQWHDDAYKY RQWHDDAYKY RQWHDDAYKY SRWHDDAYEY RRWHDDAYEY RRWHDDAYEY	DR DR DR DR TR TR	LNTTTAD LNTTTAD LNTTTAD LNTTTAD LNTTIAD LQTTATD	GYRY GYHY GYHY GYHY GYQY GYRY GYRY	LQESNET LQETNET LQETNET LQETNET LQETNET LQETNET LQQTNET	ADIAL ADIAL ADIAL ADIAL ADIAL ADIAL ADIAL ADIAL ADIAL	DIPGATDDDLLT DVPGATDDDLLT DVPGATDDDLLT DVPGATDDDLLT DIPGATDDDLLK DVPGATDDDLLT DVPGATDDDLLT	YSAF YSAF YSAF YSAF YSKF YSKF	YSSIRNH YRSVRNH YRSVRNH YRSVRTH YGSVRRH YRSVRQH	USDECVVVT USDECVVVT USDECVVT USDECVVT UTSDECVVT
99	MOH3A3 MOIIJ3 I3R9D2 MOH8W1 AOA1H7T9 MOH3PO E7QVN9 AOA110IN AOA1H1G8 AOA11427	PDT-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV PET-KVLVV	/GGGDV /GGGDV /GGGDV /GGGDV /GGGDV /GGGDV /GGGDV	VIA-VDHLRE VIA-IDHLRK VIA-IDHLRA VIA-IDHLRA VIA-IDHLRA VIA-IDHLRA VIA-IDHLRQ VA-VDHLRQ	Y-GVT Y-NVT Y-NVT Y-NVT Y-DVT Y-DVT Y-DVT Y-DVT			SEFMRDAATDPF GEFMQQTKTDPF DEFMQEAKTDPY DEFMQEAKTDPY AEFMNRTKNDPF AQFMNMTKHDPF ARFMNRTKNDSF	FRKWHDDAYEY FRQWHDDAYKY FRQWHDDAYKY FRQWHDDAYTY FRWHDDAYEY FRHWHDDAYEY FRRWHDDAYEY	DR DR DR TDR TDR TDR	LNTTTAD LNTTTAD LNTTTAD LNTTTAD LNTTIAD LQTTATD LNVTVGD	GYRY GYHY GYHY GYHY GYQY GYRY GYRY	LQESNEI LQETNEI LQETNEI LQETNEI LRKTDEQ LQQTNEI LQQTNEI	ADFALI ADFALI ADFALI ADFALI ADFALI ADFALI ADFALI ADFALI ADFALI ADFALI	DIPGATDDDLLT VPGATDDDLLT VPGATDDDLLT DVPGATDDDLLT DIPGATNDDLLK VPGATDDDLLT VPGATDDDLLT	YSAE YSAE YSAE YSAE YSKE YSKE YSEE	YSSIRNH YRSVRNH YRSVRNH YRSVRTH YGSVRRH YRSVRQH FYRSVRQH	SDECVIVI SDECVVVT SDECVVVT SDECVVVT SDECVVT TSDECVVT SEDECVVT

Figure S10. Multiple amino acid sequence alignment of the active site region of haloarchaeal members of the SpeE family. GGG(D/E)G and (E/D)(I/V)D motifs important for binding polyamine and the amino donor S-adenosylmethioninamine (red boxes) and the conserved proton acceptor of the reaction (green box) are highlighted. Uniprot number is on the left where HVO_B0357 (D4GQ04) and HVO_0255 (D4GZK0) are indicated in yellow.



Figure S11. *H. volcanii* HVO_0823 is a cytochrome *c*-oxidase (EC: 1.9.3.1) type helical bundle protein. A) The 3D structure of HVO_0823 (blue) generated by Phyre2-modeling compared to the X-ray crystal structure of the *caa3*-type cytochrome oxidase of *Thermus thermophilus* (PDB: 2YEV, tan, golden and salmon) at an RMSD value of 0.672 Å. The heam sites are in orange stick. B) The predicted secondary structure and primary amino acid sequence alignment of HVO_0823 compared to the C-terminal domain of the C-terminal domain of the polypeptide I+III (CaaA) of the *caa3*-type cytochrome oxidase complex (PDB: 2YEV).

3LF2	1X7H	HVO_2504
Lys118	Asn114	Gly114
Ser146	Ser144	Gly142
Thr159	Tyr157	Tyr155
Arg163	Lys161	Lys159



A)

F

4GZ_0ZI0	APAKIDL	NAG VA	DGRRNI	SEIDDA	MUTE A TI	TUALCE	/KVVQA	TIGUM	NUNSET	KVAMVS	STRAC	тррси	MGRSIA
20D2009_16718	EAGRIDVIN	/N <mark>N</mark> AATY	GPTGKL	DSLDSE	TIETTLF	RTNLHGPI	1LVT <mark>K</mark> H	IALSL	TVRDGG	VVTLS	sgsgç	FDGGI	DTGHLP <mark>Y</mark>
IQ2380A	AEGQUDII	/N <mark>N</mark> AGVR	GPTTSL	EDAPID	EIDQTFA	WNERGP	/VLTKY	ALPL	REQAGA	VVNLS	SGLG7	AIGPGM	DGGSPP <mark>Y</mark>
rnAC3483	EQGRID	/N <mark>N</mark> AGVM	DSREPL	DGMPTD	VIDHTLE	DTNERGAV	/LMTKY	ALPL	LAEAGG	IVAMS	SGLG7	AITESQ	SGGTPA <mark>Y</mark>
IAH_0704	EQGRID	/N <mark>N</mark> AGVM	DSREPL	DEMPTD	VIDHTLE	DTNERGAV	/LMTKY	ALPL	LAEEGG	VVTMA	SGLG7	AITESQ	SGGTPA <mark>Y</mark>
Iuta_2460	AGGQIIDLIN	/N <mark>N</mark> AGVG	GAGAGI	EDIDSE	DFDRILE	VNLRGP S	SLLSKY	ALPH	LETDAP	IVNVS	SGVG1	LAEPI	esempa <mark>y</mark>
ILRTI_01047	AGGQIIDVIA	/N <mark>N</mark> AGVG	GAGAGL	TDVDSE	AFDRILE	VNLRGP S	SLLSKY	ALPH	LETDAP	IVNVS	SGVG1	LADPI	esempa <mark>y</mark>
Hbor_09100	SAGRIDII	/N <mark>N</mark> AGIS	GGDGDI	VAESVT	DIDQTLA	ATNILRG PI	1LVCKH	IAVPL	VQDEGG	VINVS	SGMC7	ALNEAQ	SGGSPS <mark>Y</mark>
IQ2771A	IAGRIDIV	/N <mark>N</mark> AGIG	GGDGDI	VSESID	DINRTLS	STNERGPI	1LVCKH	ITVPL	LQNDGG	IVNVG	SALGN	ILSREQ	SGGNPA <mark>Y</mark>
IVO_2504	DQNRIDVI	/N <mark>N</mark> AGVG	GEGEDI	VAEPTE	RIDRTLO	VNI RGPI	1LLCKH	IAVPL	LQSDGG	VVNVS	SGMC7	ALGEEQ	SGGSPA <mark>Y</mark>
		86								1	.39		155
4G2 0210	RASKTALN	(F <mark>TI</mark> VSMK	KEAIED	OISFVI	LHPGWVK		APVEI	PDSVK	GMRNVV	DALTLE	TSCRE	TFDGE	LLPW
4G2_0210 20D2009 16718	RAS <mark>K</mark> TALN GVSKAGVN	IFTVSMK	KEAIED TOYPN-	QISFVI	LHPGWVF VCPGWVF		APVE1	PDSVK	GMRNVV	DALTLE	TSGRE	TIFDGE	LLPW V IE W
4G2_0210 2OD2009_16718 4Q2380A	RASK <mark>TALN GVSK</mark> AGVN RVSKTGLN	AFTVSMK AFTDALS SFTAYLH	KEAIED TQYPN- GEYATE	QISFVI -ILVNA GLISNA	LHPGWVK VCPGWVF ARPGWVC	TSMGGDI TDMGGS(RAPVEI GAPRSV GAPRTE	PDSVK ÆKGAE PTEGAD	GMRNVV TPVMLA TPVMLA	DALTLE REEPGN REKPNA	TSGRE PSGKI PAGRO	TIFDGE WRDRS WRDRM	LLPW VIEW PIGW
4G2_0210 2OD2009_16718 4Q2380A crnAC3483	RASKTALN GVSKAGVN RVSKTGLN RISKTGVN	AFTVSMK AFTDALS SFTAYLH SLTKYLD	KEAIED TQYPN- GEYATE GEYAAD	QISFVI -LLVNA GLISNA GLIAN <mark>S</mark>	LHPGWVF VCPGWVF ARPGWVC VCPGYVC	TSMGGDI TDMGGS TELATPO TDMTEGS	apvei Saprsv Saprte Saprte	PDSVK ÆKGAE PTEGAD PEKGAE	GMRNVV TPVWLA TPVWLA TPVWLA	DALTLE RFEPGN RFKPNA RFQPDA	TSGRE PSGKI PAGRO PSGRE	FIFDGE WRDRS WRDRM WRDRA	LLPW VIEW PIGW EIEW
4G2_0210 20D2009_16718 4Q2380A crnAC3483 4AH 0704	RASKTALN GVSKAGVN RVSKTGLN RISKTGVN RISKTGVN	AFTVSMK AFTDADS SFTAYDH SLTKYDD SLTKYDD	KEAIED TQYPN- GEYATE GEYAAD GEYAAD	QISFVI -LLVNA GLISNA GLIANS GLIANS	LHPGWVF VCPGWVF ARPGWVC VCPGYVC VCPGYVC	TSMGGDI TDMGGS TELATP(TDMTEGS TDMTEGS	APVEI SAPRSV SAPRTE SAPRTE SAPRTE	PDSVK VEKGAE PTEGAD PEKGAE PEKGAE	GMRNVV TPVWLA TPVWLA TPVWLA TPVWLA	DALTLE RFEPGN RFKPNA RFQPDA RFRPDA	TSGRE PSGKI PAGRO PSGRE PSGRE	FIFDGE WRDRS WRDRM WRDKA WRDRA	LLPW VIEW PIGW EIEW EIEW
4G2_0210 2OD2009_16718 4Q2380A crnAC3483 HAH_0704 Huta 2460	RASKTALN GVSKAGVN RVSKTGLN RISKTGVN RISKTGVN RISKAGIN	AFTVSMK AFTDALS SFTAYLH SLTKYLD SLTKYLD ALTVSLD	KEAIED TQYPN- GEYATE GEYAAD GEYAAD QTYGED	QISFVI -LLVNA GLISNA GLIANS GLIANS GLIANS	LHPGWVK VCPGWVF ARPGWVQ VCPGYVQ VCPGYVQ VDPGWV2	TSMGGDI TDMGGSC TELATPC DDMTEGS DDMTEGS	RAPVEI GAPRSV GAPRTE SAPRTE SAPRTE EAPREE	PDSVK VEKGAE PTEGAD PEKGAE PEKGAE PQKGAE	GMRNVV TPVWLA TPVWLA TPVWLA TPVWLA TPVWLA TPIWLS	DALTLE RFEPGN RFKPNA RFQPDA RFRPDA RFGPGS	TSGRI PSGKI PAGRO PSGRI PSGRI PSGLI	FIFDGE WRDRS WRDRM WRDKA WRDRA WRDRA	LLPW VIEW PIGW EIEW EIEW VIDF
4G2_0210 COD2009_16718 4Q2380A CrnAC3483 HAH_0704 Huta_2460 HLRTI 01047	RASKTALN GVSKAGVN RVSKTGLN RISKTGVN RISKTGVN RISKAGIN RISKAGIN	MFTVSMK AFTDALS SFTAYLH SLTKYLD SLTKYLD ALTVSLD ALTVSLD	KEAIED TQYPN- GEYATE GEYAAD GEYAAD QTYGED QTHGED	QISFVI -ILVNA GLISNA GLIANS GLIANS GLIANS GLIANS	LHPGWVF VCPGWVF ARPGWVC VCPGYVC VCPGYVC VDPGWV2 VDPGWV2	TSMGGDI TELATPO TELATPO TDMTEGS TDMTEGS TDLGGEI	APVEI APRSV APRTE APRTE APRTE APREE	PDSVK VEKGAE PEKGAE PEKGAE PQKGAE PEKGAE	GMRNVV TPVWLA TPVWLA TPVWLA TPVWLA TPIWLS TPIWLS	DALTLE RFEPGN RFKPNA RFQPDA RFRPDA RFGPGS RFGPGS	TSGRE PSGKI PAGRO PSGRE PSGRE PSGLE PSGLE	FIFDGE WRDRS WRDRM WRDKA WRDRA WKDRQ WKDKQ	LLPW VIEW PIGW EIEW VIDF VIDF
462_0210 2002009_16718 402380A crnAC3483 4AH_0704 4Uta_2460 HLRTI_01047 HDoT 09100	RASKTALN GVSKAGVN RVSKTGLN RISKTGVN RISKTGVN RISKAGIN RISKAGIN RISKTGLN	AFTUSMK AFTUALS GFTAYLH GLTKYLU GLTKYLU ALTUSLU GLTKYLU GLTKYLU	KEAIED TQYPN- GEYATE GEYAAD GEYAAD QTYGED QTHGED GEYGDD	QISFVI -ILVNA GLISNA GLIANS GLIANS GLIANS GLIANS GLIANS	LHPGWVR VCPGWVF ARPGWVC VCPGYVC VCPGYVC VDPGWVP VDPGWVP VCPGWVF	USMGGDI UDMGGSC UELATPC UDMTEGS UDMTEGS UDMTEGS UDLGGEI UDLGGEI	APVEI APRSV APRTE APRTE APREE APREE APREE APREE APREE	PDSVK PEKGAE PEKGAE PEKGAE PEKGAE PEKGAE PEKGAE	GMRNVV TPVWLA TPVWLA TPVWLA TPVWLA TPIWLS TPIWLS TPIWLS	DALTLE REEPGN REKPNA REQPDA RERPDA REGPGS REGPGS REKPES	TSGRE PSGKI PAGRC PSGRE PSGRE PSGLE PSGLE PSGYE	FIFDGE WRDRS WRDRM WRDKA WRDRA WKDKQ WKDKQ WRDQS	LLPW VIEW PIGW EIEW EIEW VIDF VIDF VIDF
462_0210 2002009_16718 402380A rrnAC3483 4AH_0704 4uta_2460 4LRTI_01047 Hoor_09100 402771A	RASKTALN GVSKAGVN RVSKTGLN RISKTGVN RISKAGIN RISKAGIN RISKTGLN RVSKSGLN	AFTUSMK AFTUALS SFTAYLH SLTKYLU SLTKYLU ALTVSLU SLTKYLU SLTAYLU	KEAIED TQYPN- GEYATE GEYAAD GEYAAD QTYGED QTHGED GEYGDD GEYGSD	QISFVI -ILVNA GLISNA GLIANS GLIANS GLIANS GLIANS GLLANS GLIANS	LHPGWVK VCPGWVF ARPGWVC VCPGYVC VCPGYVC VDPGWVA VDPGWVA VCPGWVF VCPGWVF	A SMGGDI A DMGGS(A ELATP(DMTEGS A DMTEGS A DLGGEI A DLGGEI A DMGGEI A DMGGEI	APVEI SAPRTE SAPRTE SAPRTE SAPREE SAPREE SAPREE DADRSV	PDSVK ZEKGAE PEKGAE PEKGAE PEKGAE PEKGAE ZEKGAE ZEKGAE ZEKGAE	GMRNVV TEVWLA TEVWLA TEVWLA TEVWLA TEIWLS TEIWLS TEVWLC TETWLC	DALTLE REEPGN REKPNA REQPDA RERPDA REGPGS REGPGS REKRES RERPES	TSGRE PSGKI PSGRE PSGRE PSGLE PSGLE PSGLE PSGLE PSGLE PSGLE	FIFDGE WRDRS WRDRA WRDRA WRDRA WKDRQ WKDRQ WRDQS WRDQS	LLPW VIEW PIGW EIEW EIEW VIDF VIDF VIDF VIDW VIEW
4G2_0210 20D2009_16718 4Q2380A rrnAC3483 4AH_0704 4Lta_2460 4LRTI_01047 4bor_09100 4Q2771A 4VO_2504	RASKTAIN GVSKAGVN RUSKTGIN RISKTGVN RISKTGVN RISKAGIN RISKAGIN RISKTGIN RISKTGIN	AFTVSMK AFTDAUS SFTAYDH SLTKYDD SLTKYDD ALTVSDD SLTKYDD SLTAYDD SLTAYDD	KEAIED TQYPN- GEYATE GEYAAD GEYAAD QTYGED QTHGED GEYGDD GEYGSD GQYGDD	QISFVI -ILVNA GLISNA GLIANS GLIANS GLIANS GLIANS GLIANS GLIANS GLIANS	LHPGWVK VCPGWVF ARPGWVC VCPGYVC VCPGYVC VDPGWVF VCPGWVF VCPGWVF VCPGWVF	A DMGGDI A DMGGS(A ELATP(DMTEGS A DMTEGS A DLGGEI A DLGGEI A DMGGEI A DMGGEI A DMGGEI	APVEI SAPRSV SAPRTE SAPRTE SAPREE SAPREE DADRSV DADRSI SANRPV	PDSVK ZEKGAE PEKGAE PEKGAE PEKGAE PEKGAE ZERGAE KTGAE	GMRNVV TEVWLA TEVWLA TEVWLA TEVWLA TEIWLS TEIWLS TEIWLS TEVWLC TETWLC	DALTLE REEPGN REKPNA REQPDA REGPGS REGPGS REGPGS REKPES RERES TEKPGA	TSGRE PSGKI PAGRO PSGRE PSGLE PSGLE PSGLE PSGLE PSGLE PSGLE PSGLE PSGLE	FIFDGE WRDRS WRDRA WRDRA WRDRA WKDRQ WKDRQ WRDQS WRDRA WRDRE	LLPW VIEW PIGW EIEW VIDF VIDF VIDF VIDW VIEW VIEW

Figure. S12. HVO_2504 is a putative oxidoreductase of the short-chain dehydrogenase/reductase (SDR, IPR002347) family. Panel A) 3D homology model of HVO_2504 (blue ribbon) overlaid on the X-ray crystal structures of an actinorhodin polyketide ketoreductase (PDB: 1X7H, tan) and a putative short chain oxidoreductase (Q9HYA2, PDB: 3LF2, brown). The catalytic tetrad Asn114-Ser144-Tyr157-Lys161 of the ketoreductase is highlighted in red in the structure and is compared to the atypical residues of HVO_2504 and Q9HYA2, with residues conserved with the tetrad indicated in green in the table inset. Panel B) Multiple amino acid sequence alignment of arCOG01266 members with conserved residues of the TGxxxGxG motif or within the central cavity of the HVO_2504 3D-model highlighted in red. Other conserved residues are in black/grey.



3CMN	HVO_2770
Tyr168	Tyr132
Asp169	Asp133
His210	His174
Glu211	Glu175
His214	His178
Glu217	Glu181
Glu294	Glu224

Figure S13. DUF2342 protein HVO_2770 is related to zinicin-like metallopeptidase type 2 (IPR018766) and F420 biosynthesis associated (IPR022454) families. The 3D structure of HVO_2770 (blue) generated by Phyre2-modeling against the PDB database was compared to the x-ray crystal structure of a putative hydrolase with a novel fold from *Chloroflexus aurantiacus* (PDB: 3CMN, tan) at an RMSD value of 0.354 Å. Conserved residues of a putative metal-binding active site including the HExH motif (red) and other residues (green) are indicated in the table inset.



His164

Asp66

His67

Asp187

His230

1

2

2

2

2

His171

Asp74

His75

Asp190

His233

Figure S14. *H. volcanii* HVO_1041 is a metallo-β-lactamase superfamily protein. The 3D structure of HVO_0823 (blue) generated by Phyre2-modeling against the PDB database was compared to the x-ray crystal structures of related metallo-β-lactamase superfamily proteins including from left to right: (i) TTHA1429 of unknown function from *Thermus thermophilus* HB8 (PDB: 2ZO4, tan) at an RMSD value of 0.589 Å, (ii) a mitochondrial glyoxalase II from *Arabidopsis thaliana* (PDB: 1XM8) at an RMSD value of 1.085 Å, and (iii) a human mitochondrial endoribonuclease (PDB: 4AD9) at an RMSD value of 0.931 Å. Residues that coordinate the two Zn²⁺ ions of TTHA1429 and the analogous residues of HVO_1041 are indicated in green in the table inset.

Strain, plasmid, or primer	Description ^a	Source or Ref.
<u>E. coli strains</u> :		
TOP10	F⁻ recA1 endA1 hsdR17(rĸ⁻mĸ⁺) supE44 thi-1gyrA relA1	Invitrogen
GM2163	F- ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6	New England
	hisG4 rfbD1 rpsL136 dam13::Tn9 xylA5 mtl-1 thi-1 mcrB1 hsdR2	Biolabs
<u>H. volcanii strains</u> :		
H295	ΔpyrE2 bgaHa-Kp ΔtrpA Δmre11 Δrad50	[76]
Transposon library	H295 with random Mu-like transposon insertions	[11]
DS70	wild-type isolate DS2 cured of plasmid pHV2	[77]
H26	DS70 $\Delta purE2$	[78]
HM1041	H26 $\Delta samv1$	[79]
HM1052	$H_{26} \Lambda_{\mu}h_{a}A$	[79]
GZ130	$H_{26} \Lambda ncm A_1$	[72]
CZ114	$H26 \operatorname{Answ}((-\operatorname{now} A2))$	[72]
GZ114	$\frac{1120}{2} \Delta psmC (-psmA2)$	[72]
GZ109	$H_{26} \Delta punA/1$	[72]
GZ108	H26 $\Delta panB/2$	[72]
<u>Plasmids</u> :		
pTA131	Ap ^r ; pBluescript II containing P _{fdx} -pyrE2	[78]
Primers:		
M1-1F (TL Inverse Forward)	5'-cetegacetegtgttcategtegeegeegacgace-3'	This study
M1-1R (TL Inverse Reverse)	5'-caaatattatacgcaaggcgacaaggt-3'	This study
M1-2F (TL Nested Forward)	5′-tcgaccgaatcatggaacaggtct-3′	This study
M1-2R (TL Nested Reverse)	5'-tacaacagtactgcgatgagtggc-3'	This study
M2-1F (HvTL_map_1st)	5'-atccggaacgcacataactgg-3'	This study
M2-1R (HvTL_map_degen)	5'-ctcggcattcctgctgaaccgctcttccgatctnnnnnnnnntgacg-3'	This study
M2-2F (HvTL_map_2nd)	5'-gcatttatcgtgaaacgctttcg-3'	This study
M2_2R (HvTL_map_nested)	5'-tcctgctgaaccgctcttcc-3'	This study
C-HVO_0255_p	5'-ggcacgtcttcgactacacgtc-3'	This study
C-HVO_0823_p	5'-gaccgtctcgtggtcgtaccg-3'	This study
C-HVO_1003_p	5'-agacgagcttccagttcgagacgttcatcga-3'	This study
C-HVO_1091_p	5'-ttccatcagcggcgagcgagagcgcttgt-3'	This study
C-HVO_1092_p	5'-acctgaccgttctcgggttcgactacgccga-3'	This study
C-HVO_1956_p	5'-tccgagagaatcttttccgtgagcgt-3'	This study
C-HVO_2194_p	5'-cgtatcgctttttcgcggctgg-3'	This study
C-HVO_2469_p	5'-cgtccgtctactcaccgaaacca-3'	This study
C-HVO_2374_p	5'-ggtgacgtcgagactcatcgctt-3'	This study
C-HVO_2504_p	5'-agcggtcagctgttcgacatcgagacgctcaccga-3'	This study
C-HVO_2653_p	5'-tacgacaacttcagcaagaccgag-3'	This study
C-HVO_2770_p	5'-ggcacgtcttcgactacacgtc-3'	This study
C-HVO_C0005_p	5'-ttcgagaaccgcgaactgcccgacttta-3'	This study
Hvo_1957 RT-fw	5′-ggaactcgacgactactcc-3′	This study
Hvo_1957 RT-rv	5'-cagtactggtagtccgtgaa-3'	This study
Hvo_2469-RT-fw	5'-tgattacctacgcctcctac-3'	This study
Hvo_2469 RT-rv	5'-aagctgttcaggaagacgat-3'	This study

Table S1. Strains, plasmids, and oligonucleotide primers used in this study^a.

n, denotes any nucleotide in the degenerate primer. *H. volcanii pyrE2* (orotate phosphoribosyltransferase, HVO_0333), *trpA* (tryptophan synthetase A, HVO_0789), *mre11* (double stranded break repair protein, HVO_0853), *samp2* (ubiquitin-like SAMP2, HVO_0202), *ubaA* (E1, HVO_0558), *psmA* (or *psmA1*, α1 of 20S proteasome, HVO_1091), *psmC* (or *psmC1*, α2 of 20S proteasome, HVO_2923), *panA/1* (PanA/1 proteasome-associated AAA ATPase, HVO_0850), and *panB/2* (PanB/2 proteasome-associated AAA ATPase, HVO_1957).

Supplemental Table S2. Cycling conditions for the inverse-nested two-step PCR (INT-PCR) and semi-random two-step PCR (ST-PCR).

PCR step	Initial	Amplification cycle	Final extension
	Denaturation		
INT-PCR	98 °C, 30 sec	35 cycles of:	72 °C (5 min)
step 1 and step 2		denaturation (98 °C, 10 sec) annealing (65	
		°C, 30 sec) elongation (72 °C, 2 min)	
ST-PCR step 1	94 °C, 5 min	5 cycles of:	72 °C (5 min)
		denaturation (94 °C, 30 sec) annealing (42	
		or 50 °C, 30 sec) elongation (72 °C, 3 min)	
		25 cycles of:	
		denaturation (94 °C, 30 sec) annealing (55	
		°C, 30 sec) elongation (72 °C, 3 min)	
ST-PCR step 2	94 °C, 1 min	30 cycles of:	72 °C (5 min)
		denaturation (94 °C, 30 sec) annealing (55	
		°C, 30 sec) elongation (72 °C, 3 min)	

*All reactions were preheated to the initial denaturation temperature immediately prior to PCR.

Table S3. Identification of transposon insertion sites on the genome of *H. volcanii* mutant strains found hypertolerant of hypochlorite stress.

Locus tag	Description	Isolate	Site of	General	General	Predicted
no. (gene)		no.	transposon	method	function	location
HVO_0255	Spermidine	15A	Intra 320 bp	INT-PCR	Spermidine	Membrane
(speE)	synthase		before stop	(BmtI to	biosynthesis	(7 TM
	(EC:2.5.1.16) of			BsphI)	from	helices)
	COG4262 super				putrescine	
	family which				(step 1)	
	includes					
	HVO_B0357					
HVO_0823	DUF420 family,	37A	Intra 45 bp	INT-PCR	Electron	Membrane
	multipass IM		before stop	(Ndel to	transfer	(5 IM
	protein with			inner)	the cell	nences)
	cytochrome c				membrane	
	type belical bundle				memorane	
	Phyre2-based model					
	to PDB: 10LE					
HVO_1003	ZIP family transport	57	Intra 22 bp	INT-PCR	Metal ion	Membrane
(gufA2)	protein		after start	(HindIII to	transport	(7 TM
				NdeI)		helices)
HVO_1041	Metallo-beta-	38A	Intra 47 bp	ST-PCR	Metal binding	soluble
	lactamase		after start	(42 or	hydrolase/	
	(IPR001279) domain			50°C)	oxidoreductase	
1001	protein	50	T (E1		200	1.1.1
HVO_1091	Proteasome subunit	59	Intra 5 bp	INI-PCK	205	soluble
(psmA1)			alter start	(FIIIIdIII to Ndel)	subunit	
HVO 1957	Proteasome	36A	Inter 37 bp	INT-PCR	AAA ATPase	soluble
(panB2)	activating		before start	(NdeI to	associated with	
	nucleotidase B2			NheI)	proteasome	
HVO_2145	Halocyanin blue	65A	Intra 49 bp	ST-PCR	Mobile	Membrane
(hcpF)	(type I) copper		after start	(50°C)	electron carrier	associated
	redox proteins					(Tat
						lipoprotein)
HVO_2194	Sır2-type NAD-	16A	Intra 260 bp	INT-PCR	Deacetylation	soluble
(str2)	dependent protein		after start	(Bmtl to	of acetylated	
	deacetylase			Bsph1)	lysine residues	
HVO 2374	N-terminal SpoVT	334	Intra 221 bn	INT_PCP	Regulation of	Membrane
(nhol]2)	AbrB DNA binding	JUA	before stop	(Ndel to	phosphate	associated
(,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	domain (IPR007159)			NheI)	uptake	(PhoU2.
	and PhoU domain				upune	regulator of
	(IPR026022)					Pi transport;
HVO_2375	ABC-type transport	83A	Intra 169 bp	ST-PCR	Solute binding	PstS1, Tat
(<i>pstS1</i>)	system periplasmic		before stop	(50°C)	protein	lipoprotein)
	substrate-binding				associated with	
	protein (Probable				phosphate	
	substrate				transport	
	phosphate)					

HVO_2441 (<i>dppB5</i>)	ABC-type transport system permease	41A	Intra 503 bp after start	ST-PCR (42°C)	Transporter (permease)	Membrane (5 TM
	DppB homolog				protein	helices)
HVO_2469	SNF family	35A, 7	Inter before	INT-PCR	Na ⁺ -coupled	Membrane
	transport protein;		start by 42	(NdeI to	transporter	(11 TM
	Phyre2-based 3D		bp for	NheI for	(e.g., Na+-	helices)
	model		isolate 35A;	isolate	dependent	
			Inter before	35A and	amino acid	
			start by 250	XhoI to	uptake)	
			bp for	BclI for		
			isolate 7	isolate 7)		
HVO_2504	Putative	62	Intra 23 bp	INT-PCR	NAD- or	soluble
	oxidoreductase		before stop	(HindIII to	NADP-	
	(Short-chain			NdeI)	dependent	
	dehydrogenase				oxidoreductase	
	family)					
HVO_2653	Uncharacterized	31A	Intra 105 bp	INT-PCR	unknown	Membrane
	protein (not		after start	(NdeI to		(3 TM
	amenable to Phyre2-			NheI)		helices)
	3D modeling					
	01/16/18)					
HVO_2770	DUF2342 protein of	40, 30	Intra 140 bp	INT-PCR	Metallohydrol	soluble
	the zinicin-like		after start	(HindIII to	ase	
	metallopeptidase		for isolate	NdeI)		
	type 2 (IPR018766),		40; 142 bp			
	and F420		after start			
	biosynthesis		for isolate			
	associated		30			
	(IPR022454) families					
HVO_A0494	TsgA6, ABC-type	63A	Intra 577 bp	ST-PCR	Transport	Membrane
(tsgA6)	transport system		before stop	(42 or	(solute	associated
	periplasmic			50°C)	binding)	(Tat
	substrate-binding				protein	lipoprotein)
	protein (probable					
	substrate sugar)					
HVO_B0012	Compatible solute	67A	Intra 591 bp	ST-PCR	Transporter	Membrane
(betT)	transport protein		after start	(42 or	(choline/glycin	(12 TM
	(Probable substrate			50°C)	e betaine)	helices)
	choline/glycine					
	betaine)					
HVO_C0005	N-acylglucosamine	93	Intra 93 bp	INT-PCR	Metabolism of	soluble
	2-epimerase (EC		before stop	(HindIII to	carbohydrates	
	5.1.3.8) orthologous			NdeI)	associated with	
	group				glycosylation	
	POG093W01SY				of proteins	
					and/or lipids	

TM, transmembrane; Tn, transposon, NheI was a high fidelity (HF) enzyme; Intra, intragenic or within the open reading frame; Inter, intergenic; stop, stop codon; start, translational start site. Method 1, inverse-nested two-step PCR (INT-PCR) with the restriction enzymes used prior to ligation in parenthesis. Method 2, semi-random two-step PCR (ST-PCR) approach with the annealing temperature indicated in parenthesis.