

## Supplementary Information

**Supplementary Table 1.** Primers used in this study.

Primer	Primer sequence	Primer pair: PCR product, vector, plasmid generated, protein encoded
1: HVO_0202 <sup>a</sup> KpnI up	5'-CGTTggtaccATGAACGTGACCGTCGAGG-3' KpnI site lowercase	Primers 1 + 2: KpnI to BlnI HVO_0202 into pJAM939 to generate pJAM949; expresses FLAG-SAMP2
2: HVO_0202 BlnI down	5'-TTAATgctcagcTACCCGCCTTTGATGAGG-3' BlnI site lowercase	
3: HVO_0383 FLAG NdeI KpnI up	5'-TTAATcatatg <b>GACTACAAGGACGACGACGACAAG</b> ggtaccA GCGCCCAGACGAACCTCG-3' NdeI and KpnI sites lowercase; FLAG-tag underlined and in bold	Primers 3 + 4: NdeI to BlnI FLAG-HVO_0383 into pJAM202 to generate pJAM939; expresses FLAG-HVO_0303
4: HVO_0383 BlnI down	5'-TTAATgctcagcTCACGCTCCACCCGCTTGTCGAGC-3' BlnI site lowercase	
5: HVO_2177 KpnI up	5'-AAggtaccAAAaGaCTCCGtGTCCTCGCCGCGAC-3' KpnI site and silent mutations in lowercase	Primers 5 + 6: KpnI to BlnI HVO_2177 into pJAM939 to generate pJAM941; expresses FLAG-HVO_2177
6: HVO_2177 BlnI down	5'-TTAATgctcagcATCAGCCCCCGCGACC-3' BlnI site lowercase	
7: HVO_2178 KpnI up	5'-AAggtaccGTGGCGCGCCGACAGC-3' KpnI site lowercase	Primers 7 + 8: KpnI to BlnI HVO_2178 into pJAM939 to generate pJAM943; expresses FLAG-HVO_2178
8: HVO_2178 BlnI down	5'-TTAATgctcagcTCATCCCCCGCGCGC-3' BlnI site lowercase	
9: HVO_2619 KpnI up	5'-AAggtaccGAGTGGAAGCTGTTCCGCCGACCTCG-3' KpnI site lowercase	Primers 9 + 10: KpnI to BlnI HVO_2619 into pJAM939 to generate pJAM947; expresses FLAG-SAMP1
10: HVO_2619 BlnI down	5'-TTAATgctcagcCTAGCCGCGCTGACCGG-3' BlnI site lowercase	
11: HVO_2619 ΔGG down	5'-TTAATgctcagcCTAGGCGCGCTGACCGG-3' BlnI site lowercase	Primers 9 + 11: KpnI to BlnI HVO_2619ΔGG into pJAM939 to generate pJAM951; expresses FLAG-SAMP1ΔGG
12: HVO_0202 ΔGG down	5'-TTcATgctcagcTATTTGATGAGGCGGAG-3' BlnI site lowercase	

<sup>a</sup>HVO\_2619 (SAMP1), HVO\_0202 (SAMP2) and HVO\_2177, β-grasp proteins related to Ub/ThiS/MoaD; HVO\_2178, MoaD-related protein; HVO\_0383, σ<sup>70</sup> region 4 homolog; FLAG-, N-terminal FLAG-tag.

**Supplementary Table 2.** Strains and plasmids used in this study.

Strain or plasmid	Description <sup>a</sup>	Source or reference
<b>Strain</b>		
<b><i>E. coli</i></b>		
DH5 $\alpha$	F <sup>-</sup> <i>recA1 endA1 hsdR17</i> (r <sub>k</sub> <sup>-</sup> m <sub>k</sub> <sup>+</sup> ) <i>supE44 thi-1 gyrA relA1</i>	Life Technologies
GM2163	F <sup>-</sup> <i>ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galk2 galT22 mcrA dcm-6 hisG4 rfbD1 rpsL 136 dam13 ::Tn9 xylA5 mtl-1 thi-1 mcrB1 hsdR2</i>	New England Biolabs
<b><i>H. volcanii</i></b>		
DS2	Dea Sea isolate	Mullakhanbhai and Larsen <sup>1</sup>
DS70	DS2 cured of plasmid pHV2	Wendoloski <i>et al.</i> <sup>2</sup>
H26	DS70 <i>pyrE2</i>	Allers <i>et al.</i> <sup>3</sup>
GZ108	H26 <i>panB</i>	Zhou <i>et al.</i> <sup>4</sup>
GZ109	H26 <i>panA</i>	Zhou <i>et al.</i> <sup>4</sup>
GZ130	H26 <i>psmA</i>	Zhou <i>et al.</i> <sup>4</sup>
GZ132	GZ108 <i>panA</i>	Zhou <i>et al.</i> <sup>4</sup>
GZ133	GZ108 <i>psmA</i>	Zhou <i>et al.</i> <sup>4</sup>
GZ134	GZ130 <i>panA</i>	Zhou <i>et al.</i> <sup>4</sup>
GZ138	H26 P <sub><i>tnaA</i></sub> - <i>psmB</i>	Zhou <i>et al.</i> <sup>4</sup>
GZ145	H26 <i>hvo_1870 hvo_1862</i>	Zhou, unpublished
GZ100	H26 <i>hvo_1870</i>	Zhou, unpublished
<b>Plasmid</b>		
pBAP5010	Ap <sup>r</sup> ; Nv <sup>r</sup> ; <i>H. volcanii</i> - <i>E. coli</i> shuttle expression plasmid	Jolley <i>et al.</i> <sup>5</sup>
pJAM202	Ap <sup>r</sup> ; Nv <sup>r</sup> ; pBAP5010 with P2 <sub><i>rnm</i></sub> - <i>psmB</i> -His6	Kaczowka and Maupin-Furlow <sup>6</sup>
pJAM202c	Ap <sup>r</sup> ; Nv <sup>r</sup> ; pJAM202 without P2 <sub><i>rnm</i></sub> - <i>psmB</i> -His6 insert	Zhou <i>et al.</i> <sup>4</sup>
pJAM939	Ap <sup>r</sup> ; Nv <sup>r</sup> ; pJAM202 with P2 <sub><i>rnm</i></sub> -Flag- <i>hvo_0383</i> ; used for inserting genes for N-terminal FLAG-tag using KpnI and BlnI	This study
pJAM941	Ap <sup>r</sup> ; Nv <sup>r</sup> ; pJAM939 with P2 <sub><i>rnm</i></sub> -Flag- <i>hvo_2177</i>	This study
pJAM943	Ap <sup>r</sup> ; Nv <sup>r</sup> ; pJAM939 with P2 <sub><i>rnm</i></sub> -Flag- <i>hvo_2178</i>	This study
pJAM947	Ap <sup>r</sup> ; Nv <sup>r</sup> ; pJAM939 with P2 <sub><i>rnm</i></sub> -Flag- <i>hvo_2619</i> (FLAG-SAMP1)	This study
pJAM949	Ap <sup>r</sup> ; Nv <sup>r</sup> ; pJAM939 with P2 <sub><i>rnm</i></sub> -Flag- <i>hvo_0202</i> (FLAG-SAMP2)	This study
pJAM951	Ap <sup>r</sup> ; Nv <sup>r</sup> ; pJAM939 with P2 <sub><i>rnm</i></sub> -Flag- <i>hvo_2619</i> ΔGG (FLAG-SAMP1ΔGG)	This study
pJAM966	Ap <sup>r</sup> ; Nv <sup>r</sup> ; pJAM939 with P2 <sub><i>rnm</i></sub> -Flag- <i>hvo_0202</i> ΔGG (FLAG-SAMP2ΔGG)	This study

<sup>a</sup>Abbreviations: Ap<sup>r</sup>, ampicillin resistance; Nv<sup>r</sup>, novobiocin resistance. Genes and encoded proteins: *panA* and *panB*, PanA and PanB AAA ATPases related to Rpt subunits of 26S proteasome; *psmA* and *psmB*,  $\alpha$ 1 and  $\beta$  subunits of 20S proteasome core particle; HVO\_1870 and HVO\_1862, S2P metalloprotease homologs; HVO\_0383,  $\sigma^{70}$  region 4 homolog; HVO\_2619 (SAMP1), HVO\_0202 (SAMP2) and HVO\_2177,  $\beta$ -grasp proteins related to Ub/ThiS/MoaD; HVO\_2178, MoaD-related protein; FLAG-, N-terminal FLAG-tag; -His6, C-terminal His6-tag.

**Supplementary Table 3. MS-identified SAMPs and SAMP-conjugates<sup>a</sup>.**

MS-identified protein (cal. kDa) <sup>a</sup>	Annotation or homologs	SDS-PAGE gel slice (kDa range) <sup>b</sup>	No. peptides/ % coverage	Mascot Score	Peptide (E value; protease cleavage)
<b>H26-pJAM947 (FLAG-SAMP1) grown in complex medium (ATCC 947):</b>					
HVO_2619 (9 kDa)	SAMP1, $\beta$ -grasp Ub-like protein modifier	5 - 10 kDa	3 / 41 %	76	G.DELALFPPVSGG.- (1.6e-3; AspN) A.DLAEVAGSRTVRV.D (4.7e-4; AspN) K.LFADLAEVAGSR.T (2e-7; trypsin)
		25 - 30 kDa	2 / 14 %	78	K.LFADLAEVAGSR.T (1.3e-7; trypsin)
		45 – 50 kDa	1 / 14 %	70	K.LFADLAEVAGSR.T (9e-7; trypsin)
HVO_0558 (29 kDa)	molybdopterin biosynthesis protein MoeB; Ub activating E1 enzyme	25 - 30 kDa	2 / 6.7 %	54	L.DGRLLFY.D (4.4e-2; AspN) P.DVSVEPVEARV.D (1.6e-3; AspN)
HVO_1864 (28 kDa)	molybdenum cofactor biosynthesis protein MoeE; Ub conjugating E2 enzyme	45 – 50 kDa	7 / 23 %	144	P.DARVPTVALAGV.D (2.4e-5; AspN) A.DVASLAAEI.D (0.1; AspN) M.DAISEELEAR.D (4.1e-5; AspN) V.DAANVVAEAETA.D (1.6e-3; AspN) R.LAAQLDGR.V (2.8e-5; trypsin) K.MDAISEELEAR.D (4.8e-08; trypsin) R.TVEDGIDR.L (5.6e-4; trypsin)
<b>H26-pJAM947 (FLAG-SAMP1) grown under nitrogen-limitation (GMM+Ala minus NH<sub>4</sub>Cl in 2.5 M NaCl):</b>					
HVO_2619 (9 kDa)	SAMP1, $\beta$ -grasp Ub-like protein modifier	5 - 10 kDa	5 / 28 %	63	A.DLAEVAGSRTVRV.D (0.00047) G.DELALFPPVSGG.- (0.0016)
		25 - 30 kDa	1 / 14 %	62	K.LFADLAEVAGSR.T (5e-6)
		45 – 50 kDa	1 / 14 %	63	K.LFADLAEVAGSR.T (4.7e-6)
HVO_0558 (29 kDa)	molybdopterin biosynthesis protein MoeB; Ub activating E1 enzyme	25 - 30 kDa	4 / 17%	79	K.AESAAAFVR.G (0.0045; trypsin) M.TLSLDATQLDR.Y (0.041; trypsin) K.LLLDEGEALDGR.L (5.8e-06; trypsin) R.GLNPDVSVEPVEAR.V (0.036; trypsin)
HVO_1864 (28 kDa)	molybdenum cofactor biosynthesis protein MoeE; Ub conjugating E2 enzyme	45 – 50 kDa	3 / 10%	54	R.TVEDGIDR.L (0.62; trypsin) K.MDAISEELEAR.D 5.2e-05; trypsin) R.LAAQLDGR.V (6.7e-06; trypsin)
HVO_A0230 (21 kDa)	MsrA, methionine-S-sulfoxide reductase	45 – 50 kDa	2 / 12%	75	R.QGPDVGTQYR.S (0.0012; trypsin) R.SIVLFHDDEQKR.Q (0.018; trypsin)
HVO_2402	glycine cleavage	50 – 60 kDa	2 / 4%	28	R.LDELSGVK.A (0.056; trypsin)

(49 kDa)	system P-protein				R.NAMADLLSR.N (0.055; trypsin)
HVO_2900 (50 kDa)	FumC, fumarate hydratase class II	50 – 60 kDa	3 / 7%	76	R.LGQEFGGYR.T (0.0012; trypsin) K.LLANSSEVFAER.F (3.4e-05; trypsin) R.AAEIMGEGIGDR.V (0.0058; trypsin)
HVO_0025 (32 kDa)	SseA, thiosulfate sulfurtransferase, Yor251cp homolog	45 – 50 kDa	3 / 13%	61	K.GPFEDIR.A (0.0077; trypsin) K.GLPLVDVR.S (0.002; trypsin) R.SPEEFSGEILAPPGLQETAQR.G (0.003; trypsin)
HVO_0359 (46 kDa)	EF-1 $\alpha$ , translation elongation factor EF-1 subunit $\alpha$	50 – 60 kDa	2 / 2%	35	K.VLEVNER.- (0.0055; trypsin) K.VLEVNER.- (0.0051; trypsin)

**H26-pJAM949 (FLAG-SAMP2) grown in complex medium (ATCC 947):**

HVO_0202 (7 kDa)	SAMP2, $\beta$ -grasp Ub- like protein modifier	10 - 15 kDa	1 / 42%	75	R.AVDLSPHEVTVLVDGRVPEDQSVEVDR.V (4.1e-07)
		30 – 40 kDa	1 / 42%	61	R.AVDLSPHEVTVLVDGRVPEDQSVEVDR.V (1.5e-05)
		60 – 80 kDa	1 / 42%	66	R.AVDLSPHEVTVLVDGRVPEDQSVEVDR.V (2.9e-06)
		100 – 125 kDa	1 / 42%	61	R.AVDLSPHEVTVLVDGRVPEDQSVEVDR.V (1.9e-05)
HVO_0558 (29 kDa)	molybdopterin biosynthesis protein MoeB; Ub activating E1 enzyme	30 – 40 kDa	21 / 70%	845	K.AESAAAFVR.G (6.9e-05) R.YLLNDVCR.F (0.00029) M.TLSLDATQLDR.Y (1.2e-07) K.LLLDEGEALDGR.L (5.6e-08) -.MTLSLDATQLDR.Y (4.8e-08) R.FEGIPLVHGAIYK.F (0.0032) R.QVVHCDDVGTTPK.A (0.00088) R.GLNPDVSVEPVEAR.V (3.4e-07) R.HIIMDEVGPEGQGR.L (6.9e-05) K.FEGQATTLVDPGPCYR.C (1.3e-07) R.LLFYDAMDMTFETVPYR.T (6.6e-10) K.SNVHEVVAGSDVVVDASDNFPTR.Y (1.3e-09) R.VDKSNVHEVVAGSDVVVDASDNFPTR.Y (1.8e-09)
HVO_0966 (35 kDa)	eif2ba translation initiation factor, aIF-2BII translation initiation factor	30 – 40 kDa	7 / 24%	357	R.GAATIADAAAR.A (2.4e-05) K.IGTSGLAVNAR.E (9.2e-06) R.YVDAIVTER.G (3.9e-06) R.GTPIMVAAQTLK.L (5.4e-06) R.TQATESDAADAFAFR.A (2.3e-09) R.DTAEVVDDDTLADLGNPTVK.N (1e-07)
HVO_0580 (36 kDa)	N-type ATP pyrophosphatase superfamily, Ncs6p homolog	30 – 40 kDa	4 / 11%	129	R.DAGEFIPR.A (3e-05) R.FADELGADK.L (7.8e-06) R.GEIQQLLK.L (0.0086) R.IREDNMLPR.D (0.0067)

HVO_1344 (27 kDa)	Shwachman-Bodian-Diamond syndrome protein	30 – 40 kDa	3 / 15%	125	R.ALEEAGFK.I (4.5e-05) K.VNNITSGTAETR.I (1.8e-06) R.NAVNPQMNDSPHPPER.I (0.004)
HVO_1250 (21 kDa)	Thiol-disulfide isomerase/thioredoxin	30 – 40 kDa	2 / 16%	124	M.VLLESDSELER.G (1e-06) R.IDDAMSPDDEVSDYEMR.T (1.4e-06)
HVO_1577 (19 kDa)	Putative winged-helix transcriptional regulator, C-terminal CBS domains	30 – 40 kDa	2 / 25%	115	K.ALQLVEGVGPK.G (0.003) R.DSEDAVKGEDIAAEVNR.N (5.1e-09)
HVO_2328 (21 kDa)	Isochorismatase	30 – 40 kDa	2 / 15%	111	R.HDSAEPDSPLR.S (1.1e-06) R.EFATVVESADLSATR.- (8e-06)
HVO_1545 (25 kDa)	DhaL, dihydroxy-acetone kinase L subunit	30 – 40 kDa	4 / 18%	109	R.EAVLDALDNVAER.L (3.1e-07) R.EEFVEMEPNEVVK.N (0.025) K.SIEQDDLPPLEALAK.A (0.00097)
HVO_A0230 (21 kDa)	MsrA, methionine-S-sulfoxide reductase	30 – 40 kDa	2 / 5%	87	R.QGPDVGTQYR.S (7.2e-05)
HVO_0481 (38 kDa)	GAPDH, glyceraldehyde-3-phosphate dehydrogenase type I	45 – 50 kDa	4 / 15%	264	R.VALNGFGR.I (5.6e-05) K.VVISAPPKGDEPVK.Q (3.9e-05) R.VPVPNGSLTELVR.L (1.8e-09) K.SYLSAGENVDES DVVR.V (1.5e-11)
HVO_0887 (31 kDa)	PorB, 2-oxoglutarate ferredoxin oxidoreductase beta subunit	45 – 50 kDa	2 / 16%	96	K.ALANSGTDPDNTFMVAGIGCSGK.I (3.5e-05) K.FGVDQDMSDIPSGAPDDAMD LVR.E (7.4e-06)
HVO_1478 (36 kDa)	TFB, transcription initiation factor B	45 – 50 kDa	3 / 11%	93	R.MASALGVPR.S (0.00012) R.SLDEVAEVS R.V (9.5e-06) R.YISQELGLELKPVDPK.Q (0.088)
HVO_2583 (41 kDa)	HmgA, 3-hydroxy-3-methylglutaryl coenzyme A reductase	60 – 80 kDa	3 / 12%	156	K.KPAAINAVEGR.G (6.6e-07) R.GCSVINSAGGATAR.V (1.3e-06) R.LHTTPEAVAELNTR.K (0.00011)
HVO_1921 (52 kDa)	SerS, seryl-tRNA synthetase	60 – 80 kDa	5 / 12%	139	K.FAEDAYR.I (0.017) R.NMEDVDIDR.V (0.00032) K.TTSMEGTGQLPK.F (0.00077) K.QEGKDEEAQEAIDR.S (0.00048) R.SGELKEELQELETR.A (9.8e-06)
HVO_0677 (49 kDa)	AspS, aspartyl-tRNA synthetase	60 – 80 kDa	4 / 9%	131	R.VFEVGP IFR.A (0.00072) K.QLMVGSGLER.V (0.00012) R.LTYEEA IER.I (0.0053) R.LIMITMLGLENIR.E (1.1e-05)
HVO_2305 (66 kDa)	MoeA, molybdopterin biosynthesis protein	100 – 125 kDa	5 / 11%	244	R.TFVAPAVR.D (0.00037) R.IEKPGVDSL R.D (4e-05) R.TDLGNAVAALADER.G (4.9e-10) R.APALLGVGEDDPALSR.V (3.1e-09)

HVO_A0379 (76 kDa)	AgaF, N-methylhydantoinase A	100 – 125 kDa	4 / 6%	148	R.LMPVGLVADGDGETLVYPVDK.G (0.006) K.TAFADLEAEGR.E (0.011) R.EDSELTPADTR.E (8.5e-06) R.ETLTAEGVAENR.Q (3.6e-05) R.AVGENDKPELER.G (6.4e-06)
HVO_1572 (71 kDa)	GyrB, DNA gyrase B subunit	100 – 125 kDa	4 / 7%	145	R.ITVEDAAAADR.M (3.7e-06) R.EFVEYLNETH.T (0.035) K.HPDPQFEGQTK.T (0.00011) R.FEVGEPQVEEFER.V (9.5e-07)
HVO_1496 (60 kDa)	PtsI, phosphoenolpyruvate-protein phosphotransferase	100 – 125 kDa	2 / 5%	120	R.LTPSDTAQLDPER.V (4.2e-06) R.SLGPDADLFETQVR.A (3.9e-07)
HVO_1000 (75 kDa)	acetyl-CoA synthetase	100 – 125 kDa	3 / 5%	113	R.IAVVGATER.E (0.00034) R.VAPVSETEAEDMTK.E (0.00022) R.GNDPVDIGGITETIQR.L (1e-05)
HVO_0980 (64 kDa)	NdhG, NADH-quinone oxidoreductase chain c/d	100 – 125 kDa	3 / 5%	100	R.NIQANVPR.T (0.0023) R.GSGIDYDLR.R (4.3e-05) R.TVGTGVLSPPEVAK.S (0.00035)

**H26-pJAM949 (FLAG-HVO0202) grown under nitrogen-limitation (GMM+Ala minus NH<sub>4</sub>Cl in 1.5 M NaCl):**

HVO_0202 (7 kDa)	SAMP2, $\beta$ -grasp Ub-like protein modifier	30 – 40 kDa	3 / 46%	141	R.AVDLSPHEVTVLVDGRVPEDQSVEVDR.V (1.3e-07)
		45 – 50 kDa	1 / 46%	69	R.AVDLSPHEVTVLVDGRVPEDQSVEVDR.V (1.6e-06)
		60 – 80 kDa	1 / 46%	69	R.AVDLSPHEVTVLVDGRVPEDQSVEVDR.V (1.5e-06)
		100 – 125 kDa	2 / 46%	123	R.AVDLSPHEVTVLVDGRVPEDQSVEVDR.V (5.3e-07)
HVO_2682 (8 kDa)	dodecin-related protein	10 - 15 kDa	3 / 39%	126	K.ELGVELASVETR.E (2.8e-06) R.SAESFDAAVDDAINR.A (1.1e-07)
HVO_0736 (16 kDa)	DUF302, domain of unknown function	10 - 15 kDa	2 / 21%	93	R.VVLEMEHEAAIER.V (2.9e-06) M.ALPIDPSAIKPEDIGEER.V (0.0003)
HVO_0558 (29 kDa)	molybdopterin biosynthesis protein MoeB; Ub activating E1 enzyme	30 – 40 kDa	11 / 37%	491	R.YLLNDVCR.F (0.054) K.LLLDEGEALDGR.L (7.7e-08) R.QVVHCDDVDGTPK.A (8.8e-06) R.GLNPDVSVEPVEAR.V (3.5e-07) R.HIIMDEVGPEGQGR.L (3e-07) K.FEGQATTLVDPGPCYR.C (1.5e-06) K.SNVHEVVAGSDVVVDASDNFPTR.Y (2.9e-09)
HVO_1289 (14 kDa)	OsmC-like protein superfamily	30 – 40 kDa	2 / 19%	171	K.TGFDDLK.V (8.2e-06) K.VQIDADADLDDSDDLER.I (2.1e-12)
HVO_B0053 (14 kDa)	conserved hypothetical protein with C-terminal H-X <sub>3</sub> -H motif protein	30 – 40 kDa	2 / 21%	114	R.IDDDDPQAGYLR.G (4.9e-06) R.LALLNITGTTDEHAK.R (2.3e-06)
HVO_0966	eif2ba translation	30 – 40 kDa	7 / 27%	300	R.GAATIADAAAR.A (3.9e-05)

(35 kDa)	initiation factor aIF-2BII translation initiation factor				K.IGTSGLAVNAR.E (0.00022) R.YVDAIVTER.G (0.0023) R.AQADLGQVGANR.L (5e-07) R.DMSSTTVEGLR.Q (5.8e-06) R.GTPIMVAAQTLK.L (1.7e-05) R.DTAEVDDDTLADLGNPTVK.N (3e-08)
HVO_1727 (20 kDa)	TATA-box binding protein E	30 – 40 kDa	3 / 21%	154	K.VVILLFGSGK.I (2e-05) R.TDDAETAVEEIVER.I (3.2e-07) K.RTDDAETAVEEIVER.I (3.3e-05)
HVO_A0230 (21 kDa)	MsrA methionine-S- sulfoxide reductase	30 – 40 kDa	2 / 12%	95	R.QGPDVGTQYR.S (4.8e-06) R.SIVLFHDDEQKR.Q (0.016)
HVO_0025 (32 kDa)	SseA, thiosulfate sulfurtransferase, Yor251cp homolog	45 – 50 kDa	3 / 13%	90	K.GPFEDIR.A (0.00031) K.GLPLVDVR.S (0.2) R.SPEEFSGEILAPPGLQETAQR.G (9.6e-06)
HVO_0861 (44 kDa)	SufB/SufD domain protein, cysteine desulfurase activator subunit	100 – 125 kDa	6 / 15%	251	R.GVLDDEAR.S (0.00079) K.AENTTADLVTR.G (1.3e-06) R.SVYEGVQDVGR.D (5.3e-05) R.SDVETELNGEASETK.I (6.2e-05) R.ENTLMLSDQSEADASPK.L (1.9e-06)
HVO_0580 (36 kDa)	N-type ATP pyrophosphatase superfamily, Ncs6p homolog	100 – 125 kDa	6 / 16%	139	R.LIESIEAV.- (0.087) R.DAGEFIPR.A (0.0011) R.FADELGADK.L (7.9e-06) R.VDLNDCER.C (0.0015) R.GEIQQLLK.L (0.0064) R.IREDNMLPR.D (0.00035)

<sup>a</sup>MS-identified proteins are reported according to the *H. volcanii* gene locus tag from the USCS Archaeal Genome Browser (<http://archaea.ucsc.edu/>; April 2007 version) and were unique to samples prepared from strain H26 expressing either the FLAG-tagged  $\beta$ -grasp Ub-like protein SAMP1 (HVO2619, pJAM947) or SAMP2 (HVO0202, pJAM949) as indicated compared to vector alone (pJAM202c). cal kDa, molecular mass estimated from deduced polypeptide based on genome sequence in parenthesis.

<sup>b</sup>Protein samples were isolated from SDS-PAGE gel slices in the molecular mass range indicated as estimated based on migration of Precision Plus Kaleidoscope standards (BioRad).

<sup>c</sup>Proteins were digested with trypsin and/or AspN as indicated.

## Supplementary Figure Legends

### Supplementary Figure 1. $\beta$ -grasp and other small proteins expressed as FLAG-fusions in *H.*

*volcanii*. a-c) HVO\_0383, HVO\_2177, HVO\_2178, HVO\_2619 (SAMP1) and HVO\_0202 (SAMP2) were expressed as N-terminal FLAG-tagged fusions in *H. volcanii* cells grown in complex medium (CM) and nitrogen-limiting conditions (– N) at optimal and suboptimal salt concentrations of 2.5 and 1.5 M NaCl, respectively, as indicated. Cells (0.065 OD<sub>600</sub> units) were boiled (20 min) in  $\beta$ -mercaptoethanol SDS-loading buffer, separated by SDS-PAGE and analyzed by  $\alpha$ -FLAG immunoblot. Molecular mass standards indicated on left. Equivalent protein loading was confirmed by Coomassie Blue staining. Strains included: *H. volcanii* H26-pJAM939 (FLAG-HVO\_0383), H26-pJAM941 (FLAG-HVO\_2177), H26-pJAM943 (FLAG-HVO\_2178), H26-pJAM947 (FLAG-SAMP1), H26-pJAM949 (FLAG-SAMP2) and H26-pJAM202c (vector alone). Cells were grown to stationary phase at 42 °C and 200 rpm. All details on experimental procedures and strains are available in ‘Supplementary Materials and Methods’ and ‘Supplementary Tables 1 and 2’. In contrast to SAMP1 (HVO\_2619) and SAMP2 (HVO\_0202), HVO\_2177, HVO\_2178 and HVO\_0383 were not detected as protein conjugates under the experimental conditions examined.

### Supplementary Figure 2. MS/MS spectra providing additional support for SAMP2-protein

**conjugate sites.** Gly<sub>1</sub>-Gly<sub>2</sub> corresponds to the C-terminal diglycine motif of SAMP2 (G65 and G66). See legend to figure 6 for additional details.

### Supplementary Figure 3. Phylogenetic distribution of SAMP1 and SAMP2.

Sequences used to construct the alignment were obtained by performing a protein BLAST search of microbial genomes within GenBank on the NCBI website. Protein sequences were N- and C-terminally trimmed and aligned using ClustalW. Pairwise comparisons were performed between sequences and mean genetic

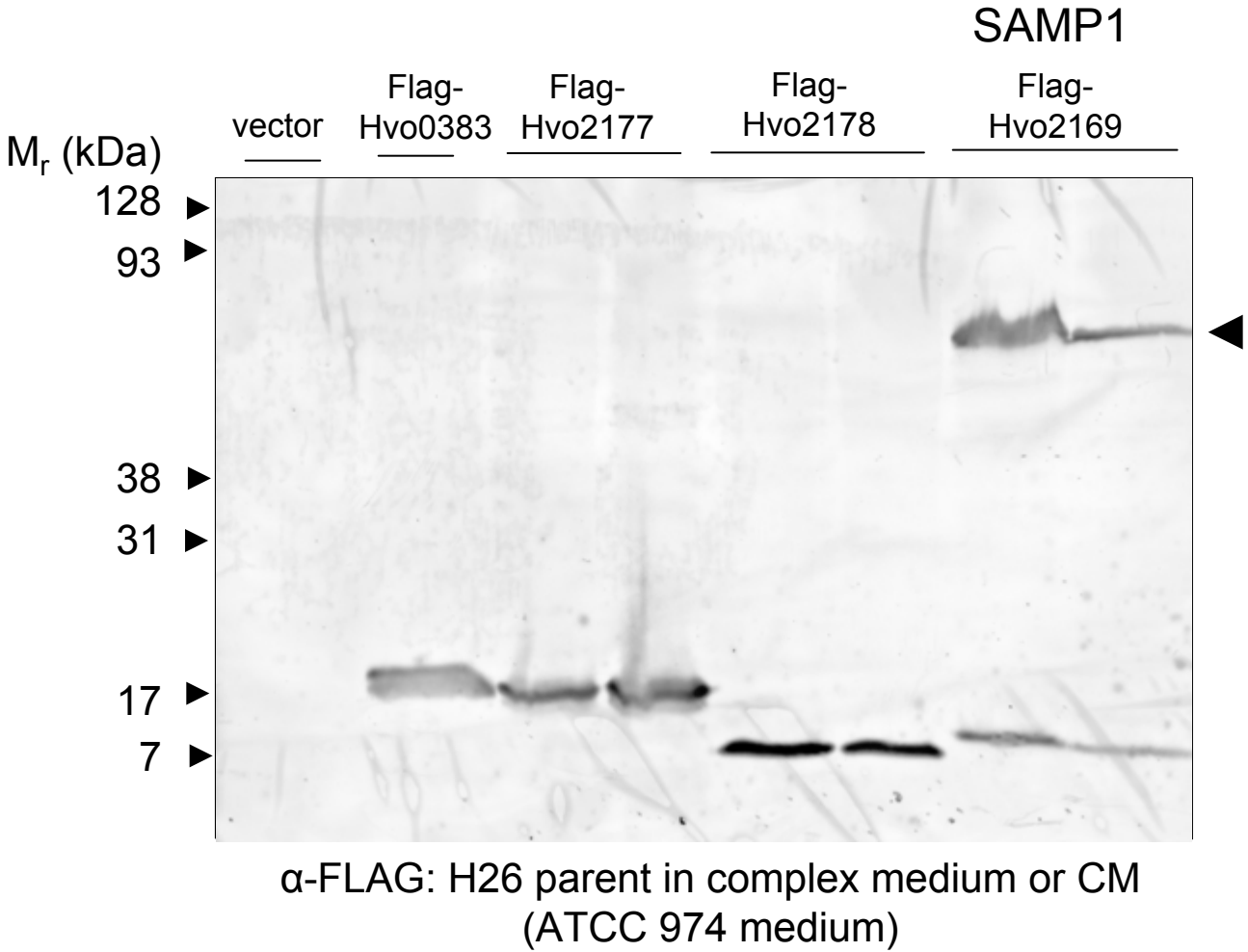


distance was evaluated using p-distance (gaps were analyzed using pairwise deletion). The best neighborhood-joining trees for *H. volcanii* (a) SAMP1 or HVO\_2619 and (b) SAMP2 or HVO\_0202 were constructed using MEGA 4.0. Bootstrap support values are indicated at the internal nodes and were obtained by performing 1,000 replicates. Organism and gene locus tag numbers for accessing sequences through the NCBI database are indicated.

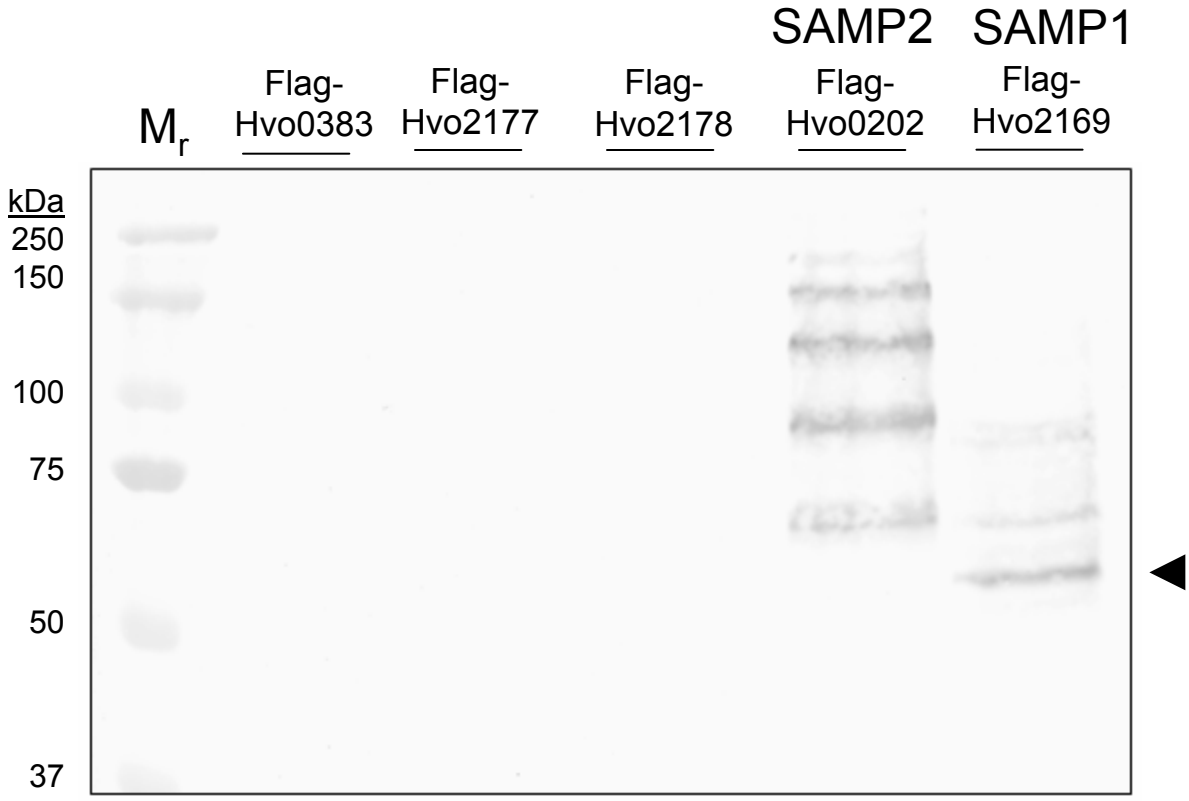
**Supplementary Figure 4. SAMP1 and SAMP2 are related to  $\beta$ -grasp proteins from diverse archaea and thermophilic bacteria.** Full length protein sequences described in Fig. 3 were aligned with ClustalW. Gene locus tag numbers indicated on left; residue numbers indicated on right and left; identical residues shaded in black; similar residues shaded in grey.

**Supplementary Figure 5. Conservation of gene order among the archaeal SAMP and Uba4p-related operons.** HVO\_0558 and HVO\_0559 are related to Uba4p N-terminal E1-like domain and C-terminal RHD, respectively. SAMP1, SAMP2 and HVO\_0558 genes are represented as black arrows. Genes that are linked to SAMP1, SAMP2 or HVO\_0558 in *H. volcanii* and other diverse haloarchaea (e.g., *Halorubrum lacusprofundi*, *Haloarcula marismortui*, *Halobacterium salinarum*, *Natronomonas pharaonis*) represented as arrows shaded with grey diagonal lines. Genes not conserved in order among other haloarchaea are represented as white arrows. Abbreviations: TrkA-N, NAD<sup>+</sup>-binding domain of K<sup>+</sup> channels, phosphoesterases and other transporters; TrkA-C, domain often linked with TrkA-N; TMH, transmembrane helices predicted; AAA or triple-A ATPase, ATPases associated with a variety of cellular activities; GNAT, Gcn5-related N-acetyltransferase domain; P1 to P3 with arrowheads, represent promoters based on the archaeal transcription factor B recognition element (BRE) and TATA-box core consensus sequences (cRnaAn:tTTTAWAtr; BRE:TATA, where R = A or G, N = any nucleotide, W = A or T and lower case letters represent less conserved bases)<sup>7-11</sup>; \*, indicates low similarity to archaeal consensus promoter.

**Suppl. Figure 1a.**

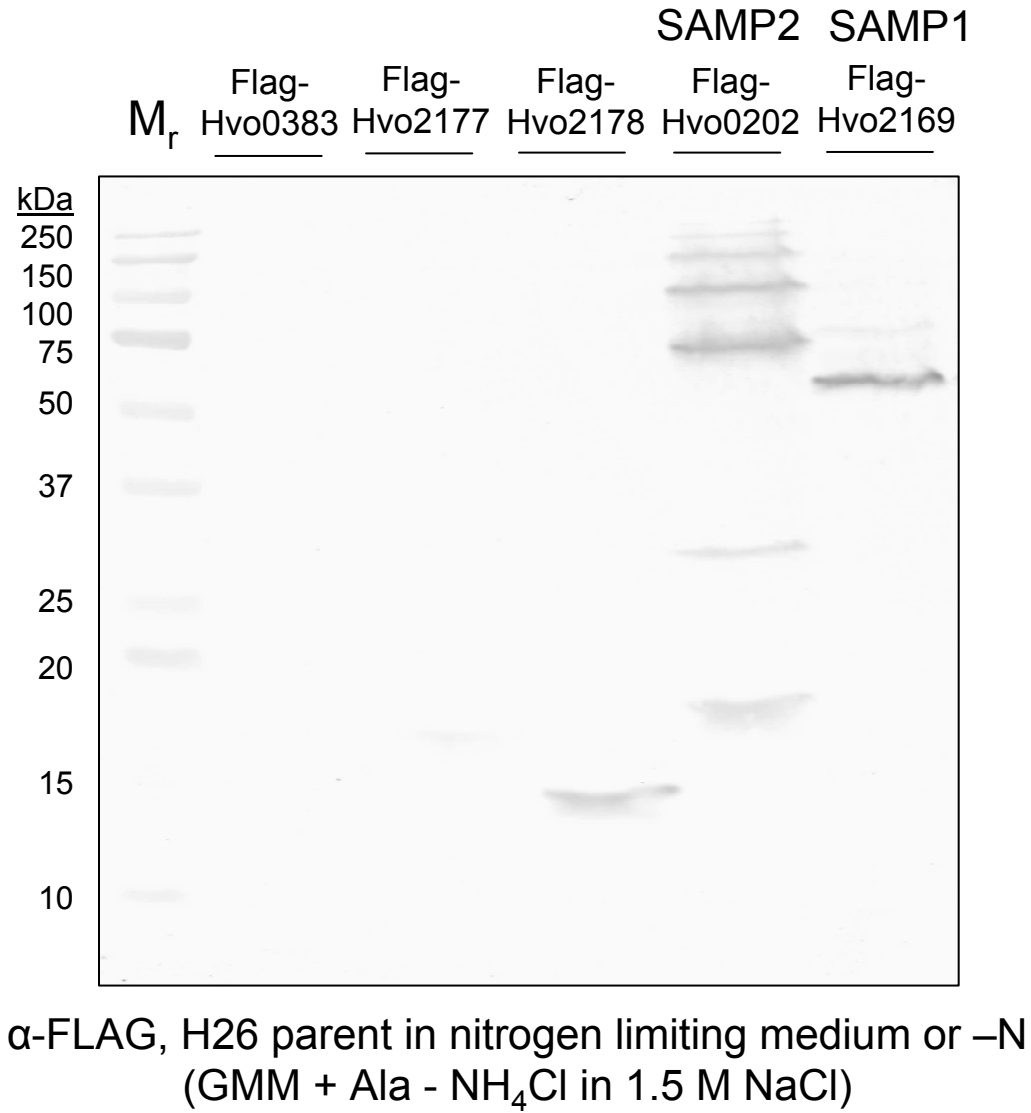


**Suppl. Figure 1b.**

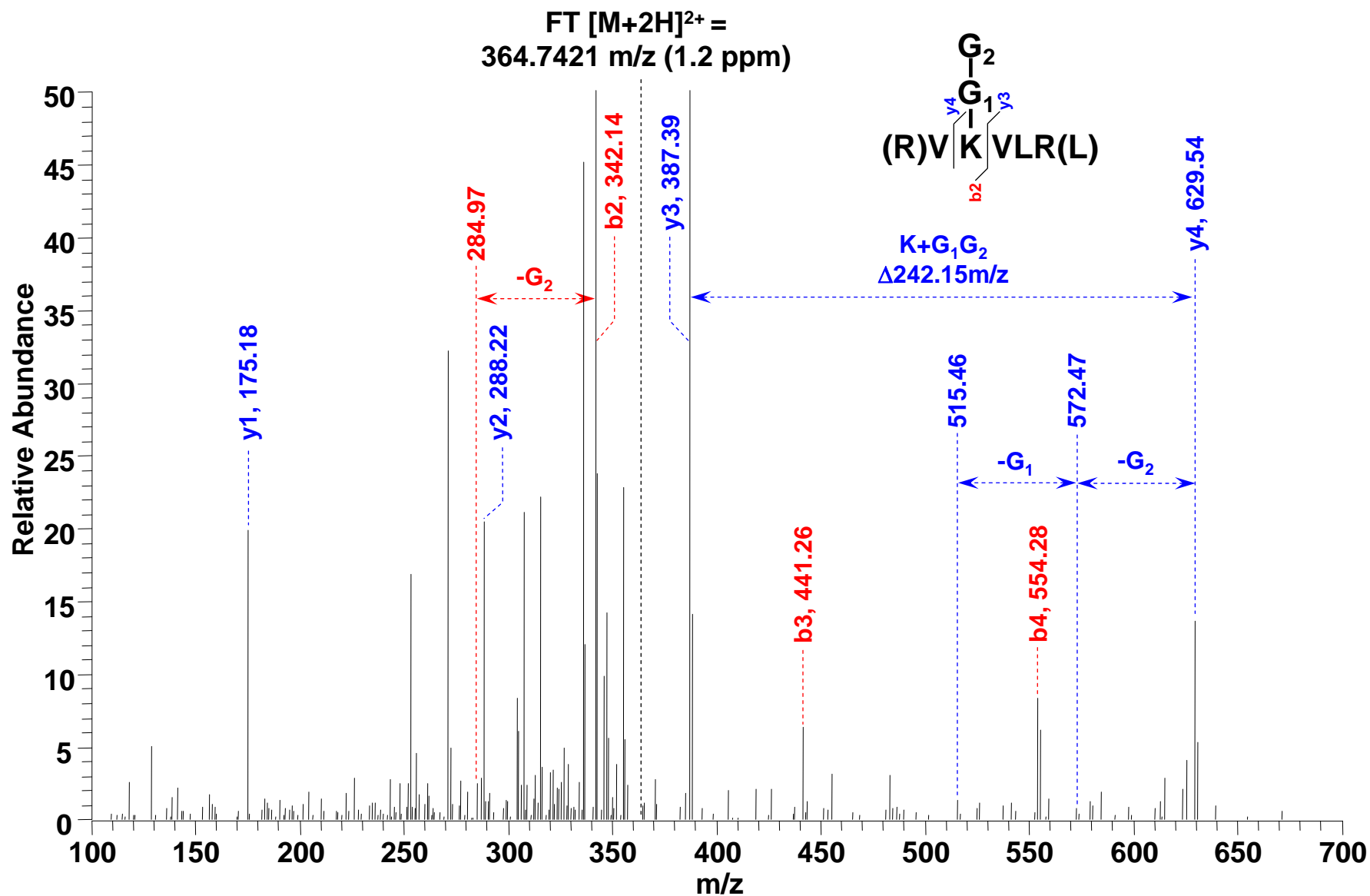


$\alpha$ -FLAG, H26 parent in nitrogen limiting medium or -N  
(GMM + Ala - NH<sub>4</sub>Cl in 2.5 M NaCl)

# Suppl. Figure 1c.

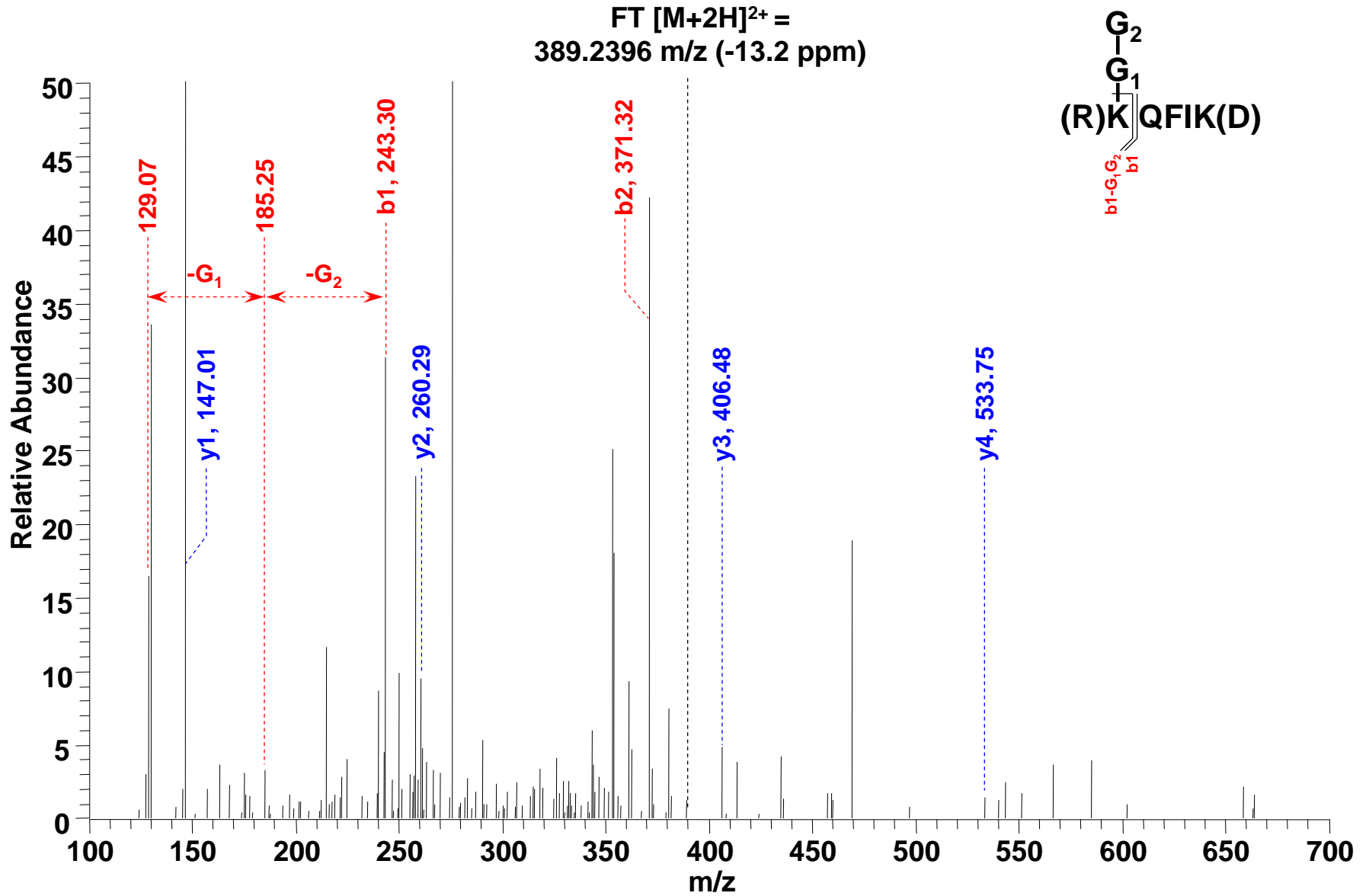


# Suppl. Figure 2a. HVO\_0202 Small Archaeal Modifier Protein 2 (SAMP2)



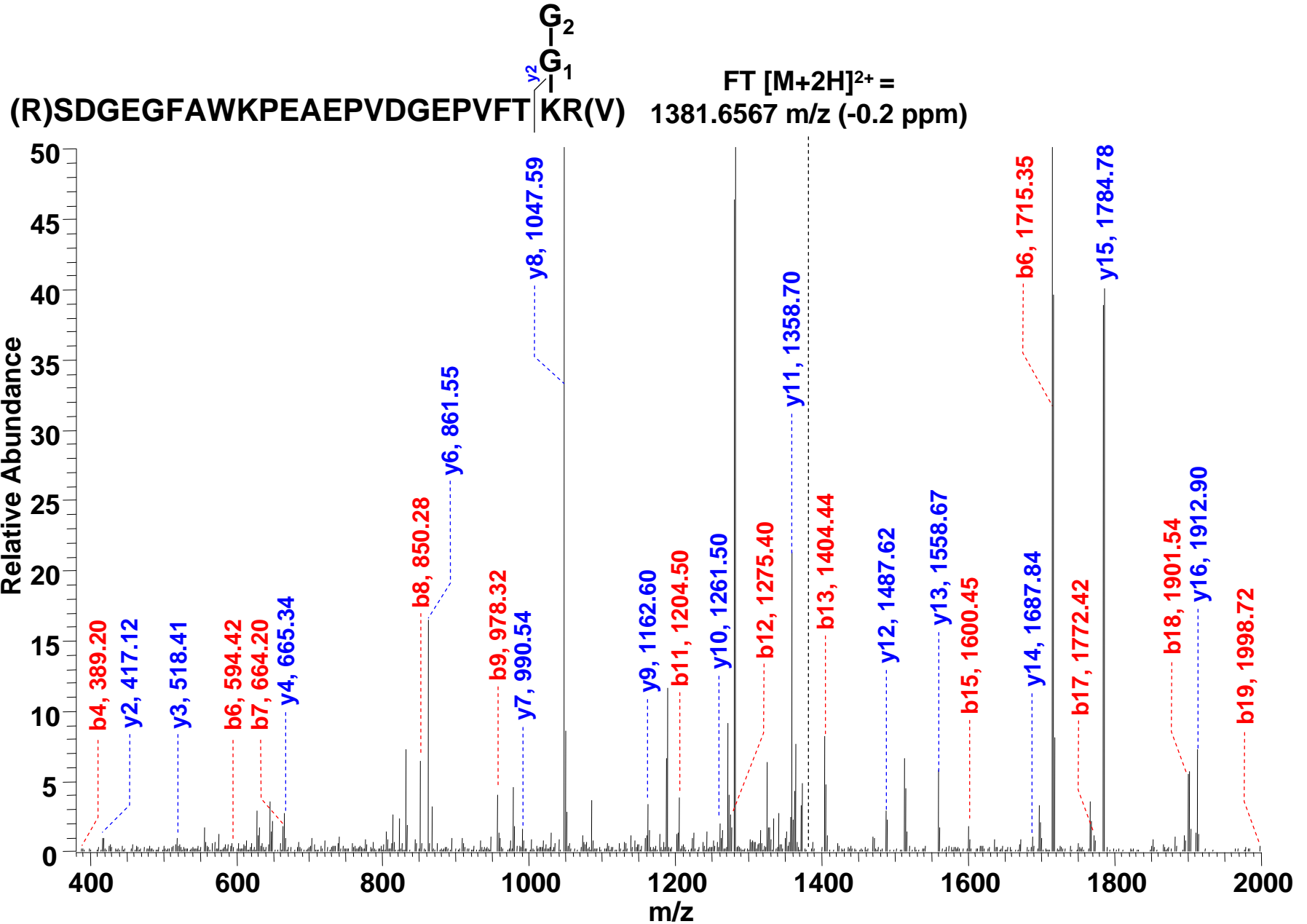
Suppl. Figure 2b.

HVO\_1572 GyrB, DNA gyrase B subunit

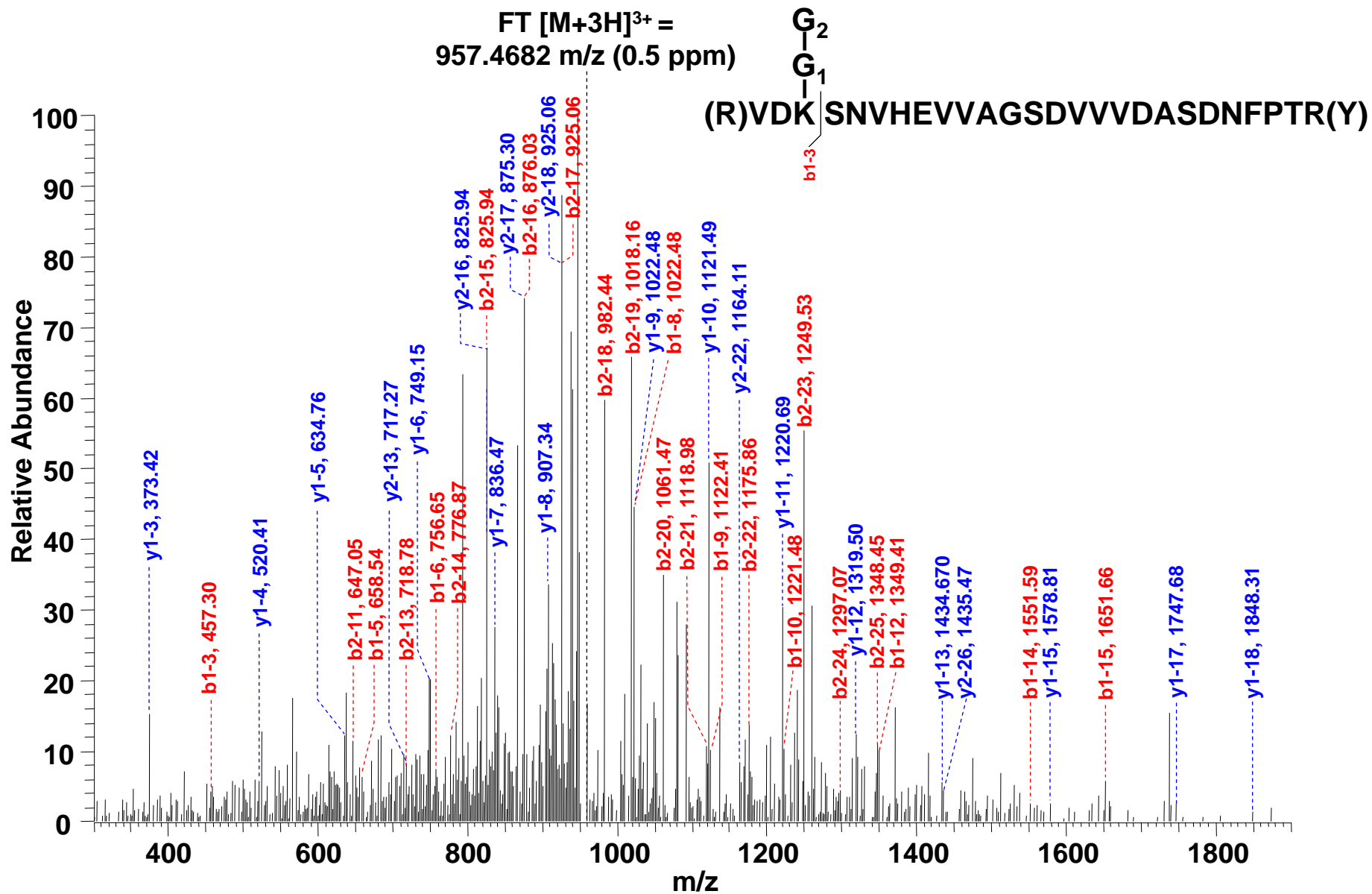


Suppl. Figure 2c.

HVO\_2328 isochorismatase

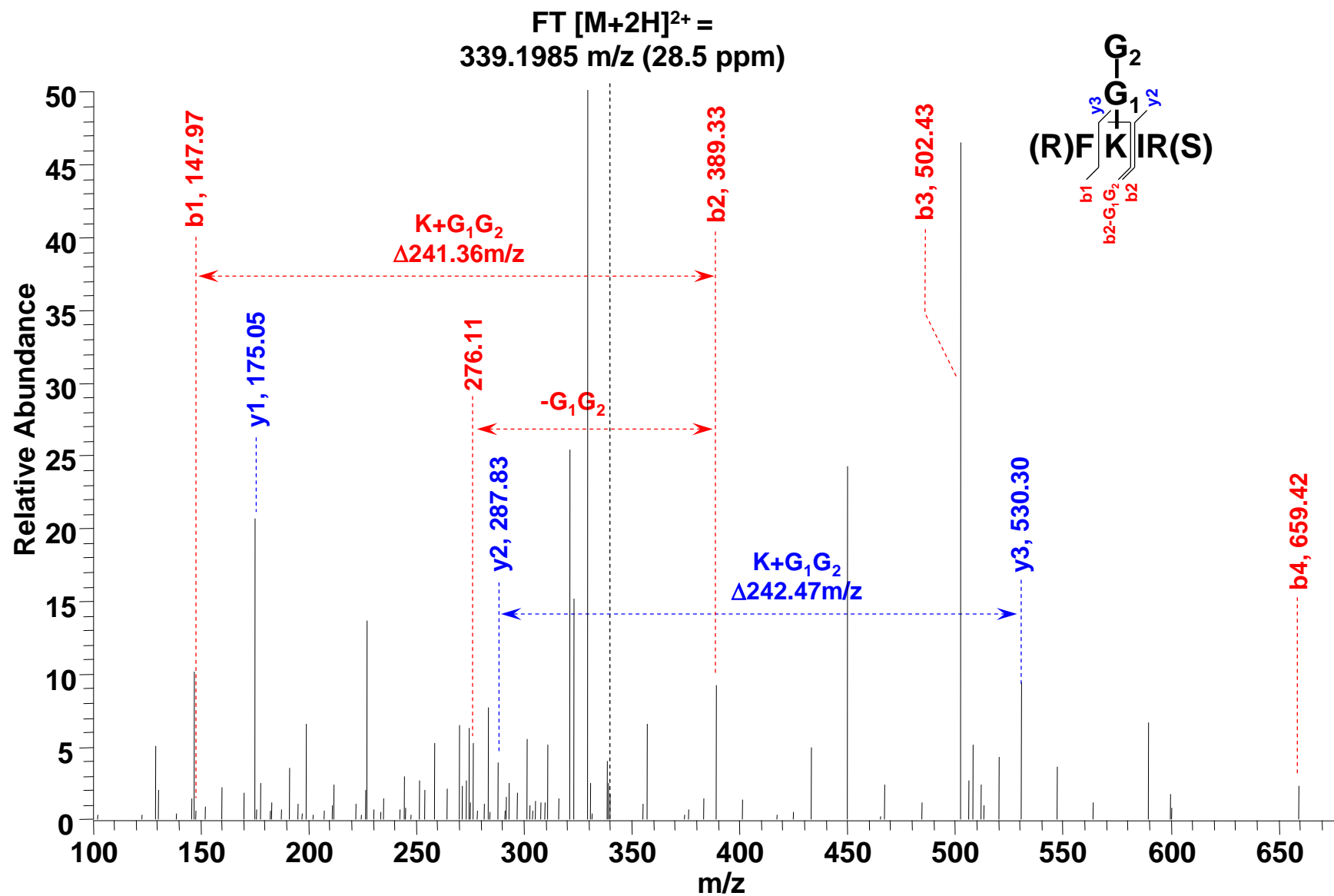


Suppl. Figure 2d. HVO\_0558 molybdopterin biosynthesis protein MoeB





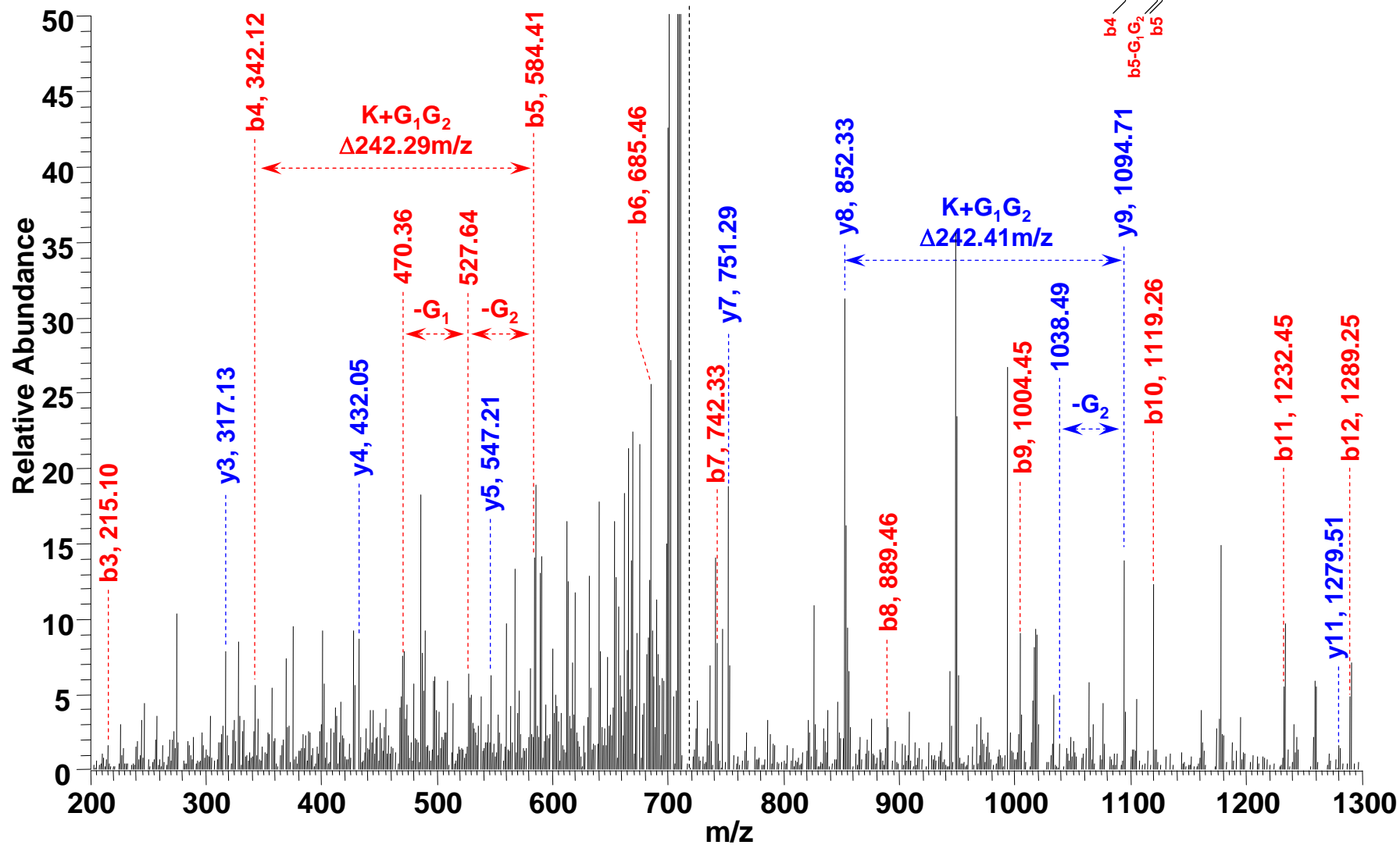
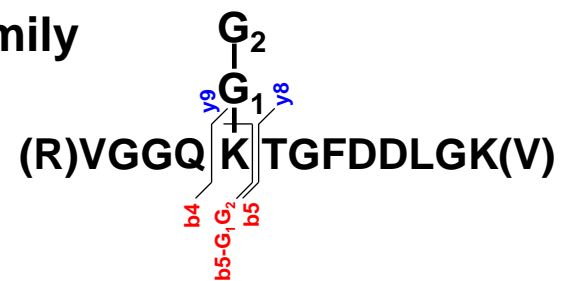
# Suppl. Figure 2e. HVO\_0980 NdhG NADH-quinone oxidoreductase chain c/d



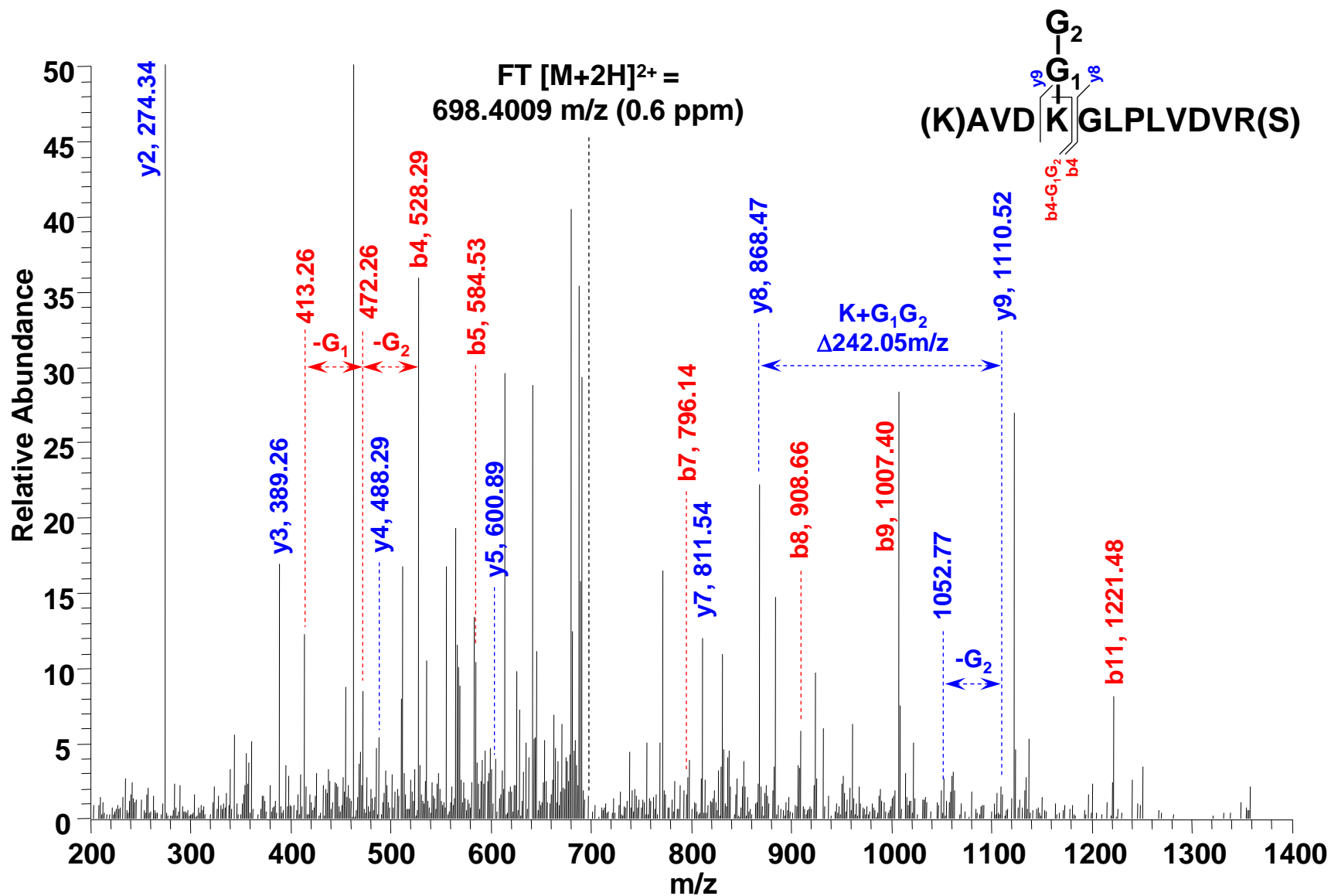
Suppl. Figure 2f.

### HVO\_1289 OsmC-like protein superfamily

FT  $[M+2H]^{2+} =$   
718.3598 m/z (3.9 ppm)



Suppl. Figure 2g. HVO\_0025 SseA/TssA thiosulfate sulfurtransferase

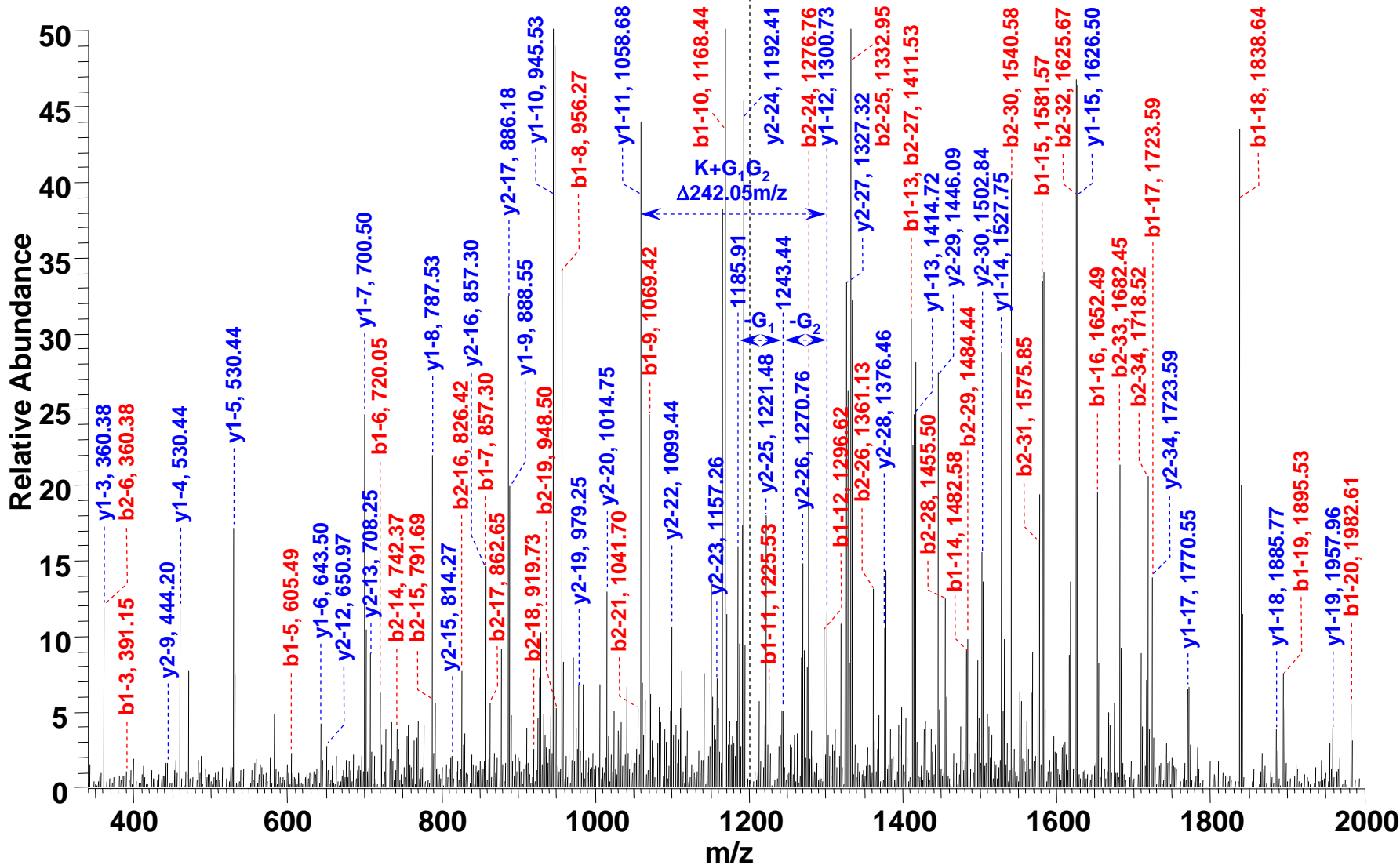


Supp. Figure 2h.

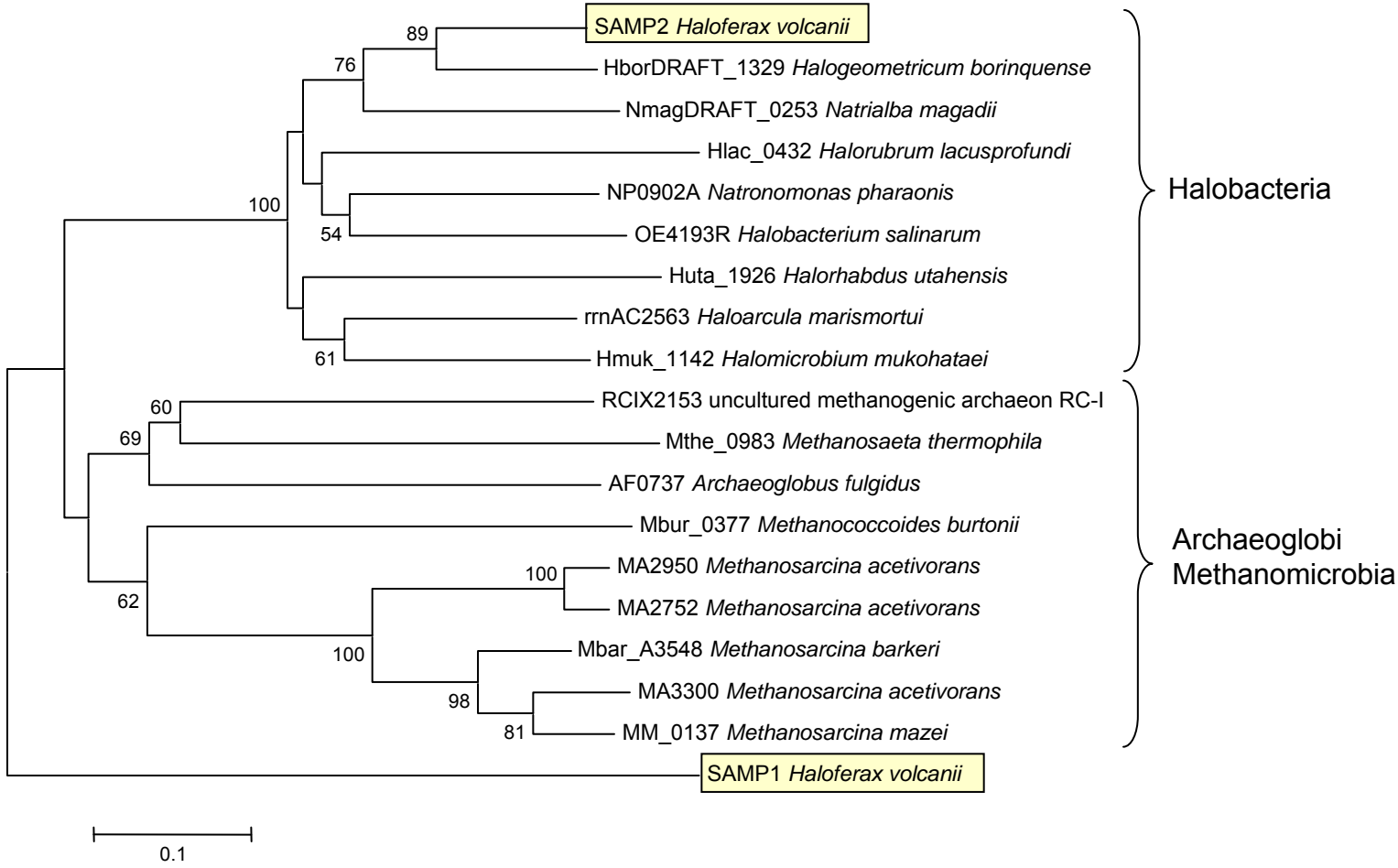
(R)YLNDVDHVLVGADAVAADGSVIN K IGTSGLAVNAR(E)

HVO\_0966 eif2ba / aLF-2BII  
translation initiation factor

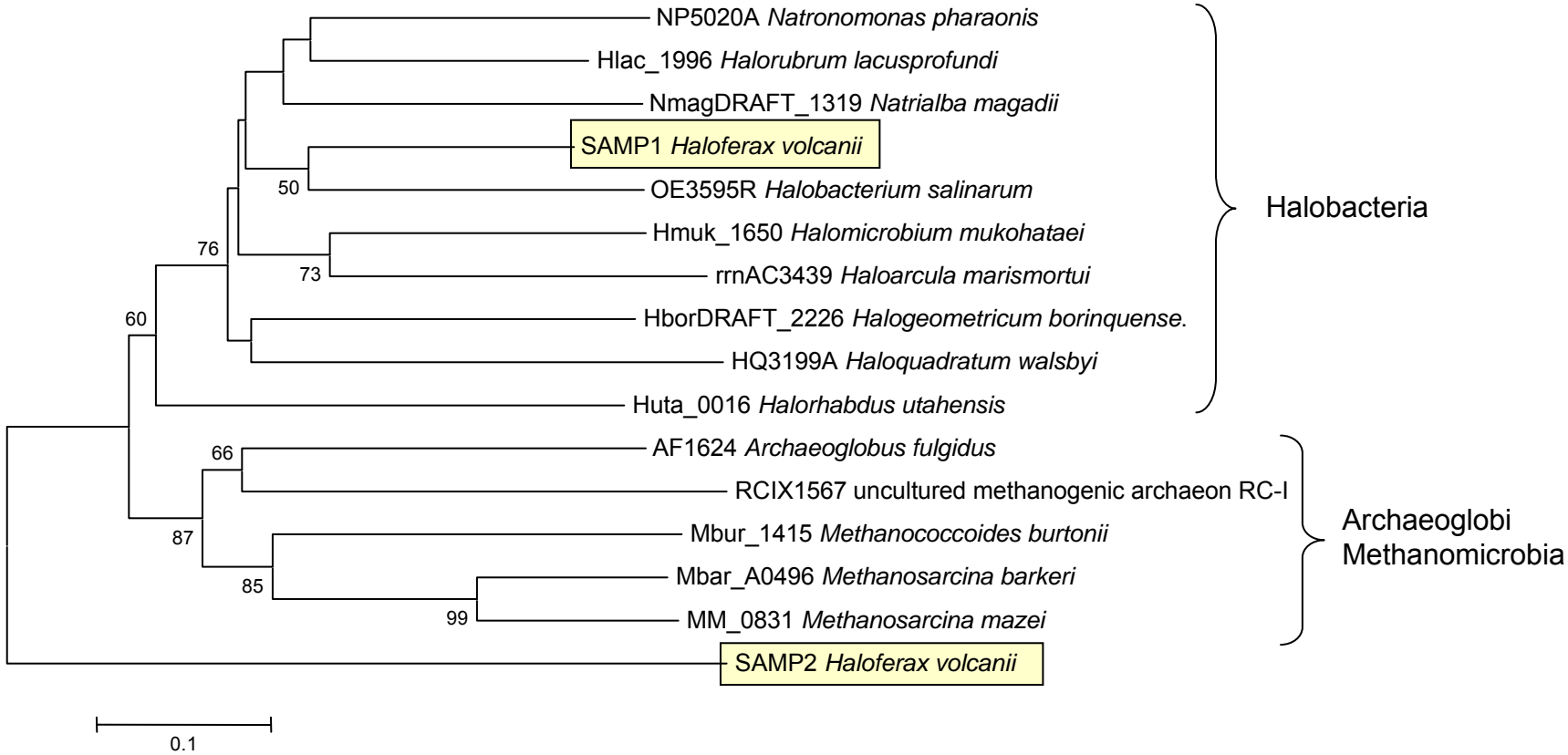
FT [M+3H]<sup>3+</sup> =  
1203.6311 m/z (-3.3 ppm)



Suppl. Figure 3a.



Suppl. Figure 3b.



# Suppl. Figure 4a.

```

HVO_SAMP1      1  -----MEWKLFDADLAEVAGSRTVVRVDV-DGDATVGDALDALVGAHPALESRVFGD-----DGELYDHINVLRNCE--- 64
NP5020A       1  -----MEWRLFANLAEAGTKRVEVDA-APGDTFGDAFEQLAAHPDLEAEVLDE-----DGE LRDHIRVLRNDRNPF 67
HborDRAFT_2226 1  -----MTTVRWRLFADLAEVAGERETAVTV-ADDATVEDADALLDNHDGLRTRVF-D-----DGV LADHVNLIRNGA--- 66
Hlac_1996     1  -----MEWKLFDADLAETIAGGREIDVAA-EPGDTVGDALDALLEAHPPDRDRVLD-----DGEVANHINVLRNQON-V 65
OE3595R      1  -----MEWQLFADLSEVAGDDTVTVAVPTDDPTVQDALDALVADYPALGDRVLAA-----DGS LAAHVNVLVNGVT-- 66
Hmuk_1650    1  -----MHWRLFATLAETAGEDRVSVDV-A--ETVGEALDALLSAHPALESEVLDD-----DGS MVDHIRLLHDGDDPF 65
NmagDRAFT_1319 1  -----MPTEWKLFDADLAEHAGDKHVTVDA-SAGDTVDDALSDLVDDRPAIADRVLTD-----DGD LRSQINVLRNGTINVL 69
rrnAC3439    1  -----MEWKLFAHLRDAADGQSVSVDI-EGDATVETALDALLATRPALAEVLEDE-----NEE LADHIRVLRVLDGEDPF 67
HQ3199A      1  MPTTEPNSESNEFNTEIEDDADRETISITWRLFADVAEITAGTREKSIDI-HTNATTEDAIDELLNDRPALANRVLLDP-GTKTETELATEINL LONGS--- 95
AF1624       1  -----MVRVKLFANFREAAAGVK-EVEVE--AG-TVGEVLQELVRRFPKLESLEYEE-----GRL RLDYVNVIMVNGRN-V 63
RCIX1567     1  -----MKIKVKLFANFREATKQKEVEISP--EGDSVKAVIKGLVTKYPNIEPLLFQD-----GAL KPYVNVLVNNGQT-V 66
Mbar_A0496   1  -----MAEVKVKLFANLREKAGTSELQLSG---EKVIDVLLAL TGKYPELKNLIFEEPEGESEAPVMCGSINVLINGNN-V 72
MM_0831      1  -----MAEVKVKLFANLREAAGTPELLLSG---EKVIDVLLSL TEKHPALKELIFEKNDGEDESPVLCASINVLINGNN-V 72
Mbur_1415    1  -----MYQTSWIVLLNMQYTIAMVGVKIKLFLANLREIAGVSELELEG---ENIQEILDILQNDHPQIQELIYDDDRGKKE---TRAYINIL INGN-I 87
Mlab_0189    1  -----MKITVKAFAATFREVM DMKVELEFP--AGATLKTLLSEL TARYMGLDALLFAAP-----GTL RDFVNI LKNGRN-I 67
TTC1835      1  -----MGRGFPLDSGGMPKVNLYATFRDLTKS--QLEL--PGATVGEVLENLVRAYPALKEELFEG-----EGL AERVSVFLEGRD-V 74
TaQDRAFT_3666 1  -----MPKVNLYATFRDLTKS--QVEV--RGRTVGEALKDLVARYPALGAELFEG-----EEL LERSVSVFLEGRD-V 63

HVO_SAMP1      64  --AAA LGEATAAGDELALFPPVSGG----- 87
NP5020A       68  VSDDCFE TLEEGDELALFPPVSGG----- 92
HborDRAFT_2226 66  --DTS LSEEVSDGDELALFPPVSGGCVPTPNPDRIERR- 102
Hlac_1996     66  RQNGGMDITLEAGDELALFPPVSGGAVGR----- 94
OE3595R      67  VDADDLDTAIAAGDELALFPPVSGG----- 91
Hmuk_1650    66  AAGDGLDTPVSPGDELALFPPVSGG----- 90
NmagDRAFT_1319 70  VEEDGLETELEEGDELALFPPVSGG----- 94
rrnAC3439    68  AAGEGLATAVDEETELALFPPVSGG----- 92
HQ3199A      95  --DANRDDQLTDGDELALFPPVITGG----- 118
AF1624       64  R--GD LNYPLSHIDEVALFPPVSGG----- 86
RCIX1567     67  KGE EGLTAKIKDGEHALFPPVSGG----- 91
Mbar_A0496   73  RHFDGLD TLLSDADELAVLPPVSGG----- 97
MM_0831      73  RHLEGLD TDLKSD EIGVLPVSGG----- 97
Mbur_1415    88  QHLEGS DTVLNEGDEI AIFPPVSGG----- 112
Mlab_0189    68  HFIAGLD TPLDEGMIALFPPAAGG----- 92
TTC1835      75  RYLQGLS TPLSPGATLDFPPVAGGG-FERTFGA FPPWLLBRYLEE WGTREGE GGVYRLPGAVVRFREVEPLK VGSLSIPOLLVEVEGEAGEAWFERIALAASRGGG 180
TaQDRAFT_3666 64  RYLEGLA TPLEEGATLDFPPVAGGSGSLGGEFGAL FPPWLF EAYLCSWGGEKEA EGVFRLPGARVRFAPPL E VGS LRIPRLRVEVEGEAAEAWGRIAL AASRGGG 170

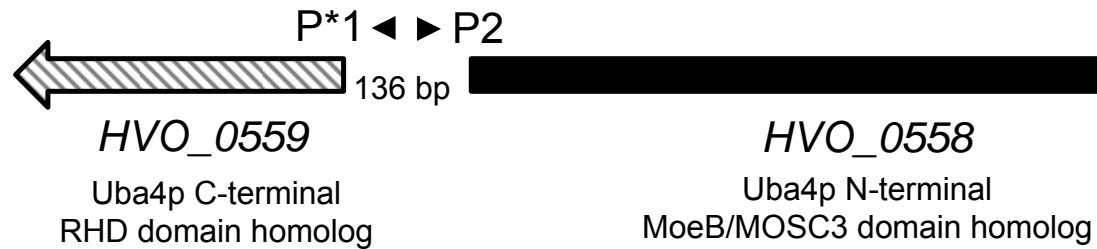
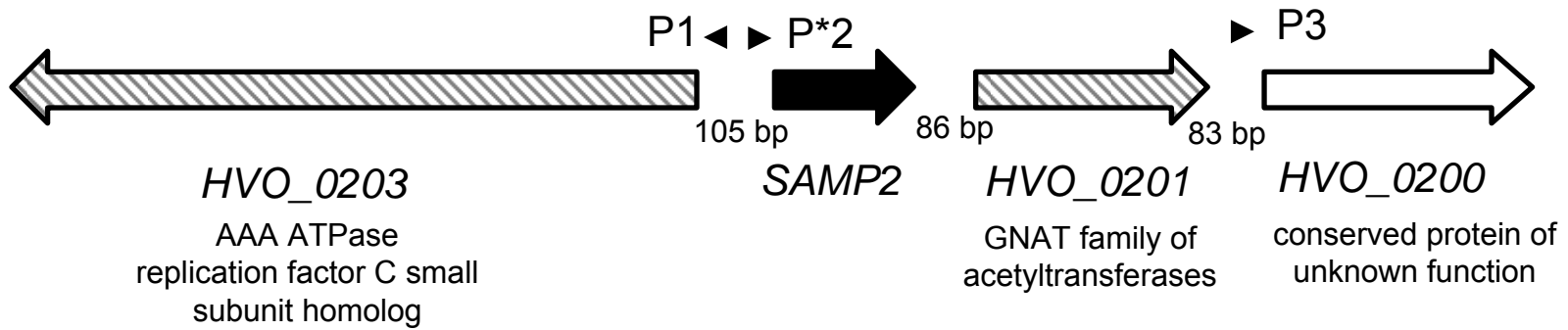
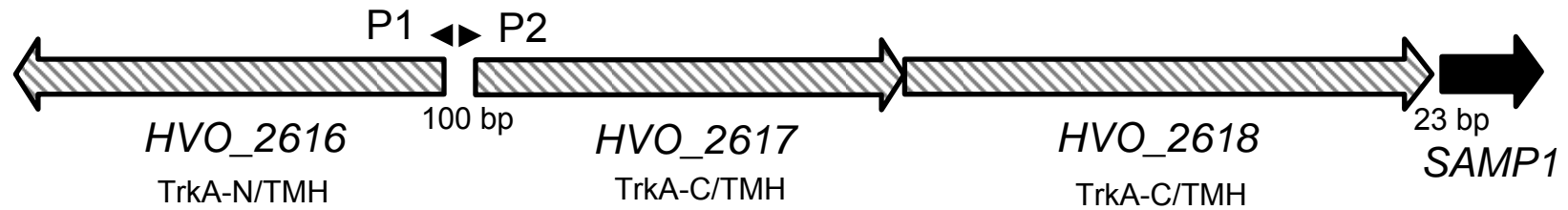
```

# Suppl. Figure 4b.

Hvo_SAMP2	1	---MNVIVVEVVGEE---ETSEVAVD-----DDGTYADLVRAVDLSPHEVTVLVDGRPVPEQSVVEVDRVKVLRRLIKGG-----	66
HborDRAFT_1329	1	---MHVSVVEVVGEE---ETHEIRVE-----DDDITYADLVRAVDLSPHEVSVLVDGSPVPEQPVAVERVKILRLIKGGAY----	68
rrnAC2563	1	---MHVIVVEIAGE---DTHELEVD-----ADATYGDLLAPVDLSPHEVSVLVDGEHVPIDQFVEHDHVRVVRRLIKGG-----	66
NmagDRAFT_0253	1	---MRVIVDVKGE---ETHEISLTDLREGNGGSSADTTADKADKPTYADLLREIDLSPHEVSVLIDGRPVPEQFVESEQVTVLRLIKGGSR----	87
Hmuk_1142	1	---MRVIVVEIVGG---ETREVEIA-----AGDSYADLLAPTDLSPHTVSLVDGTPVPEQFVETDHVRVVRRLVKGGE-----	66
NP0902A	1	---MRVITCDIVGE---DTRELDVDEAA-----TYCDLLEAVGLSSHEASVLDGSPVPEDRTVSDSVRILRLIKGG-----	66
Hlac_0432	1	---MHVIVVEVVGEE---GTAEYDLPDDA-----TYADLIREAGYHPQEASAVLDGSPVPGDRLVDAESVRLRLIKGGSTGGGA	72
Huta_1926	1	---MEVIVVELVDE---ITREFDLEPGT-----TYGELLADFEVSPHEVSVLVDDRVPVEEAVDAEHVRIVRLIKGG-----	66
OE4193R	1	---MRVITLSVVG---DTHEFTVTTAADTEPG-----PPTYADLLADVGLSPHEVAVMVDGSPVPEHFPVDADFVRVLRRLIQGG-----	74
Mbar_A3548	1	-MSRKVHTIQAG-----EISEQTVEVAE-----SATYEDLLNTLDINQETVLLVNGGNAVPLDGTVSSDRLTILRVVTGG-----	70
AF0737	1	---MKLKIKFVGF---DQKEVEVEVN-----GQRYSEIILESLGINPETVVVVKDNIPVPVDDVAEGGEVKKVVRVISGG-----	67
MA3300	1	-MSKRLYHTIQAG-----EISEQTVEVAE-----SATYEDLLDTLNLINQETVLLVNGGNAVPLDGAINDRLTILKVVVTGG-----	70
RCIX2153	1	MEVLHLTVILHAG-----RKTETQVDLPE-----GSSYFDLLRELKVNPEVTVVFKNGIPVAFDSVVEGGTVEVMRVVSGG-----	71
MM_0137	1	-MSKKVHTIQAG-----EIVEQTVEVAE-----TATYEDLLDELNLINQETVLLVNGGNAVPLDGAIGSDKLTILKVVVTGG-----	70
Mbur_0377	1	-MSTTIKVKIFPD-----TPDQQAINVDP-----GSTYEDLLNVMDINQELVILLNDGKAVPIDDVVIEGTVTILKAIISGG-----	70
MA2950	1	-MSKKVHVRIQAGALCGAMEEISEQTVEIGE-----NLTYENLLDILGINDGTVIIVLNEGNAVPLDGISSDKLTMRVVSGG-----	77
MA2752	1	-MSKKVHVRIQAGALCGAMEEISEQTVEIGE-----NFTYENLLDILGINEGTVIIVLNEGNAVPLDGISSDKLTMRVVSGG-----	77
Mthe_0983	1	--MISVSVRIFAG-----VESERQLVLP-----GTSYVEILLDLLGINPETVAITRDGVPVFNLDLVLPGEIEIIRVVSAG-----	69



# Suppl. Figure 5.



## References for Supplementary Information

1. Mullakhanbhai, M. S. & Larsen, H. *Halobacterium volcanii* spec. nov., a Dead Sea *Halobacterium* with a moderate salt requirement. *Arch. Microbiol.* **104**, 207-214 (1975).
2. Wendoloski, D., Ferrer, C. & Dyll-Smith, M. L. A new simvastatin (mevinolin)-resistance marker from *Haloarcula hispanica* and a new *Haloferax volcanii* strain cured of plasmid pHV2. *Microbiology* **147**, 959-964 (2001).
3. Allers, T., Ngo, H. P., Mevarech, M. & Lloyd, R. G. Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the *leuB* and *trpA* genes. *Appl. Environ. Microbiol.* **70**, 943-953 (2004).
4. Zhou, G. Y., Kowalczyk, D., Humbard, M. A., Rohatgi, S. & Maupin-Furlow, J. A. Proteasomal components required for cell growth and stress responses in the haloarchaeon *Haloferax volcanii*. *J. Bacteriol.* **190**, 8096-8105 (2008).
5. Jolley, K. A. *et al.* Dihydrolipoamide dehydrogenase from the halophilic archaeon *Haloferax volcanii*: homologous overexpression of the cloned gene. *J. Bacteriol.* **178**, 3044-3048 (1996).
6. Kaczowka, S. J. & Maupin-Furlow, J. A. Subunit topology of two 20S proteasomes from *Haloferax volcanii*. *J. Bacteriol.* **185**, 165-174 (2003).
7. Littlefield, O., Korkhin, Y. & Sigler, P. B. The structural basis for the oriented assembly of a TBP/TFB/promoter complex. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 13668-13673 (1999).
8. Qureshi, S. A. & Jackson, S. P. Sequence-specific DNA binding by the *S. shibatae* TFIIIB homolog, TFB, and its effect on promoter strength. *Mol. Cell.* **1**, 389-400 (1998).
9. Palmer, J. R. & Daniels, C. J. In vivo definition of an archaeal promoter. *J. Bacteriol.* **177**, 1844-1849 (1995).
10. Bell, S. D., Kosa, P. L., Sigler, P. B. & Jackson, S. P. Orientation of the transcription preinitiation complex in Archaea. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 13662-13667 (1999).
11. Kosa, P. F., Ghosh, G., DeDecker, B. S. & Sigler, P. B. The 2.1-Å crystal structure of an archaeal preinitiation complex: TATA-box-binding protein/transcription factor (II)B core/TATA-box. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 6042-6047 (1997).

***Haloferax volcanii* protein sequences described in this paper are listed below [see UCSC Archaeal Genome Browser (<http://archaea.ucsc.edu/>) or HaloLex (<http://www.halolex.mpg.de/public/>) for full genome sequence].**

>HVO\_2619-TG\_gha\_455 molybdopterin converting factor subunit 1  
MEWKLFDLAEVAGSRTVVRVDVDGDATVGDALDALVGAHPALESRVFGDDGELYDHINVLNRNGEAAALGEA  
TAAGDELALFPPVSGG\*

>HVO\_0202-TG\_gha\_455 conserved hypothetical protein  
MNVTVVEVVGEEETSEVAVDDDDGTYADLVRAVDLSPHEVTVLVDGRPVPEQDSVEVDRVKVLRRLIKGG\*

>HVO\_0558-TG\_gha\_455 moeB molybdopterin biosynthesis protein MoeB  
MTLSLSDATQLDRYSRHIIMDEVGPEQGRLSSRVVVVVGAGGLGAPAIQYLAAVGVGELV  
VVDDDDVVERSNLQRQVVHCDDDDVGTTPKAESAAAFVVRGLNPDVSVPEVEARVDKSNVHEVV  
AGSDVVVDASDNFPTRYLLNDVCRFEGIPLVHGAIYKFEGQATTLVPDGP CYRCLFPEAP  
EPGTVPDCAATTGVLGVLPGTVGCIQATEAMKLLLDGEALDGRLLFYDAMDMTFETVYR  
TNPDCPVCGEVDSIEDIDYVESCAISLD\*

>HVO\_1864-TG\_gha\_455 moaE molybdenum cofactor biosynthesis protein  
MHVLGIVGAGATTLCDRLAAQLDGRVATVESLPESATEPESTDGVAAAYGLSPDGNWVGS  
GDDRLDLGLLDDL SAEYDYALLSGFPDARVPTVALAGVDAANVVAEAETADAADVASLAA  
EIDSHEPHVTLET LVERAKADPLEVYAGAIATFTGRVRAKESEDDDP TLSLEFEKYDGVA  
ESKMDAISEELEARDGVLRLVMHHRVGVVEDGADIVFVVVLAGHRREAFRTVEDGIDRLK  
DEVPIFKKETT TDEEFVWHDR\*

>HVO\_2305-TG\_gha\_455 moeA molybdopterin biosynthesis protein moeA  
MRKQFRDLASPEDAREAIASLDLTPDAETVALDDARGVLAERIDAELDVPGFDRASMDG  
YAVRARDTFGADEADPVELDLGAVHAGNEPDVAVDPGTAVEISTGAVMPPGADAVVIVE  
RTEERDGLSVVYSSVAPGDSVMPAGTDTAAGARALGPGTRLTTPREIGLLSALGVDEVPVR  
GKPTVGVIVSTGDELVRPGDEL RPEAGEIYDVNSYTI AAGVEEAGGVPKLYPHAGDDYEEM  
ERLLRRAADECDLVLSSGSTSASAVDVIYRVVEERGELLLHGVSVKPGKPMVLVGTLGES  
AYIGLPGYPVSALTI FRTFVAPAVRDAAGLPEPTATVEGTMAARERYAEGRTRLM PVGL  
VADGDGETLVYPVDKGS GATTS LVEADGVVEVSADVEYLAEGESVTVHLFSPDVRAPALL  
GVGEDDPALSRVLD RVS SPRYLPFGSREGLRRLRDGISDVAVVAGDRGDAIDDI DAVELG  
SWTREWGLVVPAGNPDGVSGLADLVDDDLRFINRDSNSGLRTDLGNAVAALADERGTTRH  
ELVESIEGFD RGTKAFESPARAVAAGRADAGLGLRATADSLGLDFVSLGTQPVAVLANPA  
RIEKPGVDSL RDALAASDDAFADLSGFERGQTPR\*

>HVO\_0025-TG\_gha\_455 tssA thiosulfate sulfurtransferase  
MSNSDYAKDVLVSADWVESHLEDFQSDDPAYRLVEVDVDTEAYDESHAPGAIGFNWESQL  
QDQTTTRDVLTKEDFEDLLGSHGIS EDSTVVLYGDNSNWF AAYTYWQFKYYGHENVHLMNG  
GRDYWVDNDYPTTDEI PSFPEQDYSAKGPFEDIRAYRDDVEKAVDKGLPLVDVRSPEEFS  
GEILAPPGLQETAQRGGHIPGASNISWAATVNDGTFKSADELRLDYADQIGDESTIA  
YCRIGERS SIAWFALHELLGYENV TNYDGSWTEWGNLVGAPVEKGN\*

>HVO\_0861-TG\_gha\_455 sufB/sufD domain protein  
MSAQLPANLSAETVREISDARDEPEWLL EARLEALEALDALELPDVIQTPGRRWTNLEAL  
DFESLVDPLNQADETTRESPEGVVLPFTEALSEYGD LIEERFGSVVAPDNTNYLTALSAA  
LFTTTGTFIYVPEGVDAEDVKIRAEMNSRSLFSHTLVVTEESSSVTILEAIESGDSVDDAR  
YFENLVEIDAGENSIVQYGS LQNLDEETYTYSLKRADADTYSTVNWIEGNLGSRLRTRSDV  
ETELNGEASETKIVGAFGHDDQHLDVNARVWHKAENTTADLVTRGVLDDEARSVYEGVQ  
DVGRDAWSTNSYQRENTLMLS DQSEADASP KLI INNHDT EASHAASVGQVDAEDLFYMVS  
RSVPPEEQARNMLVEGFFVPLVEEIEVDEF RDDL EELIAARLR\*

>HVO\_0580-TG\_gha\_455 n-type ATP pyrophosphatase superfamily  
MECDKCGRDVAVMHAAYS GAHLCDHFCASVEKRVRRRIREDNMLPRDASPENPQTWVIGL  
SGGKDSVVLTHILDDTFGRDPRIELVALTIHEGIEGYRDKSVDACVELAEDLDIHHELVT  
YEDEFQVQMDVVEKDPENMAACAYCGVFRDLLERFADDELGADKLLTGHNLDDEAQTAL  
MNFEGDLKQVAKHFDASIGDFEKRRDAGEFIPRAKPLRDVPEKEVALYHLKDLPAHIT  
ECPHSSEAYRGEIQQLLLKLEENHPGTRHSIMAGYEELAELETAREYRGEGRVDLNDNCERC  
GSKTAGDVCRKCRLLIESIEAV\*

>HVO\_A0230-TG\_gha\_452 msrA methionine-S-sulfoxide reductase  
MGSTQTATFGGGCFWCVEAAFKELDGI SEVTSYAGGETENPSYEQVCSGSTGHAEEVVQV  
DYDPAVVG YDELLDVFFAVHDPTQLNRQGPVGTQYRSIVLFHDDEQKRQASAYIDALDE  
EYDDEVVTELVPLETFYEAEEHHQDYFEKVSTADGNTNDAYCQFNVPKIEKVRKFAFK  
VKAEAEFDA\*

>HVO\_2402-TG\_gha\_455 Glycine cleavage system P-protein  
MTAGNRRRVGSPYAPHTDADEEAMLDVAVGSDAEDLFDIPGEVRFDFGFGVDSADEQELR  
NAMADLLSRNEELTEFLGRGHHAHYVPALVDDLRRSEFLTSYTYQYQPEIAQGFLQVLF  
YQSMLEVELTGLPVANCSMYDAATALAEALLANRLRRGASGHRVLPPEHLHEGRLGVLEN  
YLDGAEMEIGEYPMDDGAVDVETLADLVDDVETVMVYAETPTVIRGVI EERLADISDLAHDN  
DALFCIGSDLVALLLEEPASVADVVVGEAGALGLGTAYGMGLGIFATREEYLRQVPGR  
LVGVSEDSADMRAFTLTLQTRERHIRKERATSNICTNQAWVALRTAMHMAWLGPDRLVDL  
ATQAVTDARDLAARLDEL SGVKAPLYDRHHFRFVVRTDQPAPAITNDLAGRGFAVHALD  
DHHLQVCITDANADAADGLVAAFEFEVI\*

>HVO\_2900-TG\_gha\_455 Fumarate hydratase class II  
MGDDDYRIERDSLGEVMPVDAYWGAQTQRAVENFPI SGITFGRRFVRALGVVKKAAAQA  
NRDLGMIDEGVAAAIVEAADEVIAGELDDQFPVDVFTGSGTSSNMNANEVIANRAAEIM  
GEGIGDRVHPNDHVNFGQSSNDVIPTAMHVASLEAVEKDLLPALETLAAALGDKEAEFD  
GVVKTGRTHLQDATPVRLGQEFGGYRTQVEKGI RRVENVREHLGELALGGTAVGTGLNTH  
PEFPAKAAEYMTTEETGVEFREADNHFEAQAAHDAMSEAHGALRTVAGSLNKIANDLRLLA  
SGPRNGLGELEQPENQPGSSIMP GKINPVVAEAVNQVHKQVVGNDAAVSAGAAEGQIDLN  
LYKPVLAHNFLS AKLLANSSEVFAERFVAKLEANEHCETRVEQSMALATALNPAIGYD  
KASKVAKKALAEKSVRQVAVDEGYL TEAEAEVLDPEKMTHRGILGDD\*

>HVO\_1289-TG\_gha\_455 OsmC-like protein superfamily  
MSDIQTSTVSEDGFACTS QVGFDFLQIDATDETGPNPNAALVATYASCFLPAFRVGGQKT  
GFDDLKGVQIDADADLDDSDDLERISFDVYVESDLSDEF AEITELAE DICHVHDALRDE  
LQADVTVVGDADF\*

>HVO\_1250-TG\_gha\_455 thiol-disulfide isomerase/thioredoxin  
MVLLES DSELERGDAAPDFELPGTDGETYSLSSFADYDAVLVVF TCNHCPYAKAKFEELN  
HLASAYDDLAVVGINPNAEDRPEDSFERMQELVADGTIGYTAYLRDESQAVAADYGAVC  
TPDPFLENTGDGFELAFHSRIDDAMSPDDEVS DYEMRTAVEALLAGDDIPVEEYPSQGC  
SIKWKDE\*

>HVO\_2682-TG\_gha\_455 dodecin-related protein  
MVFKKITLVGRSAESFDAAVDDAINRAEETLAGVHWVEVKELGVELASVETREYQAEVEV  
AFELEGEDGE\*

>HVO\_2583-TG\_gha\_455 hmgA 3-hydroxy-3-methylglutaryl Coenzyme A  
reductase  
MTDAASLADRVREGDLRLHELEAHADADTAAEARLLVSESQSGASLDAVGNYGFPAEAAE  
SAIENMVGSIQVPMGVAGPVSVDDGGSVAGEKYLPLATTEGALLASVNRGCSVINSAGGAT  
ARVLKSGMTRAPVFRVADVAEAEALVSWTRDNFAALKEAAEETTNGELLDVTPYVVGNS  
VYLRFRYDTKDAMGMNATIATEAVCGVVEAETAASLVALSGNLCSDKKPAAINAVEGRG  
RSVTADVRIPREVVEERLHTTPEAVAELNTRKNLVGSAKAASLGFNAHVANVVAAMFLAT

GQDEAQVVEGANAITTAEVQDGDLYVSVSIAASLEVGTVGGGKTLPTQSEGLDILGVSGGG  
DPAGSNADALAEICIAVGLAGELSLLSALASRHLSSAHAELGR\*

>HVO\_2328-TG\_gha\_455 isochorismatase  
MTAADLPDDAVLVCIDMQVGFDDPAWGDRNNPEMEARVADLLAAWRAADRPVVHVRHDSA  
EPDSPLRSDGEGFAWKPEAEPVDGEPVFTKRVNSGFIGTDLEAWLRERDHSSTLVVCGLT  
DHCVSTTTTRMAENLGFVYLPADATATFDREGHGGERFSADEMHRTALAHNLREFATVVE  
SADLSATR\*

>HVO\_1545-TG\_gha\_455 dihydroxyacetone kinase L subunit  
MADAETQREAVLDALDNVAERLAEEREYLTDLDSAIGDADHGANMERGFKAADKREEFV  
EMEPNEVVKVGTTLISNVGGASGPLYGGSIMFASQELEDGITAESSVAFEAAYLTKVKD  
RGAQVGAQTMVDALVPAVHTYKKSIEQDDLPLEALAKAVDAAERGVAFVPIKAMKGR  
ASFLDWRVSVGHQDPGATSTLFILEELLETAEEYLDGEVDRDARAEDA VGRGE\*

>HVO\_1496-TG\_gha\_455 ptsI phosphoenolpyruvate-protein  
phosphotransferase  
MTERTLSGIGVTPLSGVGTVVWYRPDADLPEPPAPVDVDAEAEELARFEDARAAA EDELEA  
ERERTAERVGEEEEAVFDAHVQFLNDPQITDGVSDAIESGLPAEHAVQETTFTEFVEQFEN  
MGGRMGERADDLRDVRLVRLSDGERVDLSSLPEGSVVVAERLTPSDTAQLDPERVAG  
FVTVTGGRTSHAAIFARSLALPAIVGVGEELQSVEDGTEVVVDGESGDLVVDPSDERKEA  
AAAAADVDIRHEAVETADGV DIEVAANIGTLADLGPVDRGADGVGLFRTEFLFLDRESP  
PDEDEQYEAIVEALESFDDGGRVVVRLTDIGGDKVPYLDLPDEENPFLGERGIRRS LGPD  
ADLFETQVRALLRAAASADGANLSVMLPLVSTVEELRAGRERFESVAADLDAEGVANELP  
EFGIMVETPAAAFMADQFAPHVDFFSIGTNDLAQYVMAAERGNERVSELGDYRQPAVLRA  
IDATVSAAEGEDCWWGMC GEMAGDPLTELLVGLGLDEL SMSAVTVPQVKAAVAETDTAD  
ARDLAERVLQADTKAEVAEILTL DQ\*

>HVO\_0481-TG\_gha\_455 gap glyceraldehyde-3-phosphate dehydrogenase type  
I  
MSEKSYLSAGENVDES DVVRVALNGFGRIGRNI FRAVL DNPVELVAINDVMDFDDMAYL  
AKYDSVMGRLDGVERDGDLS SIGGTSVSLYNVQSPAELPWGELDVDVALECTGIFRTKED  
ASAHL EAGADKVVISAPPKGDEPVKQIVYGVNHDEYDGEDIVSNASCTTNSVTPVAKVLD  
EEFGIENGLLTTVHAYTGSQNLIDSPHAKQRRGAAAENIVPTTTGAAQAATEILPQLDG  
KLDGMAIRVPVPNGSLTEL VVRL EDKPSVEEINDAFRAAADSGPLAGVLGYTDDEVT SRD  
ILGLPFSSTVDLNTTNQVNDGGLYKILT WYDNEYGFSNRMLDV AHFVTHQ\*

>HVO\_0887-TG\_gha\_455 porB 2-oxoglutarate ferredoxin oxidoreductase beta  
subunit  
MSSNVRFTDFKSDKQPTWCPGCGDFGTMNGIMKALANSGTDPDNTFMVAGIGCSGKIGTY  
MHSYAIHGVHGRSLPVAAGVKLANPDLTVVAAGGDGDGYSIGAGHF IHAVRRNVDMAYTV  
MDNRIYGLTKGQASPTSREDFETSTTPEGPQQPPVNPLALS LAAGGTFIAQSFATDHKRH  
AEIVQEAIEHGDGFGFVNVFSPCVTFNDVDTYDYFRDNLVLDADTDHDPTDYDAAKEKILD  
SSKEYEGIIYKDESSVS YE EKFGVDQDMSDIPSGAPDDAMD LVREFY\*

>HVO\_1000-TG\_gha\_455 acetyl-CoA synthetase  
MGELSEL FAPNRIAVVGATEREGAIGRAIMDNLID EFDGEVVPVNP KYDEL FGLQCYGDV  
GETDADLAVIVVPPKVVLPAMKSAGEAGIQNVVITAGFGETGSEGAAREQELRDIAEEY  
DLNVVGPNSL GIMNTDVGMNATFGPDMALDGNMSFMSQSGAFITAVIDWAND EDIGFKDI  
VSLGNKAVLDEADFIETWNDDPDTEVIIGYLEGISAGREFIDSARDVTKDTPIVLVKSGK  
TDAGAQAASSHTGTIAGSDAA YEAGLEQAGVIRADSVQH LFD TARVLGDQPLPENKDVAV  
ITNAGGPGVMTTDAIGESELQMA DFTDETLEAFSESLPAEGNIYNPVDIVGDADNARFKE  
ALDVALADDNVGMALVLTCP TAVLDYNQLAADTVELQEEYDKPIAACFMGGERVDAASDV  
MKDAGIPNYFDPSRAVDGLEALS KYADIRQREYDAPTEFDVDRERAREI LETVKDRDETR  
LGVEAMELLDAYGIETPAGDIVDDPADALEVAEDIDGNVVMKIVSPDILHKS DIGGVKVG  
VENEDVYDAYEDLITRKNYQPDANILGVQVQEMVNLDDGVETIVGMNRDPQFGPLMMFG  
LGGIFVEILED TTFRVAPVSETEAEDMTKEIDAAPMLRGARGNDPVDIGGITETIQRLSQ

LVTDFPAILELDINPLVALPDGVKAVDVRRLTVDPDEL\*

>HVO\_A0379-TG\_gha\_452 agaF N-methylhydantoinase A  
MFRFERSRERIVNIGANLITDCTKRGIVSQDNTQRVAVDIGGTFVDAITFDRETRDIALE  
KAATTPNQPSEGVIESVNGVDADLESANAFVHGTTLGLNAVLERDGTGIIITNEGFADV  
YEIGRTNLERTAMYDINYEKPEIVPRRRRVGVPGRDLADGAVVEEIDTDAVAEAAEYLV  
EKQDVDSIAICFLHAYQNGQHEQAAAECVQDAYPDISVSVSSDISGEYREYERTSTAVLD  
GYIKPIFENYVDTLDASLRDGGFDGSSFFVTRSGGGTLTAESAKSAPAHTILSGPAGGLIG  
ASHVGRVTDNRDLITVDMGGTSLDAAVVEDGSPVVKYDSSLEHQPLMIPVYDIRTIGAGG  
GSIWIDGDLKLVGPESAGADPGPICYDNGGTQPTVTDAALALGFLDPGDFLGGDMDTAA  
GDALDGIIEELAEPLAGMTVDEASKGVLVALANTVGAIREITVEKGLDPRDFSMVAYGGA  
GPMFVPLLARELGASEVLVQAPSVFSAWGMLMADVVDVDFSQTHLAVLDDATLNELKTAF  
ADLEAEGRETLTAEGVAENRQRIGRAVEMRYFGQEHTVEVDADGVSSLDELAERFEDQHE  
TRYGHTMDDPVQVVHLRVRVAVGENDKPELERGTREDESELTTPADTREAYCFAEDDFVEFD  
VYRRDDLKPGDEIRGPAVVTEPTTSLVFHSDQTATTTDDYGHIIITTDQ\*

>HVO\_0980-TG\_gha\_455 ndhG nadh-quinone oxidoreductase chain c/d  
MSLEEQQSDDPAELESVSRGDELAELLGDLVVGREEHLNAPGLVIRPDEVQDALFKLRD  
EAGFDHLSCVTAQEYEDRYESIYHLTKFDDRTDEVSVVVPTSKDNPVSQSAEPVFRATADW  
HEREAYDLVGIQYEDHPDLRRIILLPETWQGHPLGLDYDQDRPQIATLREHANPLEEDHRA  
GDSNTMYINIGPHHPATHGVLHVETVVDGEQVVDLESDIGYLHRCEEQMCQQGTYRHQIM  
PYPDRWDYISSGLLNEWAYARAAEDLADIEVPEYAQIIRTMGAELCRIASHMIALGTFAL  
DVYGDFTAIFMYAMRDREIVQNILEDLTGQRMFNYFRLGGVVDLPEPREEFFEKIRDF  
MDGLPQALEEYHDMITSNEILQARTVGTGVLSPVAKSYGATGPPVARGSGIDYDLRRDDS  
YGYDELEWDVVVEDGCDNFSRLVLRMREVEESAKIIQQCVDLLEDWPEDEINIQANVPR  
TLKPEDTEIYRAVEGAKGELGIYIRADGTDKPARFKIRSPCFNSLQTLPEMSEGEYIPD  
MIASLGSLLDIVLGEVDR\*

>HVO\_1727-TG\_gha\_455 tata-box binding protein e  
MSGPADSIEIQNVVASTGIGQELDLEALADDLPGADFNPNDFPGLVYRTQDPKAAALIFR  
SGKIVCTGAKSIDDVHDALGIIFDKLRELKIPVDDEPEITVQNIIVSSADLGHNLNLNALA  
IGLGLDVEYEPEQFPGLVYRMDEPKVVILLFGSGKIVITGGKRTDDAETAEEIVERID  
ALGLLG\*

>HVO\_1478-TG\_gha\_455 tfb transcription initiation factor  
MERPSRQRQREEEATAQEDEQVNCPEGSDQIVTDADQGELVCDDCGLVLDERQIDRGP  
WRAFNHSERQSKSRVGPITETMHRGLTTTTIDWKDKDAYGRSLSSSEKRSQMHRLRQWQ  
RIRTKDAGERNLQFALSEIDRMASALGVPRSVREVASVIYRRALNEDLIRGRSIEGVATS  
ALYAACRQEGIPRSLDEVAEVSVPQKEIGRTYRYSQELGLELKPVDPKQFVPRFASAL  
DLSEEVQAKATEIIDVSAEQGLLGGKSPGFAAAAIYAASLLCNEKKTQREVADVAQVTE  
VTIRNRYQEIQIEAMGFR\*

>HVO\_0359-TG\_gha\_455 tuf translation elongation factor EF-1 alpha  
MSDKPHQNLAIIGHVDHGKSTLVGRLLFETGVSPEHVIEQHREEAEEKGKGGFEFAYVMD  
NLAEERERGVTTIDIAHQEFDTDEFYFTIVDCPGHRDFVKNMITGASQADNAVLVVAADDG  
VAPQTRHVFLARTLIGELIIAVNKMDVVYDSEDKYKDVKEQVNKLLKQVRFNSDDATY  
VPI SAFEGDNIAERSDNTSWYDGDILLEALNNLPAPQPPTDAPLRLPIQDVYTIISGIGTV  
PVGRIETGTLNPGDNVDFQPSDVGGEVKTVMEMHHEEVDQAGPGDNVGFNVRGVGKDDIRR  
GDVCGPADDPPKVAETFKAQVVVMQHPSVITAGYTPVFHAHTAQVACTIESIDQKLDPAS  
GEVAEENPDFIKSGDAAIVTVRPQKPLSIEPSSEIPELGSFAVRDMGQTIAAGKVLEVNE  
R\*

>HVO\_0966-TG\_gha\_455 eif2ba translation initiation factor aIF-2BII  
translation initiation factor  
MDDRVHPEVRRATEIDTMEIRGAATIADAAARALRTQATESDAADAEAFRAELRATART  
LHETRPTAVSLPNALRYVLRDMSSTTVEGLRQSVVDSADFCARLERAQADLGQVGANRL  
RDGDTIMTHCHSTDALACVEAAVEQGKHIEAVVKETRPRNQGHITAKRLHELGPVTLIV

DSAARRYLNVDVHVLVGADAVAADGSVINKIGTSGLAVNARERGTPIMVAAQTLKLHPGT  
MTGHTVDIEMRDTAEVDDDTLADLGNPTVKNPAFDVTPPRYVDAIVTERGQFPPEI  
LMRELFEGGTSEPWAEPSPRAEP\*

>HVO\_1921-TG\_gha\_455 serS seryl-tRNA synthetase  
MIDRQLLRDEPERVRDALAARNMEDVDIDRVLDVYDEWRSLKAEGDDL RHERNEVSQQIG  
QLKQEGKDDEEAQEAIDRSGELKEELQELETRADELEAELDEALMELPNLPHE  
SVPVGADE DDNEEVRRVGFDDLREL PDEVTPHYDLGEELDI IDEGRAAKTTGSGFYFLK  
GEGAMLEHA LVQFMLDVHREQEYVDFPPIPVKTTSMEGTGQLPKFAEDAYRIGGAETENY  
DDDDLWLC PTAEPVPTNMYRDEILLKDDLPLKHQAYTPNFRREAGEHGTETRGIVRVH  
QFNKVELVNF VEPDESYDRLEALVDEAAEVLDRGLPYRVLSLCTGDLTFASAKTYDLE  
VWAPGTESEDA PEQGGRWLEVSSASNFEFQARRAGLRYRPERHESA EYLHTLNASGTAV  
GRVMVALLEY QNEDGTVDVPEPLQPYMGGREVIEGHEPVGEAAVGAGKKD\*

>HVO\_0677-TG\_gha\_455 aspS aspartyl-tRNA synthetase  
MRNRTYTADAEPGDTVTVAGWVHEVRDLGGIAFLILRDTSGKIQVKFEKDEMDDDLVETG  
LGVHRESVISVTGEVDEEPRAPTGEVTPESLDVIAEAEAEQLPLDPSGKVDAELSTRLDN  
RTL DLRKDEVKAI FEIRA EVQRAVRDKFRDLRATEINTPKIVATGTEGGTELPITYFGQ  
EAFMNQSPQLFKQLMVGSGLERVFEVGI FRAEEHNTPRHLNEATSIDFESAFIDHTEAM  
DVCEAVVTAAYEAVEENCQDELEALGLEEEFEAPSGEFPRLTYEEA IERINATGELDEQL  
VWGDDL PTEGEKALGEDVGEHYFITDWPSEIKPFYIKDHDDDETLSTGFDMHPNMELVS  
GGQREHRFDHLVAGFEQQGLDPDAFEYTKMFKYGMPPHAGFGLGGERLIMTMLGLENIR  
EAVLFPDRDRQLSP\*

>HVO\_1572-TG\_gha\_455 gyrB DNA gyrase B subunit  
MSQDNEYGAGQIQVLEGLEAVRKR PAMYIGSTDSRGLHHLVYEVVDNSIDEALAGHCDAI  
EVALHEDGSVSVTDNNGRGI PVDTHEQYDRPALEVIMTVLHAGGKFDNKSYQVSGGLHGVG  
VSVVNALSSLEVEVKRDGAVWTHRFEVGE PQVEEFERVRDLEPGEDTGT TIRFWPDDGI  
FETTEFDFKTLENRLRELAFLNSGVEISLS DERTDESSTFLFEGGIREFVEYLNETKTAL  
HDDVIYYDDESEGIEVEIAMQATDELQGS IHAFANNINTREGGTHLTGFKTALTRVVNDY  
ANTHDMLDDLDGDNLRGEDVREGLTAVIS IKHPDPQFEGQTKTKLGNSEVRGIVESVTHQ  
QLGTFEENPDTATAIISKAVEAARARKAAKQAEELTRRKSAL ESTSLPGKLADCQSRDP  
AESELFIVEGDSAGGSAKQGRDRKFQAILPLK GKILNVEKHRLDRILENDEIRALITAIG  
GGVGDEFDIEKARYQRLILMTDADVDGAHIRTLL LLLYRHRMRPLIEAGYVYAAQPPLYR  
VRYRGNTYDAMDEAERDRIIEEECN GNPTQVQRFKGLGEMNPDQLWDTT MNPENRVLKRI  
TVEDAAAADRMFNILMGDAVGRKQFIKDHANDA EWVDI\*

>HVO\_1344-TG\_gha\_455 conserved hypothetical protein TIGR00291  
MISLDEAVTARLESHGARFEVLIDPDAALS IKRGEFDGDLEEVIAAEDVFEDASRGDRPA  
ENDLEKVFGTDDPLQI IPEVVKKGEIQITAEQRREMQE QKRKSLINRIARNAVNPQMNS  
PHPPERIERALEEAGFKIDPMEPVESQVDDALDALRPVLP I KFAEVTAVAVQLPAEYAGSG  
QAQIRSYGDLERE EWQNDG SWGVITFPAGMQNDFYDKVNNITSGTAETRIVKDEDEL\*

>HVO\_1577-TG\_gha\_455 imd inosine-5''-monophosphate dehydrogenase  
MSSIELTSSQKTILTALINLYR DSEDAVKGEDIAAEVNRNPGTIRNQMQSLKALQLVEGV  
PGPKGGYKPTANAYEALDVDKMDEPAFVPLFHNDEEVEGVNVD EIDLSSVHHPELCRAEI  
HVQGSVREFHEGDKIRVGPTPLSKLVIDGTLDGKDDTSN ILILRIDDMQAPVGEPQH\*

>HVO\_0736-TG\_gha\_455 Domain of unknown function DUF302 superfamily  
MALPIDPSAIKPEDIGEERVVLEMEHEAAIERVREAF TDAGFGVATEFSPSEMLNEKVDA  
GRDPYYVLGACNPNMADRALDATDKKMGLFPCNVVIWEEEPGKQVVYHLSIMRVARLIG  
IAPDDDEMADIIADTGELVEQALANLDAADA\*

>HVO\_B0053-TG\_gha\_453 hypothetical protein (TBD)  
MACAELEALRLALLNITGTTDEHAKRHA EAELEDYLGADDPGPIQALANATTLDEAQRHL  
DAALVDLESEATRIDD DDPQAGYLRGLVAVRDAERSLRR LREGTDALLDDLGEAHHTLH  
DAFPVED\*