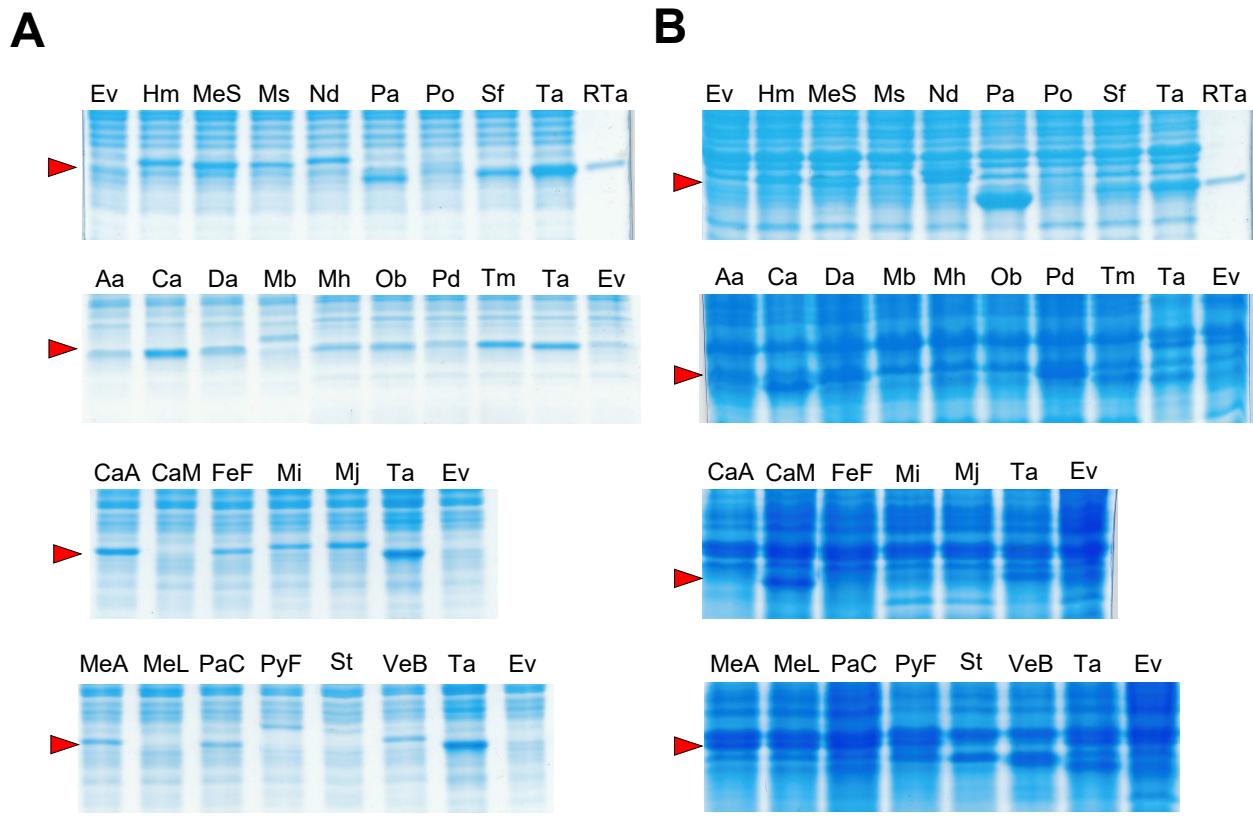


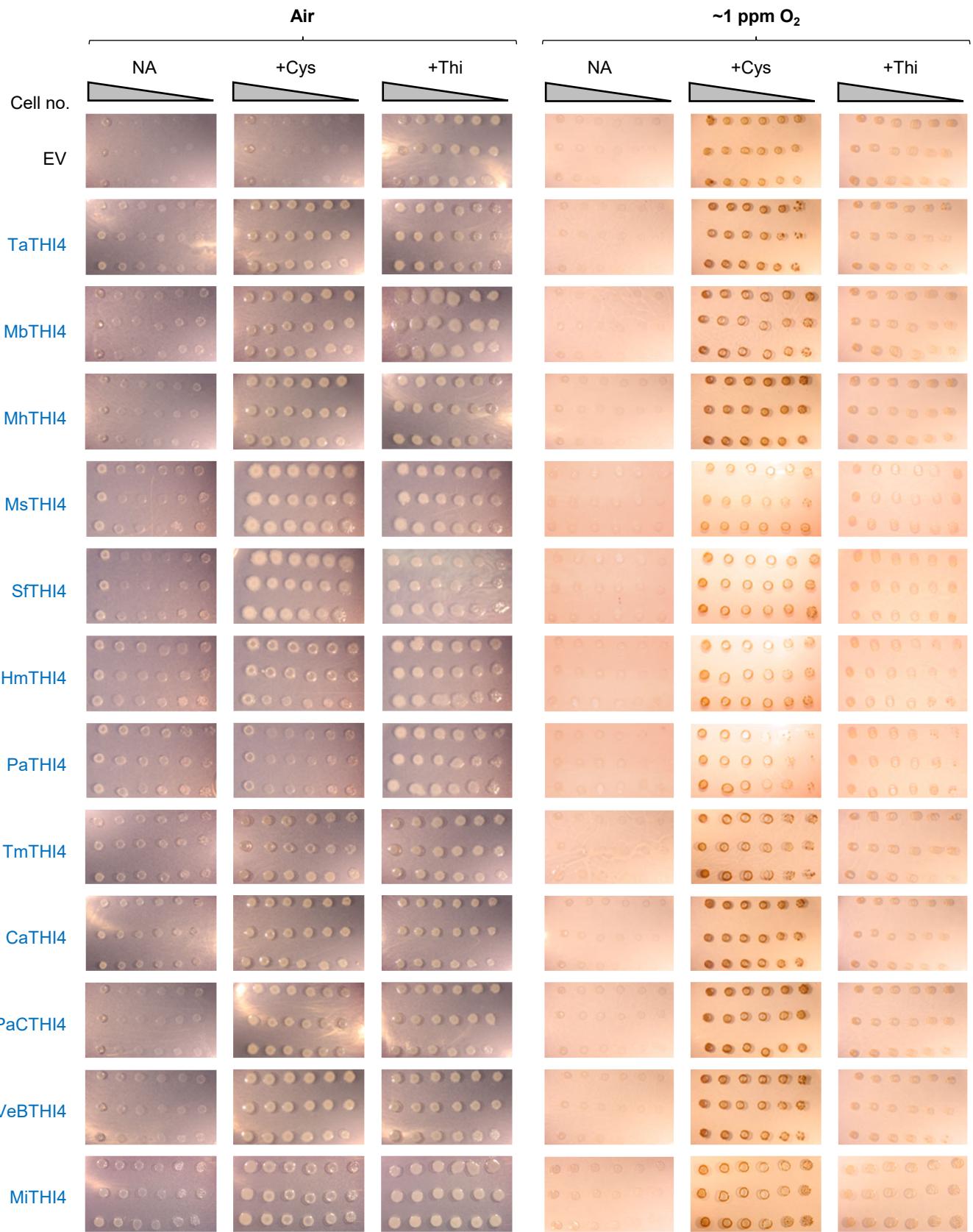
**Supplementary Figure 1. Ecology and genomic context of the 26 THI4s selected for testing.**

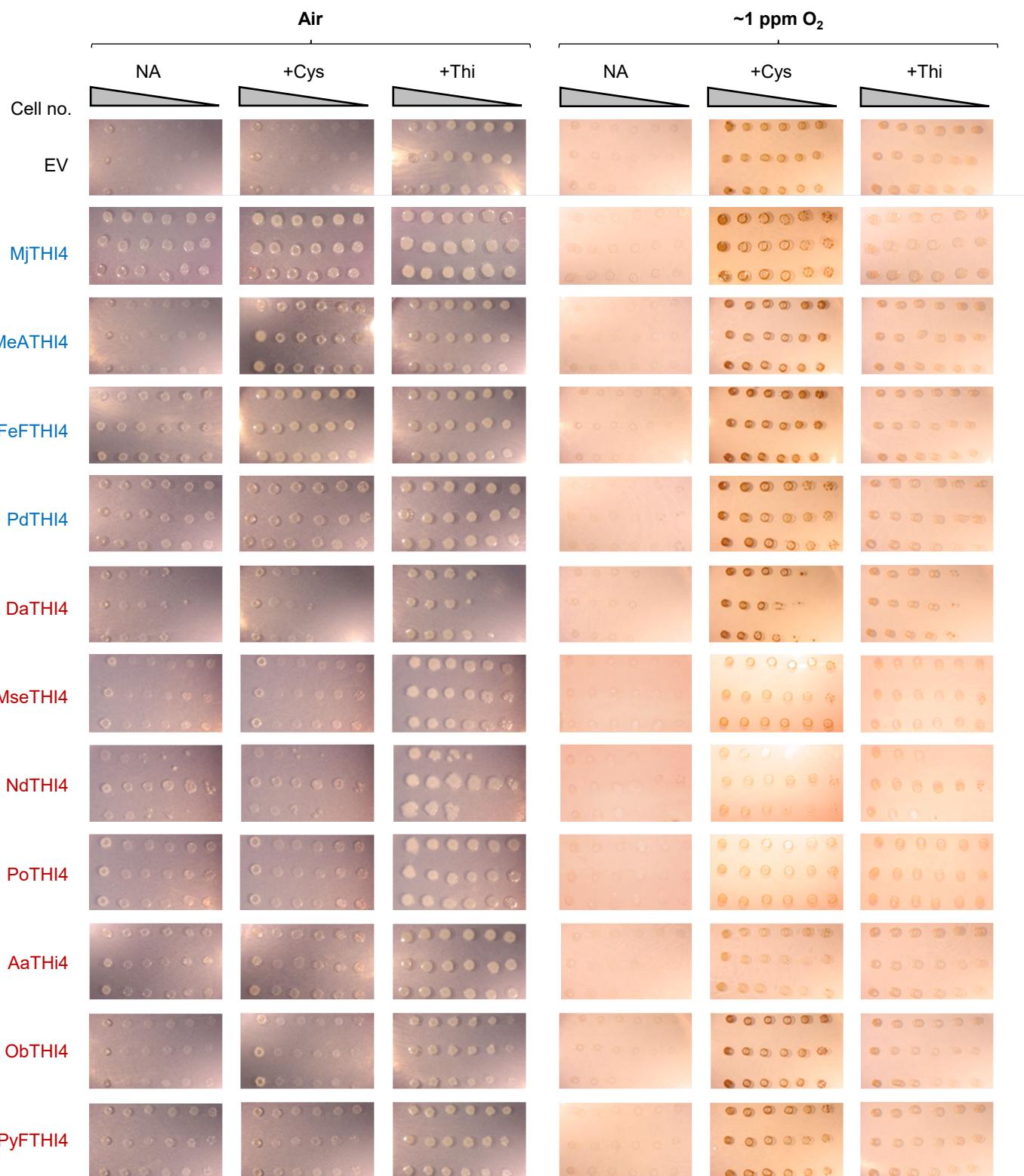
The first column (colored bullets) shows the O<sub>2</sub> adaptation of the 15 bacteria and 11 archaea whose THI4s were tested. The second column (colored squares) shows which residue replaces the active-site cysteine in each THI4. The next four columns (black or white squares) indicate the presence or absence of other thiamin synthesis enzymes (ThiG, ThiC, ThiD, and ThiE or ThiN). The last column (red or white squares) indicates presence or absence of a gene encoding a protein from the TRASH family (Trafficking, Resistance, And Sensing of Heavy metals) that is clustered with the THI4 gene.



**Supplementary Figure 2. Soluble expression of non-Cys THI4s.**

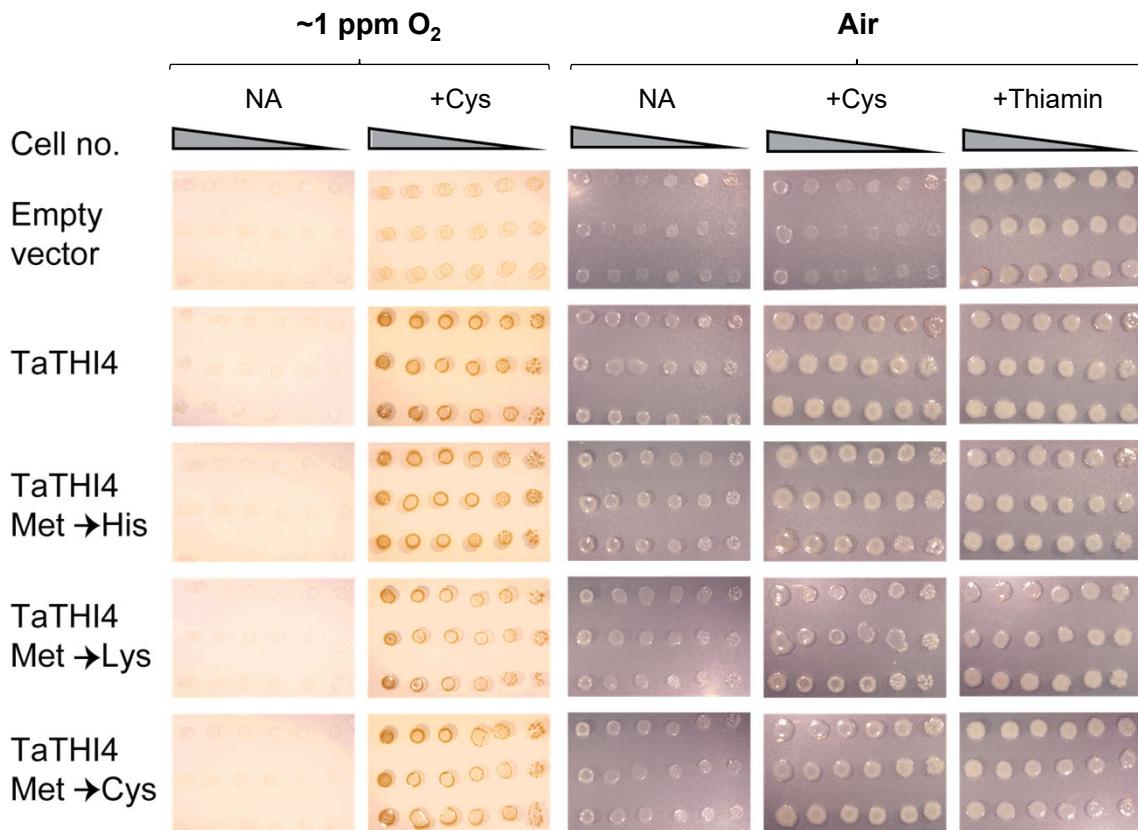
Quantitative gel analysis of (**A**) soluble and (**B**) insoluble expression in *E. coli* of 26 selected non-Cys THI4s. Soluble and insoluble fractions of cells were run on 15% gels, stained with Coomassie blue, and scanned to quantify the THI4 band, for which purified recombinant *Thermovibrio ammonificans* THI4 (RTa, arrow) served as a marker. Organism abbreviations are as in Table 1.





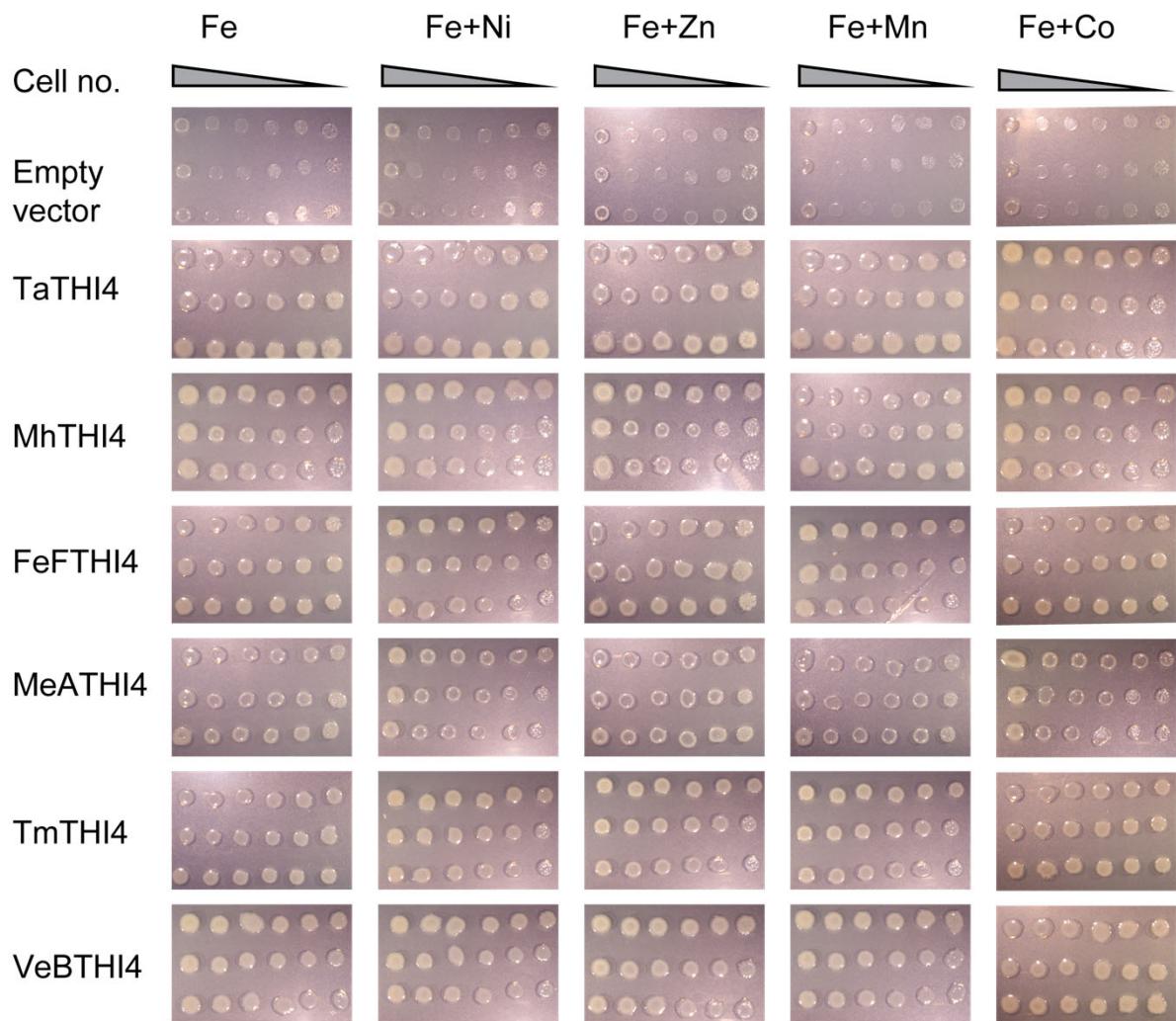
**Supplementary Figure 3. Functional complementation tests of non-Cys THI4s.**

Tests of functional complementation of an *E. coli*  $\Delta\text{thiG}$  strain by all 23 soluble non-Cys THI4s or the empty vector (EV). Organism abbreviations are as in Table 1. Overnight cultures of three independent clones per construct were 10-fold serially diluted and spotted on plates of MOPS minimal medium containing 0.2% glycerol and 0.02% arabinose with no additions (NA) or plus 1 mM Cys or 100 nM thiamin. Cells were cultured in air or  $\sim 1 \text{ ppm O}_2$ . The medium used for culture in  $\sim 1 \text{ ppm O}_2$  contained 40 mM nitrate. Images were captured after incubation at 37°C for 7 d. The high background in the  $\sim 1 \text{ ppm O}_2$ +Cys treatment is staining of the inoculum cells. Organisms whose THI4s showed clear complementing activity in air, particularly with Cys supplementation, are blue; organisms whose THI4s did not show such activity are red. Note that complementing activity was scored from direct visual inspection of plates, not from the above images, which do not fully capture growth in every case.



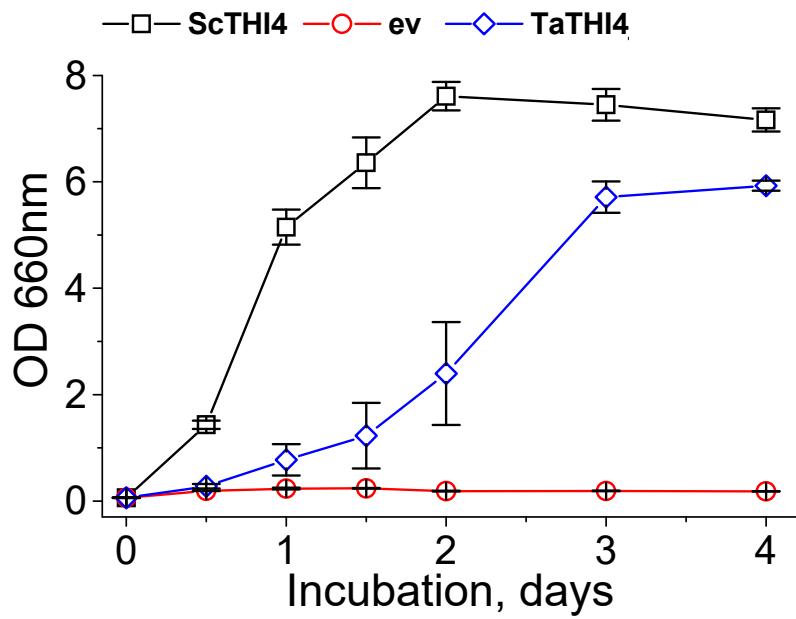
**Supplementary Figure 4. Complementation activity of TaTHI4 mutants**

An *E. coli*  $\Delta thiG$  strain was transformed with empty vector or vector harboring wild type TaTHI4 or TaTHI4 with the indicated mutations of Met158. Overnight cultures of three independent clones per construct were 10-fold serially diluted and spotted on plates of MOPS minimal medium containing 0.2% (w/v) glycerol and 0.02% (w/v) arabinose with no addition (NA) or with 1 mM Cys or 100 nM thiamin. The medium used for culture in ~1 ppm O<sub>2</sub> contained 40 mM nitrate. Cells were cultured in air or under N<sub>2</sub> containing ~1 ppm O<sub>2</sub>. Images were captured after incubation at 37°C for 7 d.



**Supplementary Figure 5 Effect of metal supplementation on complementing activity of THI4s**

An *E. coli*  $\Delta thiG$  strain was transformed with empty vector or vector harboring the indicated THI4 sequence. Overnight cultures of three independent clones per construct were 10-fold serially diluted and spotted on plates of MOPS minimal medium containing 0.2% (w/v) glycerol, 0.02% (w/v) arabinose, 1 mM Cys, and 100  $\mu$ M of the indicated metal. All media also contained the standard concentration of ferrous iron (100  $\mu$ M). Cultures were incubated in air at 37°C for 7 d.



**Supplementary Figure 6 Complementation of a yeast  $\Delta THI4$  strain by TaTHI4**

Cells of the  $\Delta THI4$  strain transformed with empty vector (ev) or vector harboring TaTHI4 or yeast THI4 (ScTHI4) as positive control were cultured in thiamin-free SC minus histidine medium. Data are means of three or four independent clones  $\pm$  S.E.

**Supplementary Table 2** *T. ammonificans* THI4 data collection and refinement statistics

TaThi4	
<b>Data collection</b>	
Space group	I121
Cell dimensions	
$a, b, c$ (Å)	89.89, 89.69, 131.81
$\alpha, \beta, \gamma$ (°)	90, 96.98, 90
Resolution (Å)	43.63-2.30 (2.36-2.30)*
$R_{\text{merge}}$	0.092 (1.552)
$I / \sigma I$	8.23 (0.77)
Completeness (%)	99.46 (96.62)
Redundancy	2.0 (2.0)
<b>Refinement</b>	
Resolution (Å)	43.63-2.30
No. reflections	93685
$R_{\text{work}} / R_{\text{free}}$	0.217/0.275
No. atoms	
Protein	7835
Ligand/ion	160
Water	58
$B$ -factors	
Protein	50.69
Ligand/ion	41.32
Water	47.34
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.12

\*Values in parentheses are for highest-resolution shell.

### Supplementary Table 3 Sequences of 199 non-Cys THI4s from SEED and UniRef90 databases

The 26 sequences selected for testing are boxed. Those that had little to no complementing activity are in **red font**. Those that were poorly expressed in soluble form in *E. coli* and were not tested are in **gray font**. The residue that replaces Cys in the active site is highlighted in **cyan**. Met and non-active-site Cys residues in each sequence are highlighted in **yellow**; the number per sequence is given in the header. Mean values were: Active THI4s: 10.6 Met, 2.3 Cys. Inactive THI4s: 9.1 Met, 3.1 Cys.

>Caldanaerovirga acetigignens (2 Cys, 11 Met)
MKSFSIPVPTDKVSSLIMKHYFKDLEDAVKSDVIVAGAGPSGLTC <sup>A</sup> WTLADQGYKVTLDRLAPGGIWGGAMSFNKVVLQKDVEWILKEADVPFV EDEGALVVSAPIFLASKLIAAKAAAHPGIRFN <sup>M</sup> LTVVDLHSSGDRITGVVVNNSAIE <sup>M</sup> AGLH <sup>V</sup> DPMVLTAKVLATGHDAVLANLYSRRAGTGLIRE SF <sup>M</sup> NAEKGEEDVVANTRMLAPGLFVAGMAANNVEGGCRMGP <sup>I</sup> FGGMLSGKKAARLIINYLSSENSK
>Candidatus Marinimicrobia bacterium (4 Cys, 12 Met)
MEKIVSFGIIDSYQKKLKENLEVDVAIVGGGPSGLIAAKYLAQAGKKVVLFERKLAPGGGMWGGAMMFNFNQIVVQEDAISILEDVDISYNLYEEGYVV CD <sup>S</sup> VEATAALIIFSAKKAGATIFN <sup>C</sup> SVEDVVVFQRGSVAGVVVNWSVHREG <sup>M</sup> YVDP <sup>H</sup> LV <sup>M</sup> AKA <sup>V</sup> LDSTGH <sup>C</sup> EVASILARKNEV <sup>K</sup> LM <sup>T</sup> RTGNIM <sup>G</sup> ER SLSIEEGELTTIENTKEIFPGLYVSGMAANAVSGSFRMGP <sup>I</sup> FGGMLSGKKAARLIINYLSSENSK
>Fervidicola ferrireducens (2 Cys, 11 Met)
MKSFSVPVPDTKVSSLIMKHYFKDLEDAVKSDVIVAGAGPSGLTC <sup>A</sup> WTLADQGYKVTLDRLAPGGIWGGAMSFNKVVLQKDVEWILKEADVPFV EDEGALVVSAPIFLASKLIAAKAAAHPGIRFN <sup>M</sup> LTVVDLHSTGDKITGVVVNNSAIE <sup>M</sup> AGLH <sup>V</sup> DPMVLTAKVLATGHDAVLANLYSRRAGTGLIRE SF <sup>M</sup> NAEKGEEDVVANTRMLAPGLFVAGMAANNVEGGCRMGP <sup>I</sup> FGGMLSGKKAARLIINYLSSENSK
>Hippea maritima strain ATCC 700847 (2 Cys, 9 Met)
MNNLDERVISRAIVERYMNKLLDYLC <sup>D</sup> VTIVGGGPAGLVC <sup>A</sup> YYI <sup>L</sup> AKANIKVAIFDKRLTIGGGMWGGAMLFNEIVVQEIGREILDEFGINYEKYTD GYTADSIEATT <sup>T</sup> LIKSTKVAGAKIFN <sup>C</sup> FAI <sup>E</sup> VEDVVFKKIDGQYR <sup>V</sup> NGLVVGWT <sup>V</sup> NNAGL <sup>I</sup> VDPLV <sup>T</sup> SKYVIDATGH <sup>D</sup> ADIANILRKGGIKLNTP EGVVIGEKPM <sup>W</sup> AEVGEQ <sup>S</sup> TIETQEV <sup>V</sup> PG <sup>L</sup> IVAC <sup>M</sup> AAAVAVSGSHRMGP <sup>I</sup> FGGM <sup>L</sup> NSGKKAARQIVIESLKK
>Marinilabilia salmonicolor (2 Cys, 12 Met)
MEQIVSSGIIDSYFSKLKENLA <sup>D</sup> VAIVGGGPSGLIAAYYLAKTGKVALFERKLAPGGGMWGGAMMFNEIMVQKEALHILKELGIEYKHYRDDYYT VDSVHATSALTYHATKAGARIFN <sup>C</sup> T <sup>S</sup> I <sup>E</sup> EDVV <sup>F</sup> HN <sup>N</sup> IVSGL <sup>V</sup> INWAPV <sup>H</sup> REG <sup>M</sup> HVDPLI <sup>I</sup> MAKAVIDGTGH <sup>C</sup> EVHTVARKNDIKIDTPSGKVM <sup>G</sup> ER SLAVEEAERTTVDNTKEV <sup>F</sup> PGLFVSG <sup>M</sup> AA <sup>N</sup> GTSGSYR <sup>M</sup> GP <sup>I</sup> FGGM <sup>L</sup> LSGQKVAGI <sup>S</sup> EKLAKEA <sup>I</sup> MESANN
>Methanocaldococcus jannaschii DSM 2661 (= Methanococcus jannaschii) (3 Cys, 10 Met)
MVNLMNIKDIKLN <sup>A</sup> DET <sup>K</sup> T <sup>T</sup> KAILKASFD <sup>M</sup> WLDIVEADVVIVGAGPSGLTC <sup>A</sup> RYLAKEGFKVVVLERHLAFGGTWGGGMGF <sup>P</sup> YIVVEEPADELLRE VG <sup>I</sup> KLID <sup>M</sup> GD <sup>G</sup> YYVADSVEVPAKLA <sup>A</sup> MDAGAKILT <sup>G</sup> IVVEDL <sup>L</sup> REDGVAGVVINSY <sup>A</sup> IERAGL <sup>H</sup> IDPLT <sup>I</sup> R <sup>S</sup> KVV <sup>V</sup> DAT <sup>G</sup> HEASIVN <sup>I</sup> LVKKNK LEADVPGEKSM <sup>W</sup> A <sup>E</sup> KGENALLRN <sup>T</sup> REV <sup>V</sup> PGLFV <sup>C</sup> GA <sup>M</sup> AA <sup>N</sup> ASHGGYR <sup>M</sup> GA <sup>I</sup> FGGM <sup>L</sup> YLSGKLC <sup>A</sup> ELITEKLKN <sup>N</sup>
>Methanotorris igneus (= Methanococcus igneus) (2 Cys, 8 Met)
MDVRLRADEYATTRAILKSAFD <sup>M</sup> WLDI <sup>I</sup> VDV <sup>D</sup> VAIVGGGPSGLTA <sup>A</sup> RYIAKEGYKV <sup>V</sup> VL <sup>R</sup> HLAFGGTWGGGMGF <sup>P</sup> YIVVEEPADEILREVGV <sup>K</sup> LEK VEGEDGLY <sup>T</sup> ADSVEVPAKLA <sup>G</sup> VIDAGAKVL <sup>T</sup> GIV <sup>V</sup> EDVL <sup>R</sup> ENR <sup>V</sup> AGVVINSY <sup>A</sup> IEKAGL <sup>H</sup> IDP <sup>I</sup> ITAKYVV <sup>D</sup> ATGH <sup>D</sup> ASV <sup>T</sup> TSRKN <sup>P</sup> ELGLE VPGEKSM <sup>W</sup> A <sup>E</sup> KGENALLRN <sup>T</sup> REV <sup>V</sup> PGLFV <sup>C</sup> GA <sup>M</sup> AA <sup>N</sup> AVYAGHR <sup>M</sup> GA <sup>I</sup> FGGM <sup>L</sup> YISGK <sup>K</sup> CAEM <sup>I</sup> VEKLKN <sup>N</sup>
>Methanococcus aeolicus strain ATCC BAA-1280 (3 Cys, 8 Met)
MDISKIDLKADEKAVTKSIFKATYEM <sup>M</sup> DNLEVDVVIVGGGPSGLTAGRYLADAGVKVL <sup>L</sup> ILERHLSFGGTWGGGMCP <sup>P</sup> YITVQSPADEIILSEVG <sup>I</sup> K LEE <sup>E</sup> GEDGLFVADSVEVPAKLG <sup>T</sup> GA <sup>I</sup> DAGAKVL <sup>T</sup> GIV <sup>V</sup> EDVL <sup>R</sup> EGKVS <sup>V</sup> VINSY <sup>A</sup> INKAGL <sup>H</sup> IDPLT <sup>I</sup> NAK <sup>V</sup> VIDATGH <sup>D</sup> ASV <sup>T</sup> CTLARKNED <sup>I</sup> GL VIPGEKSLWADEGENGL <sup>L</sup> Y <sup>K</sup> T <sup>T</sup> KE <sup>L</sup> FPGLFV <sup>C</sup> GA <sup>M</sup> SNATHGGYR <sup>M</sup> GA <sup>V</sup> FGGM <sup>L</sup> YISGK <sup>K</sup> IVADMILEKLKN <sup>E</sup>
>Mucinivorans hirudinis (3 Cys, 10 Met)
MEKIVSAGIVESYFDKLRRNLVLD <sup>A</sup> VAIVGGGPSGLVAA <sup>A</sup> Y <sup>L</sup> AKAGR <sup>R</sup> VALFERKLAPGGGMWGGAMMFNDIVVQSDALPILEELGV <sup>S</sup> YRH <sup>R</sup> DAYL VDSVHATA <sup>A</sup> LIY <sup>A</sup> TRAGATIFN <sup>C</sup> YS <sup>S</sup> VEDVVFK <sup>D</sup> ER <sup>V</sup> AGL <sup>V</sup> VNW <sup>A</sup> PIREG <sup>M</sup> HVDPL <sup>I</sup> V <sup>M</sup> ATA <sup>V</sup> LEG <sup>T</sup> GH <sup>D</sup> CA <sup>I</sup> ARL <sup>V</sup> ARKNGVR <sup>L</sup> NT <sup>P</sup> GEVIGER SLSIEEAERTTVENTKEIYPGLFVSG <sup>M</sup> AA <sup>N</sup> GVSGSFRMGP <sup>I</sup> FGGM <sup>L</sup> YLSGK <sup>K</sup> AAQ <sup>M</sup> CD <sup>S</sup> L
>Parabacteroides chinchillae (3 Cys, 11 Met)
MEQIVSTGIIDSYFAKLKS <sup>N</sup> LSVD <sup>A</sup> VAIVGGGPSGLVAA <sup>A</sup> Y <sup>L</sup> AKAGR <sup>K</sup> VALFDRKLAPGGGMWGGAMMFNDIVVQSDALPILEELGV <sup>S</sup> YHAAGNG <sup>T</sup> YI MDSVHTTSALIYQATKAGATIFN <sup>C</sup> YS <sup>S</sup> VEDVV <sup>F</sup> H <sup>N</sup> DA <sup>V</sup> AGVVVN <sup>W</sup> AP <sup>V</sup> IREG <sup>M</sup> HVDPL <sup>I</sup> MA <sup>K</sup> AV <sup>L</sup> EG <sup>T</sup> GH <sup>D</sup> CE <sup>V</sup> ARTVARKNDIKLNT <sup>P</sup> GG <sup>V</sup> IGER SLNVELGESTTVENTKEIYPGLFVSG <sup>M</sup> AA <sup>N</sup> GVSGSFRMGP <sup>I</sup> FGGM <sup>L</sup> YLSGK <sup>K</sup> AAELIC <sup>M</sup> CD <sup>L</sup> KG
>Pseudoramibacter alactolyticus ATCC 23263 (2 Cys, 19 Met)
M <sup>L</sup> SDTKISEA <sup>I</sup> LT <sup>T</sup> Y <sup>T</sup> DRFKQM <sup>L</sup> SSDA <sup>V</sup> IVGGGPSGLVAA <sup>A</sup> Y <sup>L</sup> KGAGVK <sup>T</sup> LLDR <sup>R</sup> LSVGGGMWGGGMM <sup>M</sup> NQIVVQ <sup>K</sup> SV <sup>L</sup> PILEEM <sup>M</sup> GIACKYDAEH YT <sup>V</sup> SSVAC <sup>I</sup> SGLIFRAAQSGATIM <sup>N</sup> L <sup>V</sup> T <sup>M</sup> EDAVV <sup>R</sup> GR <sup>L</sup> EV <sup>G</sup> L <sup>V</sup> IN <sup>W</sup> ST <sup>V</sup> TE <sup>M</sup> AHLM <sup>V</sup> DPL <sup>M</sup> MDAR <sup>V</sup> VL <sup>D</sup> ATGH <sup>D</sup> AA <sup>L</sup> VT <sup>K</sup> L <sup>V</sup> ER <sup>M</sup> GP <sup>V</sup> LN <sup>T</sup> PSGG <sup>L</sup> E <sup>G</sup> E <sup>K</sup> PK <sup>M</sup> WADHG <sup>E</sup> KQV <sup>V</sup> ANT <sup>T</sup> REV <sup>V</sup> PG <sup>L</sup> YVSG <sup>M</sup> AA <sup>N</sup> AT <sup>T</sup> FG <sup>G</sup> Q <sup>R</sup> MG <sup>P</sup> V <sup>F</sup> GG <sup>M</sup> YLSGK <sup>K</sup> AAE <sup>M</sup> LR <sup>R</sup> LAQ
>Pyrodictium delaneyi (0 Cys, 10 Met)
M <sup>G</sup> IASF <sup>F</sup> Y <sup>P</sup> GE <sup>E</sup> L <sup>K</sup> QY <sup>S</sup> EA <sup>K</sup> L <sup>A</sup> IK <sup>V</sup> ALE <sup>K</sup> LS <sup>Y</sup> AVE <sup>A</sup> DV <sup>V</sup> I <sup>A</sup> GAG <sup>G</sup> P <sup>A</sup> GL <sup>T</sup> LA <sup>W</sup> L <sup>I</sup> A <sup>E</sup> Q <sup>G</sup> LR <sup>V</sup> T <sup>L</sup> V <sup>E</sup> HR <sup>L</sup> ST <sup>T</sup> GGG <sup>M</sup> KGG <sup>M</sup> LF <sup>P</sup> VAL <sup>V</sup> EELA <sup>A</sup> VL E <sup>K</sup> AG <sup>V</sup> R <sup>L</sup> HR <sup>V</sup> GE <sup>G</sup> LY <sup>A</sup> M <sup>D</sup> P <sup>V</sup> EA <sup>V</sup> AK <sup>L</sup> T <sup>A</sup> RA <sup>V</sup> DAG <sup>A</sup> V <sup>I</sup> L <sup>P</sup> GL <sup>H</sup> VE <sup>D</sup> LI <sup>V</sup> R <sup>G</sup> S <sup>G</sup> S <sup>N</sup> VR <sup>V</sup> AG <sup>I</sup> V <sup>V</sup> N <sup>W</sup> AP <sup>V</sup> VEAGW <sup>H</sup> V <sup>D</sup> PL <sup>I</sup> Y <sup>E</sup> AR <sup>V</sup> V <sup>D</sup> ATGH <sup>D</sup> QL <sup>A</sup> R <sup>L</sup> L <sup>E</sup> RR <sup>R</sup> PG <sup>S</sup> LS <sup>K</sup> V <sup>P</sup> GM <sup>S</sup> SL <sup>D</sup> W <sup>T</sup> GER <sup>Q</sup> V <sup>V</sup> EHT <sup>G</sup> E <sup>I</sup> F <sup>P</sup> GL <sup>Y</sup> A <sup>G</sup> M <sup>S</sup> VA <sup>E</sup> V <sup>N</sup> L <sup>R</sup> R <sup>M</sup> GP <sup>V</sup> FG <sup>G</sup> M <sup>I</sup> ASA <sup>R</sup> LA <sup>E</sup> M <sup>L</sup> A <sup>E</sup> RL <sup>A</sup> KG <sup>R</sup> M <sup>G</sup> LAT <sup>G</sup> VA <sup>R</sup> SG

>Saccharicrinis fermentans DSM 9555 (2 Cys, 10 Met)  
**M**EQIVSVGIVDVF~~K~~KLKENLTVDVAIVGGGPGSMVAA~~Y~~YLARQGFKVSV~~E~~RKLAPGGGMWGGAMMFNEIVI~~Q~~KEALPILDELNISYKHYDKDYYT  
 LDSVHATSALIYHATQAGATFFNCTSVEDVVFLDNKVSGVVLNWAPVREKMHVDP~~L~~VIMAKAVIDGTGHDCDIARILERKNNIQLLTASGV~~E~~GER  
 SLSIDEAERTTIENTKEIYPGLYVSGMASNGVSGGFRMGP~~I~~F~~G~~GMILSGKKVANLIADNLNK

>Thermotoga maritima strain ATCC 43589 (2 Cys, 8 Met)  
**M**RDVLISRL~~V~~EYFEKLRNSLELDVAIVGAGPSGLTAAYELAKNGF~~R~~VAVFEERNTPGGGMWGGMMFNEIVL~~E~~KELENFLIKEVEIEYEVKEDHIV  
 VDSVHFASG~~L~~YRATKAGAIVFNNVSVEDAVQNGRVC~~G~~VVNW~~G~~P~~T~~RLGLHVDPITVKASFVVDG~~T~~GHPANV~~S~~LLAKRGLVEMKTEFPMDADEA  
 EKFVVDNTGEIFPG~~L~~VSGMAVCAVHG~~G~~PRMGP~~I~~F~~G~~MLSGQKVARI~~V~~SERLR

>Thermovibrio ammonificans strain DSM 15698 (3 Cys, 10 Met)  
**M**QNLSSEV~~V~~ISEAI~~T~~A~~F~~MEKLKSH~~E~~TDVAIVGGGPG~~L~~VAG~~Y~~YLA~~K~~KG~~Y~~RVA~~I~~F~~R~~RLS~~I~~GGGMWAGAMFFNEIVVQEMGREILDEFGVNYREFKP  
 GY~~Y~~LA~~D~~AVEATTIASKAVKAGATV~~F~~N~~G~~V~~T~~A~~E~~D~~V~~V~~L~~K~~Q~~V~~N~~Q~~Y~~R~~V~~C~~G~~L~~V~~IN~~W~~T~~T~~V~~E~~L~~H~~M~~V~~D~~P~~L~~V~~I~~T~~A~~K~~Y~~V~~D~~A~~T~~G~~H~~D~~A~~S~~V~~V~~STLQRKAGIKLNTE  
 TG~~V~~V~~G~~E~~K~~P~~L~~W~~A~~S~~V~~GE~~D~~TV~~K~~N~~S~~KE~~V~~FP~~G~~I~~Y~~V~~S~~GM~~A~~N~~A~~T~~C~~GS~~H~~R~~M~~GP~~V~~FG~~G~~MLMSGKKV~~A~~E~~E~~IA~~A~~KL~~N~~Q~~N~~KEA

>Verrucomicrobia bacterium (1 Cys, 11 Met)  
**M**LN~~E~~VTISRA~~I~~DAYFKKLTRH~~L~~EV~~D~~V~~A~~I~~V~~GGGPG~~S~~LVAG~~H~~DLAR~~G~~KKVALFES~~K~~LA~~I~~GGGIW~~G~~GGGMGF~~N~~EIVVQEAAREM~~L~~VEFGLR~~A~~TEFEP~~G~~Y  
 YT~~I~~DA~~V~~H~~V~~AA~~A~~LA~~A~~RA~~M~~EA~~G~~LT~~V~~FN~~L~~T~~S~~MD~~V~~V~~I~~Q~~K~~DR~~V~~AG~~L~~V~~N~~WT~~A~~IR~~H~~L~~K~~W~~H~~VD~~P~~L~~T~~I~~H~~SR~~F~~V~~L~~D~~A~~T~~G~~H~~P~~AS~~V~~A~~E~~T~~L~~V~~R~~K~~M~~N~~V~~R~~L~~D~~T~~T~~G~~GLVG  
 EK~~M~~AA~~E~~DG~~E~~R~~Q~~T~~V~~ENT~~R~~EV~~V~~Y~~P~~GL~~F~~V~~S~~GA~~A~~IT~~V~~C~~G~~G~~H~~R~~M~~GP~~V~~FG~~G~~MLMSGKKV~~A~~E~~E~~IA~~A~~KL~~N~~Q~~N~~KEA

>Candidatus Omnitrophica bacterium 4484\_171 (8 Cys, 9 Met)  
**M**LEETIISKAI~~D~~SYHNKLS~~I~~IDV~~D~~AA~~I~~CGGGPG~~S~~LV~~C~~AS~~L~~AA~~G~~KKV~~V~~L~~F~~E~~K~~KL~~S~~LG~~G~~WWGGGMF~~N~~EIVVQKKAKKILDEF~~S~~V~~R~~T~~K~~KY~~K~~EN~~Y~~  
 YLAD~~S~~SEC~~V~~C~~A~~LGYN~~A~~V~~H~~SG~~A~~V~~I~~ING~~V~~FA~~E~~D~~V~~V~~K~~KN~~R~~I~~C~~GL~~V~~IN~~W~~SA~~A~~AS~~A~~NL~~T~~VD~~P~~L~~T~~V~~R~~A~~K~~F~~V~~V~~D~~AT~~G~~H~~P~~SE~~V~~V~~K~~V~~E~~K~~S~~GV~~K~~I~~K~~T~~K~~G~~V~~LG~~E~~K~~S~~M~~W~~A~~H~~A~~E~~NT~~E~~IK~~N~~TRQ~~I~~AP~~G~~L~~F~~V~~T~~G~~C~~AN~~A~~V~~G~~AP~~R~~M~~G~~PI~~F~~GG~~M~~LLSGKK~~C~~AK~~I~~LSRL

>Poribacter~~a~~ sp. WGA-A3 (5 Cys, 10 Met)  
**M**DNLQP~~A~~PLR~~R~~ERDV~~T~~RIAREFYKEFDQ~~L~~IES~~D~~V~~I~~IVGGGPG~~S~~LV~~C~~AHDLATQ~~G~~FRT~~L~~LIEQ~~S~~SLALGGGFWSGGYL~~M~~NKAT~~I~~C~~E~~PAHSILENMGVPC  
 K~~P~~V~~K~~D~~C~~AG~~M~~R~~I~~V~~D~~P~~H~~AT~~A~~RIAS~~Y~~AG~~A~~K~~V~~N~~L~~TR~~V~~D~~L~~L~~H~~GE~~G~~V~~L~~E~~G~~V~~V~~V~~N~~TTAE~~M~~A~~G~~H~~D~~~~M~~I~~H~~VD~~P~~I~~A~~LES~~R~~V~~V~~V~~D~~AT~~G~~H~~D~~A~~V~~V~~V~~GL~~N~~Q~~R~~GL~~Y~~AT~~V~~PG~~N~~GA~~M~~V~~W~~ARSE~~A~~~~M~~V~~V~~D~~N~~T~~R~~E~~V~~FP~~N~~C~~F~~V~~T~~GL~~A~~V~~A~~AV~~D~~G~~S~~PR~~C~~G~~P~~AF~~G~~SM~~M~~LLSG~~R~~RA~~A~~D~~L~~V~~R~~H~~K~~L~~K~~G~~E~~

>Nitrospira defluvii (6 Cys, 12 Met)  
**M**EEL~~A~~RS~~R~~K~~A~~CT~~A~~VEGEYRM~~G~~K~~P~~K~~P~~A~~L~~R~~R~~ER~~D~~IT~~R~~Q~~I~~ARE~~Y~~Y~~K~~E~~F~~D~~Q~~L~~I~~E~~S~~D~~V~~I~~I~~VG~~A~~G~~P~~GL~~I~~C~~A~~HD~~L~~GR~~M~~GI~~K~~T~~L~~IVE~~Q~~SLALGGGFWSGGYLM~~N~~KAT~~I~~C~~A~~PA~~H~~K~~I~~KE~~V~~G~~V~~P~~C~~Q~~I~~KE~~C~~PG~~M~~Y~~V~~D~~P~~PHAT~~G~~ALIA~~A~~AYNAG~~A~~K~~M~~N~~L~~TR~~V~~D~~L~~L~~R~~REG~~V~~LEG~~V~~V~~N~~TTAE~~M~~A~~G~~H~~D~~~~M~~I~~H~~VD~~P~~I~~A~~LES~~K~~IV~~V~~D~~A~~T~~G~~H~~D~~A~~V~~V~~V~~LL~~H~~K~~R~~G~~L~~Y~~Q~~Q~~V~~PG~~N~~GA~~M~~W~~V~~S~~R~~SEE~~E~~VM~~D~~RT~~G~~E~~V~~SP~~N~~C~~F~~V~~I~~GL~~A~~V~~A~~AV~~F~~G~~T~~PR~~M~~G~~P~~AF~~G~~SM~~M~~LLSG~~R~~GY~~G~~A~~E~~L~~I~~R~~D~~K~~L~~K~~N~~R

>Desulfurococcus amylolyticus strain DSM 18924 (1 Cys, 5 Met)  
**M**SLES~~H~~IT~~R~~VI~~W~~EAS~~R~~D~~W~~VEL~~S~~~~C~~DIV~~V~~V~~G~~AG~~P~~GL~~T~~AA~~K~~Y~~L~~A~~E~~KL~~G~~K~~T~~L~~V~~L~~E~~RR~~L~~S~~F~~GGG~~G~~IG~~G~~GG~~G~~ML~~H~~K~~T~~V~~V~~D~~E~~R~~G~~L~~G~~I~~L~~R~~D~~F~~N~~I~~R~~Y~~K~~P~~S~~I~~K~~G~~L~~LY~~V~~V~~D~~T~~A~~EL~~T~~AK~~L~~A~~G~~AL~~D~~A~~G~~K~~A~~I~~T~~PG~~I~~S~~E~~D~~V~~I~~V~~R~~Y~~N~~P~~RV~~Q~~GV~~V~~V~~E~~WS~~A~~QL~~S~~GT~~C~~VD~~P~~L~~F~~E~~S~~K~~A~~VI~~D~~AT~~G~~H~~D~~A~~E~~V~~L~~R~~I~~E~~K~~K~~N~~P~~E~~S~~K~~V~~K~~I~~P~~G~~E~~K~~S~~Y~~E~~K~~A~~D~~V~~D~~V~~V~~E~~Y~~T~~GR~~V~~I~~P~~GL~~Y~~AT~~G~~MA~~V~~AA~~V~~R~~G~~LN~~R~~M~~G~~PI~~F~~GT~~G~~MLLSGR~~K~~V~~A~~E~~A~~VI~~R~~D~~L~~E~~S~~AP~~K~~

>Candidatus Aenigmarchaeota archaeon (2 Cys, 10 Met)  
**M**GBIIFSKV~~S~~E~~K~~V~~T~~SA~~I~~V~~S~~GF~~I~~E~~K~~FI~~K~~I~~I~~E~~S~~D~~V~~I~~I~~IVGGGPG~~S~~LM~~A~~KE~~L~~SS~~K~~G~~K~~V~~V~~I~~I~~ER~~NN~~Y~~L~~GGG~~F~~WT~~G~~GYLM~~M~~N~~K~~IT~~V~~R~~H~~P~~G~~EE~~I~~L~~K~~D~~L~~G~~I~~P~~F~~E~~E~~FG~~E~~GL~~Y~~LA~~D~~G~~P~~H~~A~~CS~~K~~L~~I~~A~~T~~~~C~~D~~A~~VG~~K~~IL~~N~~~~M~~T~~T~~LED~~V~~VL~~K~~E~~G~~AV~~G~~V~~V~~V~~N~~WT~~P~~I~~E~~T~~P~~R~~E~~IA~~C~~VD~~P~~I~~A~~LES~~K~~V~~V~~VI~~D~~AT~~G~~H~~D~~A~~V~~V~~V~~KK~~I~~E~~E~~RG~~I~~L~~K~~G~~L~~K~~G~~Y~~G~~GA~~M~~V~~W~~E~~K~~SED~~M~~MV~~V~~K~~Y~~T~~G~~EV~~H~~PG~~L~~V~~V~~T~~G~~MA~~V~~ST~~F~~FF~~G~~LP~~R~~M~~G~~PT~~F~~FG~~G~~MLLSG~~K~~KA~~E~~VT~~M~~EL~~S~~R

>Pyrolobus fumarii strain DSM 11204 (0 Cys, 10 Met)  
**M**VI~~P~~GHMTTR~~R~~TAMP~~G~~LD~~A~~IT~~R~~V~~I~~EE~~A~~SK~~E~~LV~~E~~Y~~A~~E~~S~~D~~V~~I~~V~~V~~G~~AG~~P~~GL~~A~~F~~A~~Y~~L~~A~~K~~R~~G~~F~~R~~V~~L~~V~~L~~ERR~~L~~S~~F~~GGG~~G~~IG~~G~~GG~~G~~ML~~F~~H~~K~~V~~L~~V~~Q~~E~~E~~AL~~P~~V~~L~~ND~~M~~IR~~V~~H~~P~~TS~~V~~K~~G~~I~~Y~~SL~~D~~S~~V~~AL~~I~~IT~~G~~LA~~S~~AA~~V~~N~~A~~G~~A~~K~~I~~I~~L~~GLE~~A~~VD~~L~~V~~R~~KE~~E~~HR~~R~~V~~A~~G~~V~~MA~~L~~W~~S~~AV~~G~~I~~A~~N~~L~~H~~V~~D~~P~~I~~F~~E~~A~~K~~V~~V~~D~~AT~~G~~H~~D~~A~~E~~V~~L~~R~~I~~IA~~H~~Q~~K~~L~~R~~GE~~A~~P~~V~~PG~~D~~GP~~A~~WA~~E~~E~~G~~KE~~L~~V~~V~~K~~A~~T~~G~~E~~L~~P~~I~~GL~~Y~~V~~A~~G~~M~~AA~~T~~AV~~K~~GY~~Y~~RM~~G~~PI~~F~~GG~~M~~LLSG~~K~~KA~~D~~L~~I~~E~~K~~L~~R~~KG~~K~~

>Metallosphaera sedula strain ATCC 51363 (0 Cys, 8 Met)  
**M**NI~~K~~Q~~V~~DE~~I~~K~~I~~TRY~~L~~K~~A~~T~~F~~ED~~W~~~~D~~FS~~V~~N~~D~~V~~V~~I~~V~~G~~A~~G~~P~~GL~~A~~AA~~Y~~Y~~S~~A~~K~~AG~~L~~K~~T~~T~~V~~F~~E~~RR~~L~~S~~F~~GGG~~G~~IG~~G~~GG~~G~~ML~~F~~H~~K~~V~~I~~V~~E~~SP~~A~~DE~~I~~L~~R~~I~~E~~IG~~V~~K~~L~~Q~~K~~F~~E~~EG~~V~~V~~V~~D~~S~~SE~~F~~Y~~A~~KL~~A~~AA~~T~~IDI~~A~~G~~A~~K~~I~~I~~H~~GT~~V~~V~~D~~D~~V~~I~~F~~R~~E~~N~~P~~L~~R~~V~~G~~VA~~E~~WT~~T~~Q~~A~~SL~~I~~H~~V~~D~~P~~L~~F~~I~~S~~AK~~A~~V~~V~~D~~A~~T~~G~~H~~D~~A~~E~~V~~I~~S~~V~~AS~~R~~K~~I~~PE~~L~~G~~I~~V~~I~~P~~G~~E~~K~~S~~A~~Y~~E~~I~~A~~QL~~T~~VE~~Q~~S~~G~~EV~~A~~P~~G~~LY~~A~~AG~~M~~AV~~T~~E~~I~~K~~A~~IP~~R~~M~~G~~PI~~F~~GM~~M~~LLSG~~K~~KA~~E~~DI~~I~~KN~~L~~Q~~A~~S~~T~~LK~~S~~V~~Q~~K~~E~~

>Caldivirga maquilingensis strain ATCC 700844 (0 Cys, 10 Met)  
**M**AGISIREAS~~I~~TRA~~I~~V~~N~~S~~A~~KL~~L~~SE~~Y~~SS~~V~~D~~V~~AI~~V~~G~~A~~G~~P~~SG~~M~~T~~A~~AY~~Y~~Y~~L~~A~~K~~GL~~K~~T~~L~~V~~L~~ERR~~L~~S~~F~~GGG~~G~~IG~~G~~GA~~A~~SH~~L~~PS~~I~~I~~V~~E~~H~~P~~V~~SE~~I~~LS~~K~~D~~F~~GI~~K~~I~~M~~~~D~~GD~~G~~L~~F~~T~~V~~DP~~A~~E~~I~~IA~~K~~L~~V~~A~~K~~AI~~D~~AG~~A~~K~~F~~FL~~G~~V~~H~~V~~D~~D~~V~~I~~F~~R~~E~~N~~P~~L~~R~~V~~G~~VA~~E~~WT~~T~~Q~~A~~AG~~M~~V~~G~~HT~~D~~F~~F~~I~~S~~NA~~V~~V~~D~~AT~~G~~H~~D~~A~~E~~V~~A~~AS~~R~~K~~I~~PE~~L~~G~~I~~V~~I~~P~~G~~E~~K~~S~~A~~Y~~E~~I~~A~~EL~~V~~TL~~V~~GT~~K~~VID~~G~~LY~~V~~T~~G~~MA~~V~~AV~~H~~GL~~P~~RM~~G~~PI~~F~~GM~~M~~LLSG~~K~~KA~~E~~II~~I~~ED~~L~~K~~G~~N~~H~~

>Methanofollis liminatans DSM 4140 (5 Cys, 12 Met)  
**M**ELDEV~~T~~IS~~R~~AIL~~A~~T~~Q~~ME~~V~~Y~~L~~LD~~D~~V~~V~~V~~G~~GG~~P~~SG~~I~~TC~~A~~LLA~~E~~K~~G~~V~~K~~V~~G~~L~~F~~E~~K~~K~~L~~S~~I~~GGG~~M~~W~~G~~GG~~M~~F~~P~~RI~~V~~V~~Q~~A~~E~~K~~R~~I~~L~~D~~R~~F~~G~~I~~A~~S~~K~~E~~F~~E~~P~~Y~~H~~AK~~V~~AK~~S~~VE~~A~~V~~S~~KL~~A~~AA~~T~~ACT~~A~~GA~~E~~FF~~N~~LI~~A~~VE~~D~~V~~V~~I~~K~~G~~D~~R~~L~~AG~~L~~V~~V~~N~~W~~S~~P~~V~~E~~M~~A~~GL~~I~~I~~D~~PL~~T~~I~~R~~C~~K~~A~~V~~V~~D~~AS~~G~~H~~D~~AT~~I~~A~~H~~W~~A~~K~~G~~GD~~L~~PI~~R~~GE~~G~~F~~M~~W~~A~~D~~R~~A~~E~~G~~N~~I~~L~~E~~H~~TR~~E~~V~~F~~PG~~G~~LV~~C~~MA~~A~~NA~~A~~VA~~G~~AGE~~C~~RM~~G~~PI~~F~~GG~~M~~LLSG~~K~~KA~~D~~AA~~L~~I~~A~~VL~~A~~V~~L~~H~~P~~

>Sulfurisphaera tokodaii strain DSM 16993 (0 Cys, 10 Met)  
**M**DSNSIKV~~K~~Q~~V~~DEV~~V~~K~~I~~SKY~~I~~LY~~T~~Q~~F~~D~~W~~E~~I~~V~~S~~D~~V~~V~~I~~V~~G~~AG~~P~~SG~~M~~T~~A~~AY~~Y~~Y~~L~~A~~K~~GL~~K~~T~~V~~V~~F~~ERR~~L~~S~~F~~GGG~~G~~IG~~G~~GG~~G~~ML~~F~~H~~K~~V~~I~~V~~E~~SP~~A~~DE~~I~~L~~K~~E~~M~~K~~I~~KLN~~V~~EE~~G~~V~~Y~~I~~V~~DS~~A~~E~~F~~MA~~K~~LA~~A~~AA~~I~~D~~A~~G~~A~~K~~I~~I~~H~~GT~~V~~V~~D~~D~~V~~I~~F~~R~~E~~N~~P~~L~~R~~V~~G~~VA~~E~~WT~~T~~Q~~A~~AG~~M~~V~~G~~HT~~D~~F~~F~~I~~S~~NA~~V~~V~~D~~AT~~G~~H~~D~~A~~E~~V~~A~~AS~~R~~K~~I~~PE~~L~~G~~I~~V~~I~~P~~G~~E~~K~~S~~A~~Y~~E~~I~~A~~EL~~V~~TL~~V~~GT~~K~~VID~~G~~LY~~V~~T~~G~~MA~~V~~AV~~H~~GL~~P~~RM~~G~~PI~~F~~GM~~M~~LLSG~~K~~KA~~E~~II~~I~~IK~~D~~LR~~Y~~

>Deltaproteobacteria bacterium HGW-Deltaproteobacteria-1  
MVLDEIVISKAIIERFLEKLLQATDVDVAIVGGGPGSLVAAAYYLASAGKKVALFERKLSLGGGMWGGMMFNEIVVQDEAREILDVFDIYREYQQG  
YYTADAVLAVTSICSAARAGASIFNCVSVEDVMIREGRVTGLVINWSPVEMAGLHVDPDTIAAGSVIDTTGHATEVLKIERKADMQLATPSGKLV  
GERSMWAekaerlMDNTRQICPGVYVAGMSANAAGGPRMGPIFGGMLLSGRKVAELILASS

>Thermogladius calderae (strain DSM 22663 / VKM B-2946 / 1633)  
MELESIITRLVVEESARELVELSESVDLVVGAGPSGLTAALKYLADKHVKVVLEKRLSYGGGIGGGGSLFHKVVVDERALPVLGDFKVRYKAAGVAG  
YYVVDSAELMSKLAAGALDSGAKIILGAEVEDLVRDNPLRVRVGVFMFKWSAITAAGLHVDPFLFALSRAVVDATGHEAVLVSILSRKNRVAGVAVPGE  
RSGFAERAERDVVEYTGRMPVGLYVAGMSVAAVHGLHRMGPIFTGMLLSGRKVAEAIARDLGVPQ

>Acetomicrobium thermoterrenum DSM 13490  
MKLDELVITKAIVEGYFKLMMCLEDVAIVGGGPGSLVAAALELAKAGKKVALYERKLSVGGGMWGGMLFNEIVIQHEAKEILEGVGVNRPYEVE  
GYYTADSVEAVSTLTSKAVKAGATIFNALSVEDVVVDEERINGLTVNNWTAVEMAGLHVDPFLSIHCKYIDATGHDTEVVRVVARMPGRLFTATGN  
IEEGKFMSPDRAEKLTIVNTREVFPGLYVAGMAANATFGGPRMGPIFGGMLLSGVKAAREILSK1

>Acidianus hospitalis (strain W1)  
MQSIRIKQVNEVKISKYILKYTFEDWNNLVESDVVIVGAGPSGMATAAYYLAKAGLKTIVFERRLSFGGGIGGGAMNFHKVIETPADEIIKELKIRY  
IEPEEGIFIIDSAEFMAKLATAAIDAGAKIIHGVTDDVIFRENPLRAGVAVEWTSTQMSGHLHVDPFLFISAKAVVDACTHDAEIISVASRKVPELG  
IAVPGEKSAYSEIAEELVVENTGKAVPGLYATGMACVEKSLPRMGPIFGAMILSGKKVAEEIIKDLRNS

>Acidithiobacillales bacterium SM23\_46  
MCCQSLEAWRSEPEERKRNRADVVVVGAGPSGMATAIHLARERHRVILLEKRLSPGGGIWGGGMAMSEAIIVQDDALPWLDLGVRHKPSRGGLHSADA  
VELAAALCLTKVQSGTFLFNLLTVEDVCIHQDRVTGVVNRSMIAGALPVDFPIAFRTNAVIDATGHEAVVVEAVHKRGLLAHPAVAKPLGEGPMADA  
SGEAFVVENVKEVYPGLWICGMSVCATLGPRMGPIFGGMLLSGQRVAALVSSALTEFAQKDRESRK

>Aciduliprofundum boonei (strain DSM 19572 / T469)  
MLDEVEITKLIVENYMKDLMEYADLVAIVGAGPSGLTAAYYLATAKKVAIFDRRLSFIGGMWGGMMFNKIVVQEDAKHILDDFSINYERFGDYY  
VADSVHSVTSLAYHATKEGAKIFNLIGAEDVIICKNNRSGLVINWSVIGELPIDPLSIYAKYVIDATGHESEVIKTLVRKNNIKLNTPTGSIEGEHS  
MDADTAESVIVDNVKEVYPGLFVTGMAANAVFGSPRMGPIFGGMILSGKKVADEIIIRRLS

>Actinobacteria bacterium HGW-Actinobacteria-3  
MPLSEIEVTRGILEGFSRDFLSSLQSDVAIAGAGPSGMVCAYYLAREGLKVSFERNLHVGGGMWGGMLFPRIIIQEAAIREVEFGVRLKPFKEG  
YFVGDSVETVKVTAACIDAGVRVWVGVSVEDVLIREEVNLAGVVLNWRABELANLHVDPЛАVEAKVVDATGHEAVVVRTVARKIPGCRINTDTGG  
VIGEMPMWAQVGEELIVGNTREYANLLVTGMAANAVYGAAPRMGAFGGMFLSGYKCAHLAADIVRKA

>Alistipes inops  
MIETKVSQGIVSTYFDKLQKNLELDVAIVGGGPGSLVAAAYYLAKAGLRLAQGLKLTVFERRLSVGGGMWGGAMMFNYLFQEEARPIFETMGVRYREY  
VMDVESTSALLYHAVHAGATVFCNCYSEDVYVKENRSGVNVNTPVLRGLVDPNILARVVIDGTGDSEIAATVARKNGARLNTEGTVVGE  
RSLDVTAGEDEVVKGTKEIYPGLYVCGMAASAVSGTPRMGPIFGGMILMSGKKVADEIIARLKK

>Ammonifex degensii (strain DSM 10501 / KC4)  
MAGGAIDERLVSRAIIQTYSEELLQTLTDFDVAVVAGAGPSGLTAAYYLAKAGLRLAQGLKLTVFERRLSVGGGMWGGAMMFNYLFQEEARPIFETMGVRYREY  
QPGYYVAHSVEAAFTLAACRAGARIMNLITDVLRDNRVAGLVLNWTAVDMAGMHIDPLAVHCRYVVDATGHDAEVVRILTQKNQVTVKVPGG  
HVQGEKSMWSERGEKQTLDHSGEVFPGLYVAGMAANAVAGGYGRMGPIFGGMVLSGKKVAELILEAHRREKSQTL

>Ancylomarina sp. 16SWW S1-10-2  
MEQIVSAGIVDSYFKKLKENLSVDVAIVGGGPGSLVASYYLAKKGFKVALYETKLAPGGGMWGGAMMFNEIVVQKDALHILDELNSYTNYQGDYYT  
LDSVHATSALIYHATQAGVKIFNCSSIEDVVFQNNKVCVGVLNWSPRREGLHVDPFLVIMAKAVVDGTGHECDIVSTLERKNGVKLNTKTGKVMGEC  
SLSIDEAERTTVENTKEVYPGLYVSGMASNGVSGGFRMGPIFGGMILSGEKGKLAGLIAENLSK

>ANME-2 cluster archaeon HR1  
MELDEITITRAIILEDFTSDFLQSIDTDVALVGGGPANLIAARTLARAGVKTFLFERKLEVGGGMWGGMMMPRIIVQEEARHILDDLGVRYRKYEEG  
YYVADSIECTGKLIYEASSGASIYNLISVEDVMIREGRDAVTGLVINRTIVDMQKLHVDPITIRAKVVIDGTGDSEICTTLSRKIPGALHVAGEKP  
MWADVAERIILDNTKEVYPGLIVTGMAANAVAGAPRMGPIFGGMILSGEKAQIAIAKGL

>Archaeoglobales archaeon ex4484\_92  
MEARIKSAIIIEVAKDWSNISQVDVIVGAGPSGLTAGKYLAEKGLTLILERRLCFGGGIGGGMFLHKIVIEKFAREILDFFDVRYYEHDNLLVA  
DVAEFLMAKLAAGKAIIDAGAKIIHGVSVEDVIFRDDPLGVRGVCIQWSAELSGLHVDPMFIESKAVLDATGHDAEVVSIAASKVPLDLNVGTGEKSAYA  
ELGEKLVVEKTGKVVEGLYATGMACSVFNLPRMGPIFGGMLQSGKKAAEIIYNDLK

>Archaeoglobus fulgidus  
MEAETKAIVETASEEWVEYAESDVIVVAGAGPSGLTAARYLAEKGLTLVLERRLSFGGGIGGGMFLHKVVVEREAKDILDDFGIRYTEHRNFLVA  
DSAEMAKLAAGKAIIDAGAKIIHGVSVEDVIFRDDPLGVRGVCIQWSAELSGLHVDPMFIESKAVLDATGHDAEVVSIAASKVPLDLNVGTGEKSAYA  
EVAEREIVEKTGKIVKGLYAAAGMAAAVHNLP RMGPIFGGMILSGKVAEIVADELKL

>Armatimonadetes bacterium  
MSFEKDSFKWDELTIVTRGIVETFMADFLSIDLDVAIVVAGAGPSGLTAARILAGQGHRVGIFERNLHIGGGIWWGGMLFPRIIEEEAAMPLMEAAGV  
LRPWKDGTVIADAVESATKMTAAIIDAGARIFVGIEADVVVDDSDRVCVGVINWGAVTAAKLHVDPFLVHSKVLTESTGHPCEVGDVLLRKIPGAR  
LDTGETCPGEASMNARAGEAALIANTREIYPGVVAGMAANAVSRSPRMDAIFGGMILSGQKAAEISAQIIADLG

>bacterium (Candidatus Ratteibacteria) CG23\_combo\_of(CG06-09\_8\_20\_14\_all\_48\_7  
MLDDVVISRAIVETYFQDLLNYLENDVAIAGAGPAGLTAAYFLAKKGRKVAIFERQLRVGGMPGGMMFNKIVI  
YVADSLETTSLTSKAIQAGAKIFNLIAVEDLSIQEGRVNGLVNWSAVKTAGLHVDPITIRAKAVV  
DATGHDAVLCRLLVDKGKVTLRTPGKVAG  
EGPMWAEGEEMIANTGEVFPGLFIAGMTVNAVCGGPRMGPIFGGMLLSGERLSLLIP

>bacterium (Candidatus Stahlbacteria) CG23\_combo\_of(CG06-09\_8\_20\_14\_all\_40\_9  
MMLDETIISRAIETYKEKFVNLLKSDVAIAGGGPSGLIAGYYLKKKPDLKVVLFERKLSIGGGMWGGGMM  
NEYYTADSIETVSALALNTVKAGVTILNAISVEDTIEDNIAKGLVINWTSALDIGLHVDP  
LRADHIIDATGHPCEIAHLIEKKGKKLFTRTGKI  
IGEGAMYADKGERVIENTKELFPNVWACGMAANAVFGGPRMGPIFGGMLLSGKIVAEKILQKN

>bacterium 42\_11  
MKDILSKAILESSFNKLRSLELDVAIVGAGPSGLVASYELAKKKVIAIFERNTPGGIWGGGIMFNEVV  
LEKELEDFLKELDIKYKYVEDYIV  
VDSTHFASALIYHTTWGTRIFNSISVEDIAMQNRRVCV  
VINWGPVKLGLHVDP  
ITIKASYVVDGTGHPANVVSLLVKRG  
LLEKKTEFP  
MNAEEA  
EKVVEKTGEVFPGLLVSGMA  
CEVYGGPRMGPIFGGMVL  
SGRRIAEIITERVNKR

>Bacteroidales bacterium 6E  
MEQIVSSGIIDSYFKIKESLSV  
DVAIVGGGPSGLVAAYYLAKKGLKVAMFERKLA  
P  
GGGMWGGAMMFNEIVVQKGALQILDEFKIDYTHYEGDYYT  
LDSQATSSLIYHAGKAGARIFN  
CTSVEDVVFHNNKVS  
GIVLN  
WAP  
V  
RERLHVDP  
L  
VIMAKAV  
IDGTGHD  
CIARILERKNNIQLNTVSGKV  
QGER  
SLSIDEAERTTVENTKEIFPGLYV  
SGMAANGV  
SGGFRMGP  
IFGGM  
LMSGK  
V  
A  
L  
I  
M  
Q  
L  
N  
K

>Bacteroides cellulosilyticus  
MIETKVS  
KGII  
STYFEK  
LERN  
L  
LD  
VAI  
V  
GGG  
PSG  
I  
V  
A  
A  
Y  
L  
A  
K  
G  
L  
K  
V  
A  
Q  
F  
D  
R  
K  
L  
A  
P  
G  
G  
G  
M  
W  
G  
G  
A  
M  
M  
F  
N  
Q  
I  
V  
I  
Q  
E  
E  
A  
I  
D  
I  
V  
K  
E  
F  
N  
I  
N  
H  
E  
K  
Y  
E  
D  
G  
L  
Y  
V  
M  
D  
S  
V  
E  
S  
T  
S  
A  
L  
I  
Y  
Q  
A  
V  
H  
A  
G  
A  
T  
I  
F  
N  
C  
Y  
S  
V  
E  
D  
V  
F  
H  
N  
D  
A  
V  
A  
G  
V  
V  
V  
N  
W  
A  
P  
V  
I  
R  
E  
G  
M  
H  
V  
D  
P  
L  
N  
I  
L  
A  
K  
V  
V  
D  
G  
T  
G  
H  
D  
S  
E  
I  
A  
A  
T  
V  
A  
R  
K  
N  
G  
S  
R  
L  
A  
T  
E  
T  
G  
G  
V  
I  
G  
E  
R  
S  
L  
D  
V  
I  
A  
G  
E  
E  
E  
V  
V  
N  
G  
T  
K  
E  
I  
Y  
P  
G  
L  
Y  
V  
C  
G  
M  
A  
A  
S  
A  
V  
S  
G  
T  
P  
R  
M  
G  
P  
I  
F  
G  
G  
M  
L  
M  
S  
G  
K  
V  
A  
E  
E  
I  
I  
A  
K  
L  
K  
K

>Bacteroides sp. 3\_1\_19  
MEQIVSSGIIDSYFEKL  
KS  
NL  
S  
L  
S  
V  
D  
A  
I  
V  
G  
G  
G  
P  
S  
G  
I  
V  
A  
A  
Y  
L  
A  
K  
G  
L  
K  
V  
A  
Q  
F  
D  
R  
K  
L  
A  
P  
G  
G  
G  
M  
W  
G  
G  
A  
M  
M  
F  
N  
Q  
I  
V  
I  
Q  
E  
E  
A  
I  
D  
I  
V  
K  
E  
F  
N  
I  
N  
H  
E  
K  
Y  
E  
D  
G  
L  
Y  
V  
M  
D  
S  
V  
E  
S  
T  
S  
A  
L  
I  
Y  
Q  
A  
V  
H  
A  
G  
A  
T  
I  
F  
N  
C  
Y  
S  
V  
E  
D  
V  
F  
H  
N  
D  
A  
V  
A  
G  
V  
V  
V  
N  
W  
A  
P  
V  
I  
R  
E  
G  
M  
H  
V  
D  
P  
L  
N  
I  
L  
A  
K  
V  
V  
D  
G  
T  
G  
H  
D  
S  
E  
I  
A  
A  
T  
V  
A  
R  
K  
N  
G  
S  
R  
L  
A  
T  
E  
T  
G  
G  
V  
I  
G  
E  
R  
S  
L  
D  
V  
I  
A  
G  
E  
E  
V  
V  
N  
G  
T  
K  
E  
I  
Y  
P  
G  
L  
Y  
V  
C  
G  
M  
A  
A  
S  
A  
V  
S  
G  
T  
P  
R  
M  
G  
P  
I  
F  
G  
G  
M  
L  
M  
S  
G  
K  
V  
A  
E  
E  
I  
I  
A  
K  
L  
K  
K

>Bacteroides stercorisoris  
MIETKVS  
KGII  
STYFEK  
LERN  
L  
LD  
VAI  
V  
GGG  
PSG  
I  
V  
A  
A  
Y  
L  
A  
K  
G  
L  
K  
V  
A  
Q  
F  
D  
R  
K  
L  
A  
P  
G  
G  
G  
M  
W  
G  
G  
A  
M  
M  
F  
N  
Q  
I  
V  
I  
Q  
E  
E  
A  
I  
D  
I  
V  
K  
E  
F  
N  
I  
N  
H  
E  
K  
Y  
E  
D  
G  
L  
Y  
V  
M  
D  
S  
V  
E  
S  
T  
S  
A  
L  
I  
Y  
Q  
A  
V  
H  
A  
G  
A  
T  
I  
F  
N  
C  
Y  
S  
V  
E  
D  
V  
F  
H  
N  
D  
A  
V  
A  
G  
V  
V  
V  
N  
W  
A  
P  
V  
I  
R  
E  
G  
M  
H  
V  
D  
P  
L  
N  
I  
L  
A  
K  
V  
V  
D  
G  
T  
G  
H  
D  
S  
E  
I  
A  
A  
T  
V  
A  
R  
K  
N  
G  
S  
R  
L  
A  
T  
E  
T  
G  
G  
V  
I  
G  
E  
R  
S  
L  
D  
V  
I  
A  
G  
E  
E  
V  
V  
N  
G  
T  
K  
E  
I  
Y  
P  
G  
L  
Y  
V  
C  
G  
M  
A  
A  
S  
A  
V  
S  
G  
T  
P  
R  
M  
G  
P  
I  
F  
G  
G  
M  
L  
M  
S  
G  
K  
V  
A  
E  
E  
I  
I  
A  
K  
L  
K  
K

>Bacteroidetes bacterium ADurb.Bin035  
MEQIVSNGIIDSYFNKL  
QYLS  
V  
D  
A  
I  
V  
G  
G  
G  
P  
S  
G  
I  
V  
A  
A  
Y  
L  
A  
K  
G  
L  
K  
V  
A  
Q  
F  
D  
R  
K  
L  
A  
P  
G  
G  
G  
M  
W  
G  
G  
A  
M  
M  
F  
N  
Q  
I  
V  
I  
Q  
E  
E  
A  
I  
D  
I  
V  
K  
E  
F  
N  
I  
N  
H  
E  
K  
Y  
E  
D  
G  
L  
Y  
V  
M  
D  
S  
V  
E  
S  
T  
S  
A  
L  
I  
Y  
Q  
A  
V  
H  
A  
G  
A  
T  
I  
F  
N  
C  
Y  
S  
V  
E  
D  
V  
F  
H  
N  
D  
A  
V  
A  
G  
V  
V  
V  
N  
W  
A  
P  
V  
I  
R  
E  
G  
M  
H  
V  
D  
P  
L  
N  
I  
L  
A  
K  
V  
V  
D  
G  
T  
G  
H  
D  
S  
E  
I  
A  
A  
T  
V  
A  
R  
K  
N  
G  
S  
R  
L  
A  
T  
E  
T  
G  
G  
V  
I  
G  
E  
R  
S  
L  
D  
V  
I  
A  
G  
E  
E  
V  
V  
N  
G  
T  
K  
E  
I  
Y  
P  
G  
L  
Y  
V  
C  
G  
M  
A  
A  
S  
A  
V  
S  
G  
T  
P  
R  
M  
G  
P  
I  
F  
G  
G  
M  
L  
M  
S  
G  
K  
V  
A  
E  
E  
I  
I  
A  
K  
L  
K  
K

>Brockia lithotrophica  
MFDERV  
ITRAI  
VETYLE  
FRS  
IV  
DLD  
V  
A  
V  
V  
G  
S  
G  
P  
S  
G  
L  
V  
A  
A  
R  
E  
L  
A  
R  
R  
Y  
R  
V  
A  
F  
F  
E  
K  
R  
L  
S  
V  
G  
G  
G  
L  
W  
G  
G  
G  
M  
L  
M  
N  
R  
I  
V  
V  
Q  
E  
A  
R  
P  
I  
L  
E  
E  
F  
G  
V  
R  
T  
R  
E  
Y  
A  
P  
G  
V  
Y  
V  
A  
H  
S  
V  
E  
T  
V  
A  
A  
L  
V  
F  
G  
A  
L  
Q  
A  
G  
Y  
V  
F  
N  
G  
V  
A  
M  
E  
D  
V  
V  
V  
R  
D  
G  
V  
A  
G  
L  
V  
L  
N  
W  
G  
A  
G  
L  
H  
V  
D  
P  
L  
A  
I  
H  
A  
R  
A  
V  
L  
D  
A  
T  
G  
H  
D  
A  
E  
V  
V  
R  
K  
L  
A  
A  
K  
N  
G  
V  
S  
L  
V  
S  
G  
E  
R  
S  
M  
W  
A  
R  
G  
E  
E  
A  
V  
A  
L  
T  
G  
E  
V  
Y  
P  
G  
L  
F  
A  
S  
G  
M  
A  
A  
N  
V  
Y  
G  
G  
H  
R  
M  
G  
P  
I  
F  
G  
G  
M  
L  
L  
S  
G  
K  
V  
A  
E  
E  
I  
I  
A  
K  
L  
K  
K

>Caldimicrobium thiodismutans  
MELQINQATI  
IREGMRD  
LEDFSD  
DV  
LIAGAG  
PSGL  
CAG  
Y  
L  
A  
K  
G  
L  
K  
V  
A  
Q  
F  
D  
R  
K  
L  
A  
P  
G  
G  
G  
M  
W  
G  
G  
A  
M  
M  
F  
N  
Q  
I  
V  
I  
Q  
E  
E  
A  
I  
D  
I  
V  
K  
E  
F  
N  
I  
N  
H  
E  
K  
Y  
E  
D  
G  
L  
Y  
V  
M  
D  
S  
V  
E  
S  
T  
S  
A  
L  
I  
Y  
Q  
A  
V  
H  
A  
G  
A  
T  
I  
F  
N  
C  
Y  
S  
V  
E  
D  
V  
F  
H  
N  
D  
A  
V  
A  
G  
V  
V  
V  
N  
W  
A  
P  
V  
I  
R  
E  
G  
M  
H  
V  
D  
P  
L  
N  
I  
L  
A  
K  
V  
V  
D  
G  
T  
G  
H  
D  
S  
E  
I  
A  
A  
T  
V  
A  
R  
K  
N  
G  
S  
R  
L  
A  
T  
E  
T  
G  
G  
V  
I  
G  
E  
R  
S  
L  
D  
V  
I  
A  
G  
E  
E  
V  
V  
N  
G  
T  
K  
E  
I  
Y  
P  
G  
L  
Y  
V  
C  
G  
M  
A  
A  
S  
A  
V  
S  
G  
T  
P  
R  
M  
G  
P  
I  
F  
G  
G  
M  
L  
M  
S  
G  
K  
V  
A  
E  
E  
I  
I  
A  
K  
L  
K  
K

>candidate division MSBL1 archaeon SCGC-AAA259E19  
MLDEEV  
ITKA  
I  
V  
E  
Y  
M  
R  
T  
F  
L  
E  
N  
T  
D  
V  
E  
A  
A  
L  
G  
G  
A  
P  
N  
L  
V  
A  
A  
K  
K  
L  
A  
E  
K  
G  
V  
K  
T  
A  
V  
Y  
E  
E  
K  
L  
N  
V  
G  
G  
M  
W  
G  
G  
M  
M  
P  
R  
I  
V  
V  
Q  
E  
A  
K  
R  
I  
L  
K  
E  
F  
G  
I  
N  
F  
S  
E  
Y  
E  
K  
G  
Y  
Y  
V  
A  
S  
S  
I  
E  
V  
A  
R  
L  
A  
E  
V  
K  
A  
G  
A  
E  
I  
F  
N  
L  
T  
K  
V  
E  
D  
D  
V  
M  
R  
E  
D  
D  
H  
I  
A  
G  
V  
V  
V  
L  
N  
W  
A  
S  
E  
V  
A  
K  
N  
L  
H  
V  
D  
P  
L  
T  
V  
A  
D  
V  
V  
I  
D  
G  
T  
G  
H  
D  
A  
E  
I  
C  
R  
V  
T  
Q  
R  
K  
I  
P  
D  
A  
D  
L  
Q  
V  
R  
G  
E  
K  
P  
M  
W  
A  
E  
K  
G  
E  
K  
A  
L  
M  
T  
T  
K  
E  
V  
Y  
P  
G  
L  
I  
V  
S  
G  
E  
R  
S  
M  
W  
A  
E  
A  
V  
E  
R  
L  
S  
E  
N  
T  
K  
E  
V  
Y  
P  
G  
L  
F  
A  
S  
G  
M  
A  
A  
N  
V  
Y  
G  
G  
H  
R  
M  
G  
P  
I  
F  
G  
G  
M  
L  
L  
S  
G  
K  
V  
A  
E  
E  
I  
I  
A  
K  
L  
K  
K

>candidate division TA06 bacterium DG\_78  
MIEETIIS  
KA  
I  
V  
E  
Y  
P  
G  
L  
F  
A  
S  
G  
M  
A  
A  
N  
V  
Y  
G  
G  
H  
R  
M  
G  
P  
I  
F  
G  
G  
M  
L  
L  
S  
G  
K  
V  
A  
E  
E  
I  
I  
A  
K  
L  
K  
K

>candidate division WOR\_3 bacterium SM1\_77  
MIEETVIS  
RAI  
IES  
Y  
L  
H  
D  
L  
L  
S  
I  
Q  
S  
D  
V  
I  
I  
G  
G  
A  
P  
G  
S  
G  
L  
C  
A  
G  
Y  
L  
A  
K  
H  
Y  
K  
T  
V  
L  
F  
E  
R  
G  
L  
K  
L  
G  
G  
M  
P  
G  
G  
G  
I  
M  
F  
N  
K  
I  
V  
V  
Q  
D  
E  
G  
K  
A  
T  
L  
D  
E  
F  
G  
I  
T  
Y  
H  
Q  
Y  
E  
P  
G  
Y  
Y  
V  
A  
N  
S  
L  
E  
A  
T  
A  
A  
L  
T  
D  
K  
A  
I  
K  
A  
G  
A  
K  
I  
F  
N  
L  
I  
T  
I  
E  
D  
V  
V  
V  
R  
E  
Q  
V  
V  
G  
V  
V  
N  
W  
S  
S  
V  
E  
M  
A  
K  
L  
H  
I  
D  
P  
I  
S  
F  
E  
S  
K  
F  
V  
I  
D  
A  
T  
G  
H  
P  
S  
E  
I  
V  
H  
I  
V  
E  
K  
K  
S  
G  
G  
K  
L  
L  
T  
P  
S  
G  
R  
I  
E  
G  
K  
S  
M  
W  
A  
E  
A  
V  
E  
R  
L  
S  
E  
N  
T  
K  
E  
V  
Y  
P  
G  
L  
F  
A  
S  
G  
M  
A  
A  
N  
V  
Y  
G  
G  
H  
R  
M  
G  
P  
I  
F  
G  
G  
M  
L  
L  
S  
G  
K  
V  
A  
E  
E  
I  
I  
A  
K  
L  
K  
K

>candidate division Zixibacteria bacterium 4484\_93  
MNFTDI  
K  
I  
S  
K  
A  
I  
I  
D  
Y  
F  
R  
K  
L  
T  
S  
I  
E  
L  
D  
V  
A  
I  
V  
G  
A  
G  
P  
A  
G  
L  
T  
A  
G  
Y  
F  
I  
A  
Q  
Q  
G  
Y  
R  
V  
S  
L  
F  
E  
R  
K  
L  
S  
I  
A  
G  
G  
G  
M  
W  
G  
G  
G  
M  
F  
N  
T  
I  
V  
V  
G  
K  
G  
A  
N  
D  
I  
L  
D  
E  
F  
G  
I  
R  
Y  
R  
R  
T  
E  
D  
S  
L  
Y  
I  
A  
D  
A  
V  
E  
S  
V  
G  
A  
L  
Y  
R  
S  
T  
Q  
A  
G  
L  
R  
I  
F  
N  
L  
I  
T  
V  
E  
D  
L  
L  
V  
K  
D  
G  
V  
E  
G  
L  
V  
I  
N  
W  
S  
P  
V  
E  
M  
A  
S  
L  
H  
I  
D  
P  
L  
T  
V  
Q  
S  
R  
Y  
S  
D  
A  
T  
G  
H  
P  
S  
E  
I  
V  
H  
I  
V  
E  
K  
K  
S  
G  
G  
K  
L  
L  
T  
P  
S  
G  
R  
I  
E  
G  
K  
S  
M  
C  
A  
D  
E  
G  
E  
K  
F  
V  
L  
E  
K  
T  
G  
E  
V  
A  
P  
N  
M  
F  
V  
A  
G  
M  
A  
A  
C  
A  
F  
G  
G  
P  
R  
M  
G  
P  
I  
F  
G  
G  
M  
L  
L  
S  
G  
K  
V  
A  
E  
E  
I  
I  
A  
K  
L  
K  
K

>Candidatus Acetothermia bacterium  
MKIDDVLVSRLLIEEYMAFDLDCLHIDVAIVGAGPAGLTAAYYLAKAGAKVAVYERKLAIGGGMWGGGAMFSRIVIQEEAKQILDQFKITSIAKEGDYYVADAIEAITLLAAGAIQAGAKVFNLIHIEDLLLNRDVRVEGLVLQWSPVEMGGHLHVDPITIGAREVIDATGHCEVVKKLLHKGGVKIDTETGGMLGERPMWAEKGEKMTVQYTKQIYPGLYVAGMAVGAVFGTPRMGPIFGGMLLSGKKAQIIAQTLS

>Candidatus Acidianus copahuensis  
MKVKQVDEGKISKYILKFTFEDWENIIDSVDIIVGAGPSGMAAYYLAKAGLKTFLFERRLSFGGGIGGGAMLFHKIIIESPADEILKELGIRLVKAEDVYAVDTAEFMAKLASSAIDAGAKFIHGTVDDVIFREEPLKVAGVAVEWTATQMSLHVDPIFISAKAVVADATGHDAEVISASRKIPELEISIPGEKSAYSEVAEQQVVDGTGKAPGLYAAAGMAVCEIKGLPRMGPIFGAMVLSGKKVADEIINDIRKS

>Candidatus Altarchaeum sp. CG2\_30\_32\_3053  
MIEETKITELIVRNAVDDFLNNLDVvvVAGGGPAGLTARYLAKAKKKVVLFERKLSIGGGMWGGMMFPRVVLQKGGKEILEECDVCKKSGGLWVADSIECVTKMTAKAIDEGVKIFNLVSIEDVIIRNNEDNKNKNRKNKTCIGVVIINWTAQVOMANLHVDPLSVSKFVVADATGHEASISHLVVKVGNLNTKTDVGLERSWAEKGESDIMGNTGEVYVPGFVSGMAANAVYGSERMGAIFGGMLMSGKKVSELILKKI

>Candidatus Aminicenantes bacterium 4484\_214  
MIDEIVISRAITEAYLKEFLDCLESVDIISGAGPAGLCAALNLAQEYKVVFERTLRPGGGVPGGMMFNKIVIQEEARPLLEELDVTLKPYQENYYVVALELLGALLVKAIKQGVYLFNCISVEDVLIYDKKVGVVINWSAAQAAGLHVDPARTKFVVDATGHAEVAEIVSRKGCKLFTSTGKVTGEKPMWAEEGEKILLNTKEVYVPHLYVCGMAANAVFGGPRMGPIFGGMLLSGQKVAQLIAARLKT

>Candidatus Aramenus sulfurataquae  
MQSIRIKQVDEVKISRILYKTFEDWYSLVSDSDVVVAGGPSGLATAFYTAKAGLKTFFERRLSFGGGIGGGAMNFHKIVIESPADELLREWKVKLVEAEEGVFIVDAAEFMAKLGAAIAIDAGAKVIGHNIVDDVIFRDKPLRVAGVAVEWTSTQMSGHLHVDPFLVSAKAVVADATGHDAEILSVASRKIPELGIVIPGEKSAYSEVAELVNNAGKVAEGLYTTGMAVCEVKSLPRMGPIFGAMVLSGKKVAEDIINDLRNS

>Candidatus Aureabacteria bacterium SURF\_26  
MPLDDLIVSRAIIDYEHTTLVDALNMDSVAIVGGGPAGLIVAGYLYLAKQGYKVSLFERKLSIGGGMWGGGIMFNKIVLQEDALKVLNEFNRVKKYRDNYYVADSVEVGALIYHATQAGLQIFNCMSIEDVKITADAVCGLVNNWTSVELTNMHVDPTFGAKFVIDATGHSCDMANIILKRIGKVLFTPTGDMGEGSMAELAKVVAENSREIYKNLYVTGMAANAIWGSKRMGPIFGGMLLSGKKVADDISARLAQEAVN

>Candidatus Bathyarchaeota archaeon B26-2  
MRIKEVDEAVVTKAILEGSSLKYLHELTEVDVAVVAGAGPAGLTASRYLAKAGLKTFFERRLSFGGGIGGGMQLPMLVVQSPADEFILREVCNLTTYREGVYLANSSLEMALKLAGVSAKAGAHILGTVDDILYRSEENRTRIVGVVVQWSSVIISGLHVDPFLAKAGAVVDCGHDAEVLVASRKIPELNMIQGEKAMWVSESERLIVEKTGEVSPGLYVAGMAVATLNQTPRMGPIFGGMLLSGKKVAEIIERYKQNRTQ

>Candidatus Desantisbacteria bacterium CG2\_30\_40\_21  
MQLDDVVISRAIIIESWNKDLDSLIDVAIVGGGPAGLTCGYLSKGLKVVLFERNLISIGGGMWGGMMYNKCVFQQESLPILNEFGIRTQEYQDGYFTDSLETVTLCGALKAGLKIFNLIGVEDVMIRQEGVTGLVNWTAVTMAKLHDPLTIRAKAIVDSTGHAAEVAGIIVRKIGKKLLTETGEMLGEKPMWAEVGERTIAENTKEIYPGVVFAGMSANAVFGGPRMGPIFGGMLLSGKQAAEIIAARL

>Candidatus Desulfuridis sp.  
MKIDETVVSRAIIERYTQKLLSCLDVDEIVGGGPAGLTAAYHLAKHGKVTTLYERKLSVGGGMWGGAIMMNEIVFQEQARPLFEFGIRINPYSDGYYTASSVECAALTQACQAGANIINLMTVEDVVLHEDRVSGLVLNWTAVDIAGLHVDPATRSKYVIDCTGHMEVANILSKAGVKLVTPSGEPVGEKPMWADVGENQLIGHTIEVYVPGFVAGMAANAVNGGYRGMGAIFGGMVLSGRRAGELILKRLQS

>Candidatus Fermentibacteria bacterium  
MENIVTSAAEDYHRKIQESTVTAHVAIAGGGPSGLVAAEALASRGISVNLYEKNLTPGGGMWGGAMLFNSILIEEEFAETAELGMKLRKYRDSVLLADSVQATAALISRACSAGVRMFNGMAVEDVTVDYDRVNGVVNWAPVMKLGMMVDPLMTCAGAVLDATGHPAEIVTRFAAKNNTIEITVPGEASLNVEMGEKHTVEHTGMVHPGLFVSGMSACATAGGYRGMGPVFGGMLKSLKAAEITEYLNKSQ

>Candidatus Korarchaeota archaeon  
MFVVEESVISSAIIERGSKFLVDLVSVDIIVGAGPSGLVAGRYIAKGLKTCIERRLSFGGGIGGGMLFPRIIVVQEPAQEILEEVGVKLEPYSKGVWIADVAETIKAAGAIDSGARILLGANVEDLIVRNSRVCVVVQSAVTSAGLHVDPFLAFESRAVIDCTGHNAEVVAAIAARKNPELGIRVLGEHSMDAVRAEKEVVELTGEVLPGLWVAGMAAAAVRGGPRMGPIFGGMLLSGKKVAELVSRELEV

>Candidatus Nitrospira inopinata  
MHKPKPAPLRLERDVTRHIAREYYKEFDQRIESDVIIVGAGPSGLLCAHDLAAMGFRTLIVEQSLALGGFWHGGYLMNKATICEPANEILEEVGVPCKRIACDCGMYMVDPPHATGALVAAAYRAGAKILNLTRVVDLILRQDGLLEGIVVVNNNTAEMAGHDVIIHVDPIALESKIVVADATGHDAVVLVELLHKRNLYKPVPGNGAMWVSRSEEEVMDRTGEVYVNCFVIGLAVAAYVTPRMGPAGFSMILLSGRYGAQLIKKKLKQE

>Candidatus Nitrospira nitrificans  
MAKPRPAPLRLERDITRHIAREYYKEFDQRIESDVIIVGAGPSGLICAHDLAAMGFRTLIEQSLALGGFWGGYLMNKATICEPANEILEEVGVPCKKIKECEGMYMDPDPHATGALIAAAKGGAKIMNLTRVVDLILRNGGLLEGIVVVNNNTAEMAGHDVIIHVDPIALESKIVVADATGHDAVVLVELLHKRNLYNKVPNGAMWVSRSEEEVMDRTGEVYVNCFVIGLAVAAYVTPRMGPAGFSMILLSGRYGAELIKKKLKQE

>Candidatus Nitrospira nitrosa  
MTKPRPAPLRLERDITRQIAREYYKEFDQRIESDVIIVGAGPSGLICAHDLAAMGFRTLIEQSLALGGFWHGGYLMNKATICEPANEILEEVGVPCKKIRECEGMYMDPDPHATGALIASAYKAGAKVLNTRVVDLILRRDGILEGVVVNNNTAEMAGHDVIIHVDPIALESKIVVADATGHDAIVVVELLHKRNLYQKIPNGAMWVSRSEEEVMDRTGEVYVNCFVIGLAVAAYVTPRMGPAGFSMILLSGRYGAELIKKKLKQE

>Candidatus Syntrophoarchaeum butanivorans  
MDEVTISKAITESYMKDLIDSMLVLDVVVGAGPAGLLAAYNLAREGVKAVAFERRLSVGGMWGGMMFSRIVVQDAGREILDEIGVRCSEYEPEGYYY  
TADAEAVTTITSEVIRAGARIFNLMSEVEDVVVRDDRIHGVINWSAELSKLHVDPMTVIADYVIDATGHAAEVARIVEQKLGGALTVERPM  
WAEAGEAAVENTKEIYPGLIVAGMAANAVLGS prmGPVFGM LLSGRKAAELVLSRL

>Chloroflexi bacterium RBG\_13\_51\_52  
MVKFSPVG EVVITRAIVEFAKEFN EYVESDCIIIGGGPSGLVAGRDIARAGKKVIIERNNYLGGFWSGGYLMPKVTVRPGEKILDELGV PYKT  
VAKGLVVCDA PHACAALIAAACAAAGVKIFNMTMLEDLVVKDGRVCGAVINWSPIASLPRQVAALDPVAIEAKVVIDATGH DATVVA KLEKRNL IKMK  
GEGAMWIEKSEDLIVEHTGECFPGLIVTGMAGVAYGLPRMGPTFGSMFLS GEVAAKVALEKMK

>Clostridium drakei  
MYLEDTKISKAIIDTYKDKLEDILHSDVII VGGGPGSLV GASYLAKAGIKTTLERNSL IGGGMWGGMMMNQIVIQESAKSILDEFNIGKKYEE  
YYTADSIECVSALTLSASQSGARILNSISVEDVIVKDKCISGLVINWAAVEKTRMPIDPIMIESKYVLDATGDASVNVNL VTRMGNVLNTPNGTLE  
GEKPMWADRGEQVIKNTREVY PGLYVSGMAANATFGGQRMGP IFGGMLISGQKVAQELIKKIKNC

>Clostridium ragsdalei P11  
MYLEDTKISKAIIDTYKNKLEDVLHSDVII VGGGPGSLV GASYLAKAGIKTTLERNSL IGGGMWGGMMMNQIVIQESAKSILDEFNIGKKYEE  
YYTADSIECVSGLTFNAAQAGARILNLITVEDVIVKDKCISGLVINWAAVEKTRMPIDPIMIESKYVLDATGDASVNVNL VTRMGNVLNTPNGTLE  
GEKPMWADRGEQVIKNTREVY PGLYVSGMAANATFGGQRMGP IFGGMLISGQKVAQELIKKIKNC

>Clostridium sp. JN500901  
MYLEDTKISKAIIDTYKDKLEDILYSDVII VGGGPGSLV GASYLAKAGIKTTLERNSL IGGGMWGGMMMNQIVIQESARSILDDFNVN YKKYEE  
YYTVD SIECVSALT KAVKAGAKILNLISVEDVIEKDNCIAGLVINWAAVEKTRMPIDPIMIESKYVLDATGDASVNVNL VTRMGNVLNTE NGSFV  
GEKTMWADRGEQVIKNTREVY PGLYVSGMAANATFGGQRMGP IFGGMLISGQKVAQELIKKIKNC

>Coxiella sp. DG\_40  
MEQITTLGIVDSYYQKLKDNL FIDVAI VGGGPSALVAAYYLAKLQKVAIFERKLAPGGMWGGMMFNQIVVQSEALSILDEFKISYALFKDNYYL  
VDSIESTASLIYHTIHAGAKVFNCYSVEDIVLKNNKVGIVVNWTGVDQGLHVDPVVAKC VIEATGH SCEVAKVLA KNGIKLHTETGGVVG EK  
SLAMEQAERSTIENTKEIYPGLYVCGMAANGVSGDFRMGP IFGGMLMSGKKVAEIIVKDIT

>Dehalococcoidia bacterium  
MPLFHPVTEGEITRAIVNSFLRQFEEYVSSDVI I VGGGPGSLV MAGREL GKGAGL KVIVIERN NYLGGFWAGGYFMNKLT LREPAQEVL DELGV PYSR  
AGEGLYVADAPHACSKLIGAAADSGVKFFNLT LEDLVVREDKRVAGAVINWSPIA LPREIA ALDPVPLETKVII DATGDASVARKLERRGMLK  
AGE GALWIEESEEAVV EHTGEVY PGLV VTGM AVASVY GLPR MGPTFGM LLSGK RAAEVAL AAVL TDSR

>Desulfacinum hydrothermale DSM 13146  
MALDERIITRAIMDRYIAKLKEAIDL DVAI VGAGP SGLVAGM LLAEGKKVALFERKL SVGGMWGGMLFNEIVVQEEAKTILDQVGIRAQHYTDG  
YYTADAVE SVSTLT SRVKGARIFNCVS EDVMMR PERIMGLV INWSA VEMAGL HVDP LTVRCVVV DATGDTEVVKVVERK VPGSLSTPSGKRA  
GERSMWAEEAERLTLENTCQVY PGLYVAGMAANATFGGPRMGP IFGGMLLSGQKVARLILEQLQS

>Desulfarculus sp.  
MLEEVITTRAIRYLGKLDQSLEDDAAIVGGGPAGL VAGKKLAQAGYKTA LFERKL SVGGMWGGMLN EIVVQQEARRI LEFGVPSSEFAPGY  
YTADSVLATSTLCVAAKAGLTIFNLV SVEDVVIR AQRV TSLV INWSA VQ MAGL HVDP LT IKARV VIDATGDSEVLH VIARKV DAELL TASGKVMG  
ERSLWAEQAESDTLANTREAFPGVYTAGMCANAVFGSYRMGPVFGM LLSGEKAAA EVAARLAAGE

>Desulfatibacillum aliphaticivorans  
MEERITS AIVRTYFEKLQNFLEV DLAIVGAGPSGLVAAA ALAKEGKKV AIFERLLAPGGVWGGMLFNEIVVQEEALHILDDFNISYKSAGD GLYT  
ADSVEVASGLIFGAKKAGVMINNAV SVEDVVC REGRICGV VNWT PVER LGMV HVDPLV VMSKAV LDGT GHPGEITDLATRKAGIKIDTPTGKIMGEK  
PMW MELGEASTVENTKCL YPGLYVSGMA ANN ASGGFRMGP IFGGMF MSGRK VAKMILEDIDG

>Desulfobacca acetoxidans (strain ATCC 700848 / DSM 11109 / ASRB2)  
MGLDEIIISRAITERFMEKFLDNLELDVAI VGGGV SGLVAGWRLA QKGRKAAIFERKL SVGGMWGGMMFNEIVVQEEAKHLLDELGITSR PYDRG  
YYTADAEIESTTTLASQAMKAGVKIFNL I HVEDV MVREN RIDGLV ILWTA VN MAGL HVDP LTIRAKH VIDCTGH DVEV I KIFLRKNQ PASL KTETGGI  
MGERSMWAEVGEAKTVEY TSEVY PGLW VAGMTATGTL GTFRMGP IFGGM LLSGEKA ANLIDERLKKG

>Desulfobacteraceae bacterium  
MQLDDVAISKSIL DAYFEKL LARL DV DVALVGAGPANLVAGY YLGKSGFKAVV FESKL APGGMWGGMMFNEIVLQDDAVHIAEELGIHCNP GGDG  
YYTMD SVEATSTIISRCV RGTIVFNL KIVEDVLF RQDDRQ PRV SGLV INWSPV EKLYV DPLS IRASFV VDGT GH PADIC RTVARKMDV KLN TKT  
GNVVGEMPLWAEKGEQFTVNTA E VFPGLYVAGMAANAFGGPRMGP IFGGMLRGK VAEILA EKLRSD

>Desulfobacterium sp. 4572\_20  
MAINEVVISKA IIDRSQKF MEY TEVDTAIVGAGPSGLIAAYFLARAGQKVALFERQLS IGGGMWGGMMFNEIVVQTQ GKELLEMFG ISAREYEPG  
YYTADAVECVTTIC SNAV KAGAKI FN CMSV EDVSI REDRV MGLV LT WSAVEAARMHV DP LTIA AKYV IDATGDTEVIRLIEKKADIALQ TETG KIM  
GERSMWADKAEQLTIENTKEI CPGV FVSGMAANAA FG GPRMGP IFGGM LLSGK VAE LIMA EGSAFET SAEDFD SWFNRNQ AIFESELLAQQQFIY  
SL

>Desulfocarbo indianensis  
MLDEITITRAIIDRYFEKL NRNLEDDAAIVGGGP SGLIAGYKLAKAGYRVAMFERKL S IGGGMWGGMMN EIVVQEEAKRILDEVQVPTREFQPGY  
YTADSVLCTSTLCQA AKAGLTIFNLV SVEDV MVRE QRV VGLV INWTA VEMAGL HVDP LTIRAKYTI DATGHAAEV MHVIARKVDAKLFTDDGKVAG  
ERSLWAEVAETNTV NNTREAFGGVFTAGMCCNATFGSYRMGPVFGM LLSGEKAQALV AERLQAEK

>Desulfococcus oleovorans (strain DSM 6200 / Hxd3)  
MELNEVTISRAIIDRFYEKLIANLEVDVAVGPPSGLVAAWRLARAGRKVALFERKLSIGGGMWGGAMLFNEIVVQKSALHVLDAMEIGYRLYAED  
YYTADAVEAISTLTSQAAKAGVAIFNCVTVEDVMIRPDRIVGLVLNWPVEMAGLHVDPDLMRASFVIDATGHATEVVHVVAKVPGTLRTDSGKIE  
GEKSMWSdraesitlentrevypgLyVAGMAGNATFGGPRMGAIFGGMLLSGEKVAAEILERLE

>Desulfofundulus australicus DSM 11792  
MMHLEDVVISKAIIISRYQEELLEALESDVAVVGGPSGLVAAYYLARANKKVVLFERKLSIGGGMWGGMMFNQIVIODEALPLLEEFKISYRVFEE  
GYTASSVEAVAALTLGAVRAGAKIFNLISVEDIMVRDNRVAGLVINWTPVDSLHVDPDLMRASFVIDATGHATEVVHVVAKVPGTLRTDSGKIE  
EGEKPWAARGEMATVANTREVYPGLIVAGMAANAACVCGGHRMGPVFGMLLSGQRAARIILEGDKT

>Desulfofustis glycolicus DSM 9705  
MLNEVTISTAIINRYMTKLTSALDDVAIVGGPSGLVAGYYLAKAGRKVALFERKLSIGGGIWGGMMFNEIVVQEAGAAVLAEGFLAGSPFEPGY  
YTLDHSVTTATLKHAMAAGLLIFNLIGVDDVVKDERVAGLVINWGAUTLGWHIDPLTFARYVLDATGHDAEIASVLRKMGRNLNTETGGLVG  
EKSMWAERAERETVNTREVYPGLIVAGMAANAACVCGYRMGPVFGMLLSGKRAAESILEGLA

>Desulfonatronospira thiodismutans ASO3-1  
MALEIIISRAIETYTKLMDSELVDVAICGAGPSGMVAAYYLASAGKKTAVFERNLAPGGGMWGGMMFNEVVVQEEAREILDELDIKSVEYTPG  
YYTADSVEAVCTLGSKAAKAGARFFNLVCIEDVMIRENRTGLVINWNSAVESAGLHVDPDLMRASFVIDATGHDAEIASVLRKMGRNLNTETGGLVG  
GEKSMWAKEAEEHTIENTTEAFPGVYVCGMSANATFGSFRMGPVFGMLLSGKVAQEIIINKAK

>Desulfonauticus sp. 38\_4375  
MSLDEKIISEAIIISKYFEDFKRCLNLDVAIVGGPSGLTAAYYLAKEGFKVALFERKLSIGGGMWGGMTFNYIVVQEKGQILEEMDIICEEYKPG  
YYVDAVLATTLSKACKAGAKIFNCMSVEDVVIREEGVKVRAGLVNVNSPVEIAGLHVDPDLMRASFVIDATGHDAEIASVLRKMGRNLNTETGGLVG  
GGIEGEKSMWAEVAAEENTLKNTRAEFPGIYVCGMAANACFGSYRMGPVFGMLLSGVKAEEISTRLKEGK

>Desulfurudis audaxviator (strain MP104C)  
MKLDENIIISRAIIESYVTRLSCLEVDTVIEVGGPSGLTAAYYLARAGLKTTVYERKLSIGGGMWGGMTWNSIVVQESAKSILEDAGVALSEFKPG  
YYTASSVECAVAILTGACRAGANIMNLLTVEDVVLHNRRVSGLVLNWSAVEISGLHVDPDIATRSKFVVDATGHDSVVGVLARKAGVQLDTPSGKVG  
GEKPMWADLGEAQIMENTSEIFPGLYVVGMAANAVHGGYRMGAUTFGGMVLSGRRAEMIIDRLKV

>Desulfovibrionaceae bacterium CG1\_02\_65\_16  
MIIDERIVSEIAIASAYFEKFQCLLDVAIVGGPSGLTAAWKLAKAGRKVALFERKLSIGGGMWGGMTWNSIVVQESAKSILEDAGVALSEFKPG  
YFTADSVATAALAYQATHAGAHVFNCMSVEDVVIREVEGVKRVIGLVNVNSPVEIARLHVDPDLMRASFVIDATGHDAEIASVLRKMGRNLNTETGGLVG  
GGIEGEQSMWADVAEANTVRYTREVFPGVWAGMAANAACFGSYRMGPVFGMLLSGVKAETIDALL

>Desulfovibrionales bacterium GWA2\_65\_9  
MIIDERIVTEAIAASAYFEKFQCLLDVAIVGGPSGLTAAWKLAEAGRKVALFERKLSIGGGMWGGMTWNYIVVQEEAKGILEEAGCAMSEYKPG  
YFIADSVATAALAYRATKAGAHVFNCMSVEDVVIREIDGEKRVGMIGLVNVNSPVEIARLHVDPDLMRASFVIDATGHDAEIASVLRKMGRNLNTETGGLVG  
GGIEGEQSMWADVAEANTVRHTREVFPGVWAGMAANATGGSYRMGPVFGMLLSGVKAEEINARL

>Desulfurella amilsii  
MALDERIISRAIERYFQKLLANIDCDCAIVGAGPAGLVCYELVKNGLKVTLFDKRLSVGGGMWGGAMMFNEIVVQEEGKILDEFDIKCSLFEPN  
YYTDSIEAITLISKTVKAGVKIFNGIEIEDVVLKKVGDYKVGVVINWTTVNMAHPVDPPIVISSFTVDAUTGHDAEIASVLRKGKVLNTDS  
GAVIGEKPMWAQIGEQDTVNHTKEIFSGLYVCGMAANAVSGAHRMGPVFGMLLSGKCAQLILEKWSRK

>Desulfurella multipotens  
MALDERIISKAIERYSQKLLSQLDCDCVIVGGPGAGLICGYELAKNGLKVTLFDKRLSVGGGMWGGAMMFNEIVVQEDGKAILDEFDIKTVLYEPN  
YYTADSIEAITLISKTVKAGVKIFNGIEIEDVVLKKVGDYKVGVVINWTTVNMAHPVDPPIVISSFTVDAUTGHDAEIASVLRKGKVLNTDS  
GGVGEKPMWADVGEQDTVNHTKEIFSGLYVCGMAANACSGAHRMGPVFGMLLSGKCAQLILEKWSRK

>Desulfurobacterium atlanticum  
MELSEVVISRAIVERFMNKLLSNLKVDAIVGGPSGLVAAYYLAKEGFKVSLFERKLSIGGGMWGGAMLFNEIVVQEMGREILDEFDVGYEKFQEG  
YYTDSVEAVTTIASKAVKAGAKVFNGBTVEDVVLKKENGDYRVCGLVINWTPVETGMHVDPDLMRASFVIDATGHDAEIASVLRKGKVLNTDS  
GCVVGEKPLWASVGEEDTVKNSREVFPGIYVCGMAANAVCGSHRMGPVFGMMSGKKIAKEIAERLKHNVEE

>Desulfurobacterium indicum  
MENLSEVKISKAIIERFTEKLLSNLEVDVAIVGGPSGLVAAYYLAKEGLKVSFERKLSIGGGMWAGAMFFNEIVVQEMGREILDEFDSVSYRKYDE  
YYTADAVEAVTTIASKAKAGAKIFNGVTAEDVVLKKVNGQYRVCGLVINWSTVDMTGLMVDPLVVTNSYVIDATGHDAEIASVLRKGKVLNTDS  
TGCVGEKPLWASVGEEDTVKNSREVFPGIYVCGMAANATCGSHRMGPVFGMMSGKKIAKEIAERLKHNVEE

>Desulfurococcales archaeon ex4484\_42  
MVKELESRVTELIVKHASRDWAELASTDViVAGPSGLTAALKYLAEGFKVSLFERKLSIGGGMWAGAMFFNEIVVQEMGREILDEFDSVSYRKYDE  
LYVVDASELMAKLAVGALNAGAKITHGVTVEDViFRTNPLRITGVAIQWSAVPLANLHVDPDLMRASFVIDATGHDAEIASVLRKGKVLNTDS  
KSAYSELGERLVEKTKVPGLYVTGMAAALNNLPRMGPVFGMLLSGKVKANEVLKDLRT

>Desulfuromonas sp. SDB  
MKDINITNHIIAEFYKDIQDRVSSDVIIIGAGPSGLVASYLLAQDNFKVTVFEPKRNPQPGGGIWGGMMFNQLVLPDDLQDFLNQMSIKFKLHPDNL  
SVDSVFSSALLYHATEVGKIFNNIGVEDLLVVDDMVRGVVVINWNDVIKNKIPIDPLTFEAKVVDSTGHPADGVEKLARRGLVEISQEFPMNADV  
AEKFVVEATGQLYPGLYVCGMAATAAKGGPRMGPVFGMMSGKKIAKEIAERLKHNVEE

>Dethiosulfatarculus sandiegensis  
MMLDEVTITRAIIDRYMEKLNHANLDDVAIVGGGPGLVAGYYLAKKGYNVAMFERKLSIGGGMWGGMMNEIVVQEEAKRILDEFGIPCREYVEGYYTADSVSTSLTSKATLAGLSVFNLITVEDVMVRDNRVNGLVINWSPVEMAGLHVDPDTLRARYTIDATGHPAEVINVISKVDALKSTDGKVIGERSLWAEVAESTTIENTKEAPGVYTAGMCANAVFGAHRMGPVFGMLLSGEKVAQVLDRLQEED

>Dethiosulfovibrio peptidovorans DSM 11002  
MELDERVISKAISRSFFERLTDHLENDVVIVGGGPAGLVAAGYVLADAGVKVSLFDRLSLGGGMWGGMLFNEIVVQSEGARILDDLGVSILREFEPGYYTAGSVEAVSTLISSAVRAGTVFNGMVAEDVVMREDRVRIVGLVINWSTVETSGLLVDPLAVRSDFIIDATGHDSNTSTVEKKVPGRLLTETGKVEGEKSLWCERAEKLTVDNTKEVYPGLVAGMSANAVFGGPRMGPIFGGMLLSGEKAKEILLRLNGKRV

>Dissulfuribacter thermophilus  
MREIDITKAIIDKHIEELNKCLCSVVIVGAGPSGLVAGSILAQKGYTITIFEKRLAPGGIWGGGMGFKYVIQKEALDIVEFNIPYEKYSDDLYAVDAINFASGLILEAGKRGVHIFNLIAVEDLLVREGRVQGVVINNTFAKMNQFPIDPLTIEAKAVVDATGHEHEVVKTLSQNDVTLNTPTGKPLGERSLFAETAEKAVVNTKEVYPGLYVCGMATAAVYGGYRMGPIFGGMILMSGKKLAGLLEEALK

>Elusimicrobia bacterium CG\_4\_10\_14\_0\_8\_um\_filter\_37\_32  
MKLDIVISKAIMETFTKDFDVYLEVDAIVGGGPAGLTAAGYVLAKKGKVVLFERKLSIGGGMWGGMMYNKCVFQEDAKKILDEFGVTHKYQEGYYVTDSLTVSILCSKAIAKGLKIFNLISVEDVMIRKEKITGLVWNWSAVQLAHKLVDPMTIRAKYVIDATGHDAEVVKIVVRKIGKKLYTKGDMLGEKPMWAEVGEKDIKNTKECPGLYICGMASNAVFGGPRMGPIFGGMILSGKRISGLV

>Euryarchaeota archaeon ADurb.Bin165  
MTLDEVTISRAIITDHNLTVQYMEMDVAVIGAGPSGLVCATILAEGKLKVGLIEKKLSVGGGMWGGMMFPRIVVQQGAKRLLDRFGIRSSEFSPGYYTARSIEAVAKLAAAASADVEFFNLTTEDVMVKGDGLLSGLVINWQPVEATGLHVDPMTVRCRMV рат GHDAIIAHYVSKCGGLEIKGETMWADNAEAAVVAHTKEVYPGLYVCGMAANAVMGGNRMGPVFGMLLSGESAAEQLSRF

>Ferroglobus placidus (strain DSM 10642 / AEDI12DO)  
MPFSEKNITRVIREAAKEWEEISETDVVVGAGPAGLTAAYHLADFGFDVVVFERRLSFGGGIGGGMLFHKIVVKEAKEIAEEFGIKTREVEDGLYVIDAEMLAJKLSSAGAIDSGAKVILGTVDDVYRPEPLRISGVLVQWSAVQIAGLHVDPMLIESKAVV рат GHDAEVV рат VARKIPELEIYVAGEKSAYSELSEKLVVEKTGKVVGDGLYVAGMAVSAYVGLPRMGPIFGGMILSGRKVAEQIMFDLKK

>Fervidicoccus fontis  
MSENLEFKITKLILEHSMKDYLIEFADSIIVGAGPSGMAAKYLADKKLKVLVLERKLSFGGGIGGGGNLMHKIVIKSDALKIIKDFIEYKKTEFEDLYTLDASELISKLATGAINSGAKILFGYSEDIVREKPLRVSGVVWKSAIDLQLAQHVDPPIFTGKAILDАТ GHDAELIKILAKKNPSFAINVКNESSAHAELGEKQVVEFSKGKVDGLYAGMSVATLHGLYRМGPIFSGMLISGKKVAELISKELGK

>Fervidobacterium changbaicum  
MGKDLTISKLIVENFFEKLSNALEVDAVIAAGCGPSALTSLSELSKKGYKVAIFEAKNEPGGIWGGGMFNEVVESELEGYLKELGIRFFKDFEFITVDVHLASALLYHTTLAGTMIFNNVFEDLVMYDRVSGVVINWPTLREKLHVDPISIVSKFTV рат GHDAV рат VLSKRGIISSIGSTEASYNFGIVGYEFPMDAENGERFVENTREIYGLYIVGMAAVSVGAGPRMGPIFGGMIMSGLRAELISNELRKMGGSDDER

>Gemmatimonas sp. SG8\_17  
MRGRGGGVTDSEQQITRAIITAYHEKLWQVVGDVVVVGAGPSGLVAAATDLARRGLKTVLEKRLSPGGIWGGAMAMNEVVV рат QDAALPLLAЕFSVSERSVKGGLHVINAVELASALSILKAVQTGAVIINLTVAEDVCVHRGRTGVVANRTNLAEALPVDPVSFEAKAVL рат GHDAALVQMLQRGLLKGЛTEMQGEGPMDAAGGESFVVDKVTEVYPGLWVSGMAVVATLGGPRMGPIFGGMILSGKRAADLISDTLSGE

>Geoglobus acetivorans  
MSYSERNITRIIVREAAKDWEISDTDVVIVGAGPAGLTAAYLREFGFДVVVFERRLSFGGGIGGGMLFHKIVIEEEAKEIAEGFGMKLKEVESGLYSVDSSDFLAKLSYSAVESGAKVLLGVTVDDVFRPDPLRISGVLVQWSAVQISGLHVDPMLIESRAVV рат GHDAEVISIAARKIPELEIFIHKESAYSEMSEKLVVEKTGKADGLYAGMAVAHVGLPRMGPIFGGMILMSGKKVAEQIMFDLKK

>groundwater metagenome  
MIDETKITELIVRSAVDDFLGNLKVDVVVVGPGAGLTTARYLAKAKRVLFERKLSIGGGMWGGMMFPRV рат LQKGKEIЛЕЕCNVRYKKFDDLWVADSIECVTKMTAKAIDEGVKIFNLISIEDVIIRSQSNKNNKEGKTKICGVV рат GHDAV рат VLSKDFVV рат GHHEASICHЛVVKVGНLNTKTDV рат GERSMWAЕKGESDIMNNTKEVYPGLVSGMAANAVGSERMGAIFGGMILMSGKKVSELILEKEK

>Hadesarchaea archaeon DG-33-1  
MGGIEDTEITAIIKRFMRDFEDVTNLDAVIAAGAGPSGITAASFLASGGAKVAVFERNLHVGGGMWGGGILFSRVVIQEAKVMLEEVGVKLKPTAAGYYTADSVЕAVTKSTAAVDAGARVMVGLTVEDVMIREKDRVAGIAVNWKABELAGLHVDPVGISAKIVIDATGHДАMIARIVQRKVPNAKFPTSTGGV рат GEKPVWAЕVGETEIVNNTREIYGLV рат GHDAV рат VLSKDFVV рат GHHEASICHЛVVKVGНLNTKTDV рат GERSMWAЕKGESDIMNNTKEVYPGLVSGMAANAVGSERMGAIFGGMILMSGKKVSELILEKEK

>hydrocarbon metagenome  
MWYLVELDERVISRAIIAVQMEKMLRYTDMDAVIAVGGGPAGLTAASFLGAEGFSVALIEKKLSVGGMWGGMMFPRIVVQEGRQLLDHFAIRYTRYEЕGYYVASSVEAVAKLTAACDAGVEFFTЛV рат GHDAV рат VLSKDFVV рат GHHEASICHЛVVKVGНLNTKTDV рат GERSMWAЕKGESDIMNNTKEVYPGLVSGMAANAVGSERMGAIFGGMILMSGKKVSELILEKEK

>Hyperthermus butylicus (strain DSM 5456 / JCM 9403 / PLM1-5)  
MVNAVQAPEHSQLPHNVTSLREGALAALIIRKTAЕKLTISITSVDVAIAGAGPAGLTAAWLLAEKGLRVVVVEHSLGVGGMRRGSMLMPVGLVEDGLPAELLRRAGARLDRVADGLYAVDPTEAVV рат GHDAV рат VLSKDFVV рат GHHEASICHЛVVKVGНLNTKTDV рат GERSMWAЕKGESDIMNNTKEVYPGLVSGMAANAVGSERMGAIFGGMILMSGKKVSELILEKEK

>Ignicoccus islandicus DSM 13165

MIDEKGKVTSIIIIEESSKELSQMAKGVDVIVGAGPAGLTASHYLAKAGLKVLIERRVSLGGGSGGGSLFKVVVVEDVELEGYNPEIAEELGVPLKKVDDNLYTTDAALVAKLSNASVSAGAKIVLGMHVEDLIYRIEEGVTKVGVALWSPYIYSLGLHVDPIFFKAKAVVDAKGHDAEILKIAASKKLPNVNFEVGREYGAWIDEAEKLVVKYTGKVLLEGLYAAGMSVASFYRLPRMGPVFGGMLASGKKVAEKIIGDLEVS

>Ignisphaera aggregans (strain DSM 17230 / JCM 13409 / AQ1.S1)

MKELELRISRAILRNSVRELIEYSDVVIVGAGPSGLTAARYLAMNGFRVVLERRLSFGGGIGGGGMLFHKIVVSSEALPILNDFDIKYYRDEE  
DLYMIDSSELMAKLAVGAINAGAKIFHGHIHVEDVIYRENPLRITGVVIQWSAVVMSGLHVDPFLITSRAVVDATGHDAEVLQIYRSKIPVEVGISLPG  
ESSAYSELSEKIVVEKTGMVIPGLYVAGMVAALYKLPRMGPFSMILLSGKRVAEIIANDLKKK

>Korarchaeum cryptofilum (strain OPF8)

MESLESRSKAIWESTYKDLDIIDS D V V I V G A G P S G L T A A S Y L A K S G F K T T V I E R R L S F G G G I G G G M Q L H K V V V D G R A L K V L E D F K V R Y S Y L E K Y  
DLYVLDLSELMAKLASGAIDSGAKLIHGLTVEDLIVREDPFRVEGVVVQWSSVLLAGLHVDPFLIHSRVVVDATGHDAEVIRILERKNPNSLGKVP  
ERSAYSELSVVERTGKVVEGLYVTGMAVAALNQLHRMGPFGMMLSGRKVAEEIIRDLS

>Labilibaculum filiforme

MEQIVSAGIVDSYFKKLKENLSVDVAIVGGGPGLVASYYLAKKGKFVALYESKLAPGGGMWGGAMMFNEIIIVQKDALHILNELGVSYQHYQEDYYT  
LDSVHATSALIYHATQAGVKIFNCFSIEDVVVFQNNDKVCVGVLNWSPVRREGLHVDPVLVMAKAVVDGTHDCCDIARTLERKNDVKNLNTKTGKVMGEC  
SLSIDEAERTTVENTKEIYPGLYVSGMASNGVSGGFRMGPFGGMLRSGEKLAGLIAENSLK

>Latescibacteria bacterium DG\_63

MKLLDDVEISKAIIIESFYAKLDSLMDCVAVVGGGPAGLTAAYYLAHKGRKVFLFERKLSIGGGMWGGGIMFNEIVVQRDGKKILDEFGVRTTLVKEGYFCADSVEAVSTICSKASQAGARIFNLFSVEDVMMTTEERVTLGLVINWSAELNSLHVDPISIKAEHVIDATGHAAEVAHITQTKSGSKLLPTGTGERPMCAEVAKSILENTKEIFPGVIAAGMCCNAVGAPRMPGPIFGGMLMSGKKAAELIIDKPAAPKRRCLPDE

>Lentisphaerae bacterium ADurb.Bin082

MAMENIITTAIIRQFADKLSAGTDLDFVGGGSPALVAAAKLAKKGLKTAIFEKSLAPGGGVWGGMLFNEIVVQENVLGILEQIGISYQAVPDAK  
GYYTVDSEMASGLIFNAVKAGAKIFNAMSVEDIVFKEGRVNLVINWAPVRKLAMPVDPLTVIAKAVV рат GHPCЕIIRIACEKAQVKIATETGGV  
LGERPMWVQHGEQQTVDSTAЕYYPGLFACGMSATNVТGGYRMGPIFGGMILSGLKAADLIASKLASLEKR

>Metallosphaera yellowstonensis MK1

MEIRQVDEVKITKYILKATFEDWMDIAENDVVIVGAGPSGLSAAYYLAKKGLKTTVFERRLSFGGGIGGGAMLFHKIVIESPADQVLREMNRNLQRVEEGVYIVDSSEMAKLASSAIDAGAKIVHGTVDDVIFRENPLRVTGVAVEWTATQMASLHVDPFLFIHAKAVV рат GHDAEVISVAARKIPELGIAI PGEKSAYSEVAEKLTVDTNТGEVAPGLYAGMAVTEVKGLPRMGPIFGAMVLSGKKVAEDIASTLLMKARNT

>Methanobacteriales archaeon H

MELDDITISRAIVEEFMNDFMDDYMDIDVAIGGGGPAGLTAGYYLAKAGLKVALLYERKLSIGGGMWGGGMFMNKIVVQEEGKRILDEFGIQSKKYQEN  
YYVSDSVEATSTLC SKATQAGLKIFNLM SIEDVMIRGDDISGLVNWSSVEMGLHVDP LSIRSKAVIDATGHPCEVVKVVQNKGPKLNTPTGEII  
GEKSMWAEVGEPAIMENTREVYPNLYVAGMAANAVYGA PRMGPVFGGMLLSGEKIANMLIEKLK

>Methanobacterium subterraneum

MKLDDDIIVSKGIVAGYMEELLDYMEMDVAIGGGPGSLTAGYYLAKAGLKVLFKEKKLSMGGMWGGGMFNKIVVQEEGKRILDEMGIRNQEYEEG  
YYLADSVESASTICSKACQAGLKVFNLMEIEDVMIKGEGVEGLVINWSPVEMAGLHVDPITVGARAVIDATGHPCEVVVKLERKMEAPLKTETGKIM  
GEKSMWADVAEQNIMGNVGEIYPGMYVTGMAANAVHGSPRMGPFIGGMLLSGEKVAEMLIEKLK

>Methanobrevibacter woesei

MKKLDDITVSKAIQEYMDFLDVTMDVAIGGGPGSGVTAGYYLAKAGYKVALFERKLSIGGGMWGGGMMFNKVVVQEEGKRILDEFGIKSKKFED  
NYYTVDSSIECTSTLCSKATQAGLKIFNLMSIEDLMLVRENGINGIVLNWSSVEMSGLHIDPLTVRAKAVIDATGHPTEITKIVEQKMGANLKTETGK  
MGEKSMWADRAEGKILDNVTEVYPLGLWTGMAANAVHGSQRMGPIFGGMLLSGEYVAQKIEKLENE

>Methanocalculus sp. 52\_23

MQLDEVTISRAILETHAEISSRYLDLIAIVGGPGSLVCAALAAEDGRKVAVIEKKLSVGGMWGGGMFTPRIVVQEEGKRLLDQFGIRSRVYKPGYHVASSVESVAKLTAACDAGAEFFNLTSEVEDVVIKEKGDRVSGLVITTPVEMTGLHVDPITLAAKVTVDATGHDAVVAHCVLRKGDDITIHGESFMWAERAETNNIINHTREIFPGLIACGMAANAVAGEARMPVFGGMLLSGEHAAVLAREISERV

>Methanocella conradii (strain DSM 24694 / JCM 17849 / CGMCC 1.5162 / HZ254)

MELDETLISRAIIDDRLTSLSDYVSDVGIVGGPGSLVCATYLARAGVKVAVFERKLSVGGGMFPRIVVQQEATRILDDFGIRYREYRP  
YYIAGSIEAVGRLTSAAAGAGAEIFNLMHSVDMIRENVGVGLVINWSAVIDAGLHVDPITVRTRVVVATGHPAEVCRIVERKVSGGAFKVPGEQ  
SMWADRGERALISTTKEVYPLGLVVAAGMAANAVAGGPRMGPFGMLLSGEIAARIVKEKLGS

>Methanococcoides burtonii (strain DSM 6242 / NBRC 107633 / OCM 468 / ACE-M)

MKLDEVTISRAIIIEFSKVFLDYTDVVALVGGGPANLVAALKYAEAGLKTVIYEKKLVAGGGMWAGGMMFPRIVQEDALHILDEFGISYHEYENG  
YYVANSIESVVGKLISGATSAAGEIFNLNVNVEDVMIRENDEICGLVINWTAEIGKLHVDPAIRSKVVVDGTGHPAVVCSTVQRKVPGAKLGELGVV  
GEKPMWADVGEMKLLDTKEVYPNLYVAGMAANAVAGAPRMGPVFGGMLLSGKQVAELIIERLG

>Methanococcoides methylutens MM1

MKLDEVTISRAIIDEFSKVFLDÝTEVDVALVGGGPANLVAALKYAEAGLKTVIYEKKLSVGGGMFPRIVQEEARHILDDFGIDYHEYEG  
YYIANSVESVGKLISAATSAGTEIFNLNVEDVMIRDNDNEVCGLVINWATEIGRLHVDPRAKVVVDGTGHEAACNTVQRKVPGAKLGELGVV  
GEKPMWADVGGERMLVETTREVYPNLYVDGMAANAVAGAPRMGPVFGGMLISGKQVADLIIERLK

>Methanococcoides vulcani  
MKLDEVTISRAIIDEFSKVFLDYTEVDVALVGGGPANLVAALKYLAEGALKTIVIYEKKLSIGGGMWAGGMMFPRIVVQEEARHILDDFDITYHEYEKGYIANSVESVGKLISGATTAGTEIFNLVNVEDVMIRENDEVGLVINWTAVEIGRLHVDPLAIRAKVVVDGTGHEAAVCNTVQRKVPGAKLGDLGVVGEKPMWADVGERMLLETTKEVYPNLYVDGMAANAVAGAPRMGPVFGMLLSGKQVAELIIERLK

>Methanococcus maripaludis  
MDGKLRADEVAVTKSILKSTFDMWMDLIDVDVVIVGAGPSGLTAALKYLAQNGVKTVVLERHLSFGGGTWGGGMGFPNIVVEKPADEILREAGIKLDEVIGEPELTADSVEVPAKLVAAIDAGAKILTGIVVEDLILKEDKVSGVIQSYSIEKAGLHVDPITISAKYVIDSTGHDSSVIHTLARKNKDLGIEVPGKSMWADKGENSENLTRNTREVFPGLYVCGMAANAYHAGYRMGAIFFGMYLSGKCAELILEKLENK

>Methanocorpusculum labreanum (strain ATCC 43576 / DSM 4855 / z)  
MDLEVTKAITESWFARLQENLCFDAIVGTGPGSLIAAVKLADAGYKVSMFESKLAPGGMWGGAMLFSSIAVQNEAVYLLDELEIPYKRYNENLVVCDSDLATSALIYQASKRGVVIHNGMSVEDVVFMDNRVSGVNVNWGPVREGLHVDPPLSFRAKIVDATGHPCMISETAARKNNITLNTPTGKVCGEC SLNAVEGEAMTVNTKEIYPGLYVCGMAANGVFGSPRMGPIFGGMLLSGEKVAKLIIIEKL

>Methanoculleus thermophilus  
MTINEVTISRAILESHRALIEHLEMDDAVVVGGSGLACAALLGEKGLSCALIEKKLSIGGGMWGGMMFPRIVVQEEARRLLDRFGIAYKEFEPGYYVAKSVEAVAKLTAACDAGVEFFNLTTVEDVMIRGDGRVGGLVINWTPVDMAGLHVDPPLTACTCTVDASGHDAVVARMIERKGGLQVKGESFMWAERAESRILDHTKEVFPGLVAGMAANAVAGECRMGPIFGGMLLSGERAELVAESLER

>Methanohalobium evestigatum (strain ATCC BAA-1072 / DSM 3721 / NBRC 107634 / OCM 161 / Z-7303)  
MELDDITITKAIVDDFSKTFIDYTEDVALVGGGPANMIAATRLAQEGYKVALFEKKLALGGMWGGMMFPRIVVQDEARKILEEFDINHYEYDNEKGYIANSIESVSRLINKTVTSVQVFNVNFEDVMIREDDRVTGIVINWTAVSIANLHVDPPLTIRAKVVIDGTGHEAVVCNTVQRKIPNAKFEGVVERPMWADAGEKSLLKETTREVYPGLIVTGMAANAVAGAPRMGPVFGMLLSGEMAAKIAMSCLD

>Methanohalophilus euhalobius  
MELDERIITRAIVEFTNVFLDYTDVDFVALVGGGPANLVAARYLAEGALKTFLFEKKLSVGGGMWGGMMFPRIVVQEEARRILDDFDVPHYEYEEGYYVANSVGTVGKLISAAVSAGVEIFNLVSFEDVMIRDNEVCGLVINWTAVEIARLHVDPPLTIRAKVVIDGTGHEATVCNTVQRKIPGAFFGKEVVGEEKPMWADTGERLVMKNTREVYPGLIVTGMAANAVAGSPRMGPVFGMLLSGEKAQALASRLKD

>Methanolacinia petrolearia (strain DSM 11571 / OCM 486 / SEBR 4847)  
MKLDEVTISRAILSEQHKIMTEYLDIDCAVVGGSGLTCAAILQAQNGVKVALIEKKLSIGGGMWGGMMFPRIVVQEEARRLLDHFGIKYTEYEKGYYVASSVEAVSKSLAAACDAGAEVFNLTTVEDVVVKEDGGVSGLVINWTPVEMAGLHIDPLTMRTKTVDATGHDSMIAHMVRKKGAALEIKGEGFMWAERAETNILSHTKEVFPGLIVAGMAANAVGGETRMGPIFGGMLLSGEKAANMIIERLK

>Methanolinea sp. SDB  
MELSETTIITRAIVSSQMKİLEYSELDAVAVVGAGPSGLTAAAILGDAGYKVGVIKKLSVGGMWGGMMFPRIVVQEPAARRLLDRFEISYQPFEEGYYVASSIEAVARLTSAACRGAEFFNLTSVEDVMVKDDGRVSGLVINWTPVEMAGLHVDPPLTICRYTIDATGHDAVVATLVERKGRNLEVKGEGFMWADRAESEIISHTREVYPGLIVTGMAANAVAGEHRMGPVFGMLLSGEFAASLVREKLNR

>Methanolobus profundi  
MELDETIITRAIVEEYSKVFPLYIEVDVALVGGGPANLVAALKYLAEGALKTFLFEKKLSIGGGMWGGMMFPRIVVQEDAKHILDDFNINYHEYEKGYYVASSIESVGKLICGATDAGAEIFNLIDVEDVMIRENDTVCGLVINWGPVSMNRLHVDPPLAIRAKVVIDGTGHDAGICSTVQRKIPGTDIKLDVVGEEKPMWADVGEKILMDTTKEVYPGLIVTGMAANAVAGAPRMGPVFGMMLSGKKAELAIEKLRK

>Methanomassiliicoccales archaeon PtaB.Bin215  
MEIDEVLVLTRKIVERYTEEFLENVDVVIAAGAGPSSLTAARYLAKAGLRVVIKERKLTPGGMWGGGMFTPIIVVQEGSKDLLGEIGVRLRDAGDGYFTADSVEASAKLISAAVTAGARLYNTISVEDVMIRQDSICGVVINSSAVEAGLHVDPPLAVRSKYVIDGTGHPAEEVHHVVKVGRNLNTPTGQIEG EKSMWAEQGEKDVVELTGEVLPGLYVTGMANAVAGAPRMGPIFGGMLLSGRKVAEMIAREKKKGKK

>Methanonatronarchaeum thermophilum  
MNVDKFKVSKAIIDEFSKDFLDSLSDVDAIGGAGPAGMVAALKYLAENDIKTAVFERKLTSVGGMWGGMMFPRIVIKEKSLPILDDLNNINYREYQDGYYIANSIESVGQTAAEAVKAGAEIYNLMTVEDLHYKENKVNGVVINWSSVLDAGLHVDPPLTIESKITIDATGHDCELVKVAQERINKLNKTGKIMGEKSMWAEQGEKDVVELTGEVLPGLYVTGMANAVAGAPRMGPIFEGMLLSGKVAEQCICKLK

>Methanoplanus limicola DSM 2279  
MKTKVYESENKMLDEVAISRAIVSEQSVMFLYYDDLCIAIVGAGPSGLTCAMLGEEGLKVGVIKKLSVGGMWGGMMFPRIVVQEEARRLLDHFGIKYREYESGYFVSSSVEAVAKITSAACDAGAEFFNLTYVEDVVIKGDNRISGLVINQTPQMTGLHIDPLLATKVTIDATGHDSVVAHLVRDKGGSVEIKGEGFMWADRAESNILSHTKEIFPGLIVTGMAANAVGGETRMGPVFGMLLSGEKAALKSALKK

>Methanopyrus kandleri (strain AV19 / DSM 6324 / JCM 9639 / NBRC 100938)  
MEREITPIVLREGYEFINDCESDViVVGAGPAGLTCAYELAKSDVDTIWERKLYVGGGMGGMLFPAGVIMEETAEVLEEVGVELRPAEAGLLAFNVEAAIKLANALEAGARILVGLIEVEDVIERRGRVCGVVNVTAVKAANMHVDPLALEAEYTVDATGHEAAVCKLAGIEVKGEPMWAERGEELVKHTQEVKPGFLVAGMAASAVKGAYRMPGIFGGMLESGKAAEELERL

>Methanoregula formicina (strain DSM 22288 / NBRC 105244 / SMSP)  
MELDELTISRAILASQTNVLINHLEDDAVVGGGPAGLTCALIAGQGKKVGVIEKKLSVGGMWGGMMFPRIVVQEEARRLLDGFIRYTFESGYYVARSVEAVSKLTAACDAGVEFFNLMSVEDVMIKADKRISGLVINWTAVEMGKLHVDPLOVMSRTVDATGHDAVVARLVEKKGGDIRVKGEKGFMWADRAETNILNHTKEIFPGLVVAGMAANAVAGESRMGPVFGMFLSGERAQIVLREMKA

>Methanoregulaceae archaeon PtaB.Bin009  
MELDEITISRAILSSQVEKLLFEMDVAVGPGPSGLTAAALIGEQGFRVGLIEKKLSVGGMWGGMMFPRIVQEEAKRLLDQFDIAHTSYEGYYVASSVEAVSKLTASACDAGVEFFNLFSVEDVMIRGDSRLSGLVNNWTPVEMAGLHVDPMTGCRAVDATGHDAVLARLVERKGDVKRGEGLWWADRAESEIVSHTREVFPGLVVCGMAANAVAGEHRMGPVFGGMLLSGERAALATSSLRQENSAA

>Methanosaeta harundinacea  
MALDEVITITKAIVESYMESFLKYTDVDVALVGAGPANLVAAKKLAEADAKTVVFERNLSVGGGIWGGMMFPRIVQKEGCRLDEFGVWYREYEGLYYIASSIETVAKLTAGVIDAGAEIINLVTVEDVMIREDERIAGLVINWEAVERTRLHVDPLSVRARVVIDGTGHDANICKVVQRKIPGAKVGSGLGVVGEKPMWADVGEKTVVETQEVYPGLATGMAAAAAGVAGGPRMGPIFGGMLLSGEKAAMLAKEKLGL

>Methanosalsum zhilinae (strain DSM 4017 / NBRC 107636 / OCM 62 / WeN5)  
MELDEVVITRAIVEDEFNLNVFLDYTDVDVALAGGGPANLVAAKYLAEAGYKTVLFEEKLSSIGGGMWGGMMFPRIVQEEARRILDDFNITYKEYEDGYYVANSIESVSKLAAGATSAGAEIFNLVSEDVMIREDRVSGLVINWTAVGIGKLHVDPMTIRSKVVIDGTGHDASVCNIVQQKVPGAQLGELGVVGEKPMWADVGEKLLMETTREIYPLIVSGMAANAAAGAPRMGPVFGGMLLSGEKAELAISKLD

>Methanosarcina acetivorans (strain ATCC 35395 / DSM 2834 / JCM 12185 / C2A)  
MELDEVIITRAFDEYSKTFDIDVALVGGGPANLVAAKYLAEAGVKVALYEQKQLSLGGGMWGGMMFPRIVQEEATRILDDFGIRYKEYESGYYVANSVESVGKLIAGATSAGAEIFNLVSEDVMIREDRVSGLVINWTAVGIGKLHVDPMTIRSKVVIDGTGHEAVVCNTILRKIPNAKIGELGLLGEEKPMWSEVGERLAVNATQEIYPLIVAGMAANAA TRAPRMGPVFGGMLLSGEKAALKALDLRKLT

>Methanospirillum stamsii  
MTLDEITISRAIISDMYHMTLLEYMEMDVAIVGGGPSGLVCALSIAEKGYKVGLIEKKLSIGGGMWGGMMFPRIVQSEAKRLLERFNITHSEFSPGYYTARSIEAVSKLTAAVDAVGVEFFNLTTVEDVMVKGDGRSLGLVINWQPVEATGLHVDPMTIRCRMIVDATGHDAVIAHYVSKMGKPDIKGEGTMWADNAESAVVTHTKEVFPGLVFCGMAANAVSGGHRMGPVFGGMFLSGESAAVQILQQL

>Methanothermobacter defluvii  
MEIHAGVKMKLDDIKISRAIVEGYMEDLLDYMEMDVAIGGGPSGLTAGYYLARAGLKVALFERKLSIGGGMWGGMMFNKIVVQDEGREILDEFGIRSEPYDEGYHVADSVEATSTLCSCAQAGLKIFNLMISIEDVMIRDEGITGLVNLNWSVEMAGLHVDPMTVRAGAVIDATGHDCEIVKVVERKIGPELNTPDGRIQGERSMWADVGEEAALIENTREVYPNLVYAGMASNAVYGA PRMGPIFGGMLVSGRRVAEMIEKLK

>Methanothermococcus okinawensis (strain DSM 14208 / JCM 11175 / IH1)  
MDKFKIEEKDVTTSILKATFNMWMDIVDVDVIVGAGPSGLTAARYLAKEGVKVVVERHLSFGGGTWGGGMGHPYITVQKPADEILREVGVKLEEDGGLYVADSVEVPAKLGVGAIDAGVKILTGVIVEDLILKENKVS GVVINSYAIKDAGLHIDPLTINA KYVIDATGHDASVTNLARKNDLGLLEVPEGKSLWAEKAENSILRH TREIFPGLVFCGMAANATHGGYRMGAIFGGMLSGKKVAELILEKLKNND

>Methanothermus fervidus (strain ATCC 43054 / DSM 2088 / JCM 10308 / V24 S)  
MVLNEVTISKAIISKYMEELIDNTNL DVAIAGGGPSGITAGYYLAKEGFVKVALFEKRVSVIGGA WGGMMFNKIVVQEEGKKILDEFDVNTERYENNYYVADAIEMITTLASKACKSGLKIFNLINE DIVINKKKISGIVVNWTAAEMA KIHVDPLVIKSFKVIDATGHDCEVVKA VEKKLGPVLTETGRIVGEKPMWAEKGEKAVIKNTGEVYPNLVYAGMAANSVYGSYRMPGP IFGGMLLSGKKVAELIRERLL

>Methanotherrix soehngenii (strain ATCC 5969 / DSM 3671 / JCM 10134 / NBRC 103675 / OCM 69 / GP-6)  
MSLDEVMTKVAIVEGYLESFLENTEVEAALVGAGPANLVAAKRLAEBANI KTVLFERKLSVGGWLWGGMFPRIVVQQAIRILEEYGYRHEHCKGYYVANSIETVAKLTARAIDAGAQIVNLTVEDVMIREQDRVVG LVINWTAAEMA QIHVDPLCIRARYVIDGTGHEASVCRVVAR KIPGAIGIDGVVGEKPMWAEVGERTV VEMTQEVYPGLV VAGMAAAAVCGGPRMGPIFGGMLOQSGEKAAGIVIENLNK

>Methanotherrix thermoacetophila (strain DSM 6194 / JCM 14653 / NBRC 101360 / PT)  
MALDEVKITRAIVESYLESFLKCTDVDVALVGAGPANLVAAKRLAEADVRVVLFEKRLSVGGWLWGGMFPRIVVQKEACRILDEIWIYREFEEGYYVADSIEVVAKL TAGAIDAGAE LINLVSEDVMIREGDRIVGLVINWTAAADMAGI HVDP LAIRARV VIDGTGHEASVCRVVAR KIPGAIVGESGVI GEKPMWAALGEKIVV DATREVYPGLIVAGMAATTVAAGPRMGPIFGGMLLSGEKAASIALEKLQAQSVD

>Methanotorris formicicus Mc-S-70  
MDLRLKADEYTTKAILKSAFNWMDDIDVDVAIVGGGPSGLTAARYIAKKGYKVVLERHLA FGGGTWGGGMGF PYIVVEEPADEILREVGKLEKVDGEEGLYTADSV E VPAKLA VGSIDAGAKILT GIVVVEDLILREN R VAGV VINSYAI EKAGLHIDPITITAKYVVDATGHDASVATL SRKNPELGL EVPEKSMWAEKGENALLRNTRREVYPGLVFCGMAANATYGGNRMGAI FGGMYL S GKCAEMVVEKLNNE

>Nitrospira bacterium SM23\_35  
MELDEVVITKAIVDQFCKKLTNHLET DVAIVGGGPSGLVAGYFLAKAGRKTVLFERKLSVGGMWGGMFLNEIVVQKA VRILKEFGITYCEFQKNYYTADSVESISTLISRAVQAGVTIFNCITAEDVLMRTSRTGLV LWNWSAVE MARLHVDP LA VR SRFV VDATGHETAVV RL VQN KVP GTL KTL SGK VGEKSMWSDKAES TLK NTREVFPGLVYAGMAANATFGGPRMGPIFGGMLLSGEKA KL LQAL SQKSK

>Nitrospira japonica  
MGKPKPAPL RERDITRQIAREYYKEFDQ LIESDVIIVGAGPSGLICA HDIADM GFKTVIVEQNL ALGGGF WHGGYLMNKATICAPAHK ILDEIDVPC KRIKDCEGMYIVD PHATGALIAAA YRAGAKVNLTRV D LILR RDGSLDG VV VNN TAE MAGHDLI HVDP IALES KIVV DATGHD A VV VNL LH KR SLYTEVPGNGAMW VSRSEEDVMDHTGEVYPNCFVIGLAVSAVHGTPRMGPAGFSMLLSGRYGAELKKKLQ

>Nitrospiraceae bacterium  
MNPINVAPL RERDITRHIAREYYKEFD S LIESDIIIVGGGPSGLL CARD LATSGF RTLLIEQSLALGGGFWSGGFLMNKATICEPADQI LEELGIPFKPIKDCPGMTMVDPPHVT SRLISAAYEAGVKIMNLT KVV D LILR QDHRIEGVV VNNSTVEMAGHDTI HVDP IALES QIVV DATGHD A VV VNL LH KR NLYQKPGNGAMW V ARSEALV VENTREIYPNCFVAGLAVA AVDGS PRMGPAGFSMLLSGRYAA ELVR QKLGE

>Nitrospirae bacterium  
MPKPPTPAPLRLERDITRHIAREYYKEFDQLIESDVIIVGAGPSGLICAHDLAAMGFKTVVVEQSLSLGGFWSGGYLMNKATICEPANEILEEIGVP  
CKKITECAGMMYMDPPHATGALIAAYRGAKIMNLTKVDLIIRRDIILEGVVVNSTTAEMAGHDIAHVDPPIALESKIVVDAKGDAIVVELLHKR  
NLHKAVPGNGAMWVAQSEQEIMDRTEVYVNCFVIGLAVAAYGTPRMGPAGSMSLLSGRYGADLIKLLKG

>Omnitrophica bacterium RBG\_13\_46\_9  
MDEALISRAITESFTKDFIDAFNVDAIAAGAGPSGLICACYYLAKQNWKAVFERHLRVGGGMPGGMMFNRIIVQEEAMPILEFGVSAKRYKKDLY  
IVDALEAISTFCSTIKRGAKIFNLINVEDVVIRKDRIAGVVLNWSAVSWAKLHVDPMAVRSKAVVDAKGHDSEIARIVERKTGPVLRTEGGVIGE  
KSMWAEIGEKMILENTKEIYPGLIVCGMAANAVFGSPRMGAIFGGMLLSGKAAEVARKVIKSK

>Omnitrophica WOR\_2 bacterium RIFCSPHIGH02\_02\_FULL\_68\_15  
MFARAQEAQITRAVRAFAKEFDGLVRSVLIVGAGPSGLVAAMDIARRGRRLVVEQTNYLGGGLWLGGYLBNKLTVRAPAHRLKELKVP瞿Q  
PGLYVADAPHVCARLIAAACDAGVKFAQMTEMVDVVVREGGRVEGLVINWSPVSALPKLAHVDPVALEAKVVDAKGHDAAVVRLLAKRGLAAPVP  
GDGAMWVERGEQAVMDKTGEVHPGLFAAGLAVSAVHGTPRMGPAGSMSLLSGRCAQMIERAYFA

>Peptococcaceae bacterium SCADC1\_2\_3  
MPLEDITISKAIITTRYNQELLALESVDVAIAAGGGPSGLVAASYLAQQGAQVVLVERNLSLGGMWGGMMFNQIVVQEEAIPILDTGFVRYRTFEPG  
YYTAHATEVAAIILGAVRKGVKILNLISAEDVMVRNERVCGLVLNWTAVGALARLHVDPIAASCSCVIDCTGHDAQIANIVRKMGAVLKTSRGKIE  
GEKPMWAERGEAAIKNTGEIYPGLVAGMAANAVGNHRMGPVFGMLLSGKRAELIKGER

>Phorcysia thermohydrogeniphila  
MQNLNEVIISQAAIESFMEKLKNSLEVDVAIVGGGPSGLVAGYYLAREGFKVSIYERHLAIGGMWAGGMLFNEIVVQEMGREVLDEFGVRYREFQP  
GYYVADSVEAVTTIASKAVKAGAVIFNGVTAEDVVLKKVNDEYRCGLVINWTSVERSRLPVDPVITAKEYVIDATGHDAVSVSTLQKKAGIKLATE  
TGCVIGEKPLWASVGEEDTVKNTREVFPGLFVSGMAANATCGSHRMRGPVFGMLVSGKAAQEIAEKLLKGKNEE

>Planctomycetes bacterium DG\_20  
MDDDFDETDSQAILRAYYAKVADALQGDVLVVGAGPSGLVAAWRLAQAGHRVVVLEKRLSPGGGIWGGSLGMNEVAVQKHIALILDEAGVHQPSGR  
LFTADAMELASCLKALHAGAVILNLMTAQDVCRSGRTGVVANRSLLGESLPIDPIVSARAIDATGHEAVLANCIQRGLLKNSLGRLPGE  
PLDAPAGERFVVDHVAELYPGLWTGMSVCASLGGPRMGPIFGGMLLSGEKVAALVGQALKTRPQVRHE

>Porphyromonas sp. CAG:1061  
MEKLVSQGIITTYFEKMEKSLDLDVAIVGGGPSGIVAAAYYLAKAGLKVAQFDRKLSPGGMWGGAMMFNEIVIQQEEALEIIKEMGINYEQYQDKLYT  
MDSVESTALLYNAVHAGARIFNCSVEDVVYKENRVSGVVNNWTPVLRGMDPLNIMAKYVIDGTGHDSEICRVVAKNGATLNTSTGGVGEQ  
SLDVITGEKMVEGTKEIYPGLYVCGMASSAVGTTPRMGPIFGGMILMSGKKVANLIIDQLK

>Prevotella amnii  
MIEKEISKGIITTYFEKMEKSLDLDVAIVGGGPSGIVAAAYYLAKAGLKVALFDRKLSPGGMWGGAMMFNFQIVIQKEALDIKEFEINYEQYSDNLF  
TTDSIECTAAILYKAHVAGATIFNCYSVEDVVFKNNIVSGVVNNWTPVLRGMDPLNIMAKFVIDGTGHDSEICKVVARNNITLNTSTGKVVG  
RSLDVIEGEQQVVEGSKEIYPGLYVCGMASSAVGGTPRMGPIFGGMILMSGKKVADMILKRIQS

>Prevotella nigrescens  
MIEKKISKGIITTYFEKMEKCLELDVAIVGGGPSGIVAAAYYMAKAGLKVALFDRKLSPGGMWGGAMMFNFQIVIQEEALEIIKDFDINYQAFEDGLY  
TADSVESTSALLYKATHAGATIFNCYSVEDVVFKNNIVSGVVNNWTPVLRGMDPLNIMAKFVIDGTGHDSEMCQVVARNGIKLNTATGDIV  
RSLDVAEGERQVVEGTKEIYPGLYVCGMASSAVGGTPRMGPIFGGMILSGKKVADMIILERLK

>Prevotella stercorea DSM 18206  
MIETQVSKGIITTYFDKLQNLLDLDVAIVGGGPSGIVAAAYYMAKAGLKVAQFDRKLSPGGMWGGAMMFNFQIVIQEEAMHIVKDFDINYQAFEDGLY  
TIDSVESTSSLLYHAVHAGATIFNCYSVEDVVFKNNIVSGVVNNWTPVLRGMDPLNIMAKCVIDGTGHDSEMCVVARNGIQQLDTATGGVIGE  
RSLDVVEGERMVEGTREVYVPGLYVCGMASSAVAGTPRMGPIFGGMILMSGKKVADMIIEKLKK

>Prosthecochloris sp. ZM  
MEEKISKFIISQFFAKLEDSTLVDVAIVGAGPSGLIAAKELAKAGKVAIFESKLAPGGVGWGGMLFNEIVLQENIIPILDEYAIRYKTGEFYVT  
ADAVEVSSALIYGAHVAGVRIFNAVRVEDLAMRDERVCGVVINWNPVSRLLEMHVDPVITSRAVLDTGHFSELINLASNKAGITLDTPTGKVMGEK  
PMWMENGESSTVINTKRLYPLGLYASGMAANNAMGGFRMGPIFGGMILSGKKVAGLILEDIQG

>Pseudothermotoga lettingae (strain ATCC BAA-301 / DSM 14385 / NBRC 107922 / TMO)  
MKDTMISTLIVNRYFKKLRSFLELDVAIVGAGPSGLTAAYELAKKGKVAIFEEKNTPGGGIWGGMMFNEIVLEKELEDFLNELGITYVIQENHVL  
VDSVHFASALLYRTTMVGATVFNNISVEDVAMQDGKVCGVVINWGPTMRLGLHVDPITVKASFVIDGTGHF PANVASLLAKRGLIEMKMELEPMNADEA  
EQFVENTGEIFPGLMASGMAACAVHGGPRMGPIFGGMILSGKKIAQIIIEKLR

>Pyrobaculum aerophilum  
MELKIGRAIIRHAKLDDEYSDVDVAIVGAGPSGLAAKYLAEKGKLVVYERRFSFGGGIPGGNMLPKIVVQEEAVPILRDFKVRYKPAEDGLY  
VDPAAELIAKLAAGAVDAGAKIILGVHVDDVIFRGDPVRTGLLWIWTPIQMSGMHDPLYTQAKAVIDATGHDAEVVSVAARKVPELGIQVVGEKSA  
WSEVSEKLVVEHTGRVAPGLYVAGIAVCAYGLPRMGPIFGGMILMSGKKVAEVVKDLMMAEAHAVRA

>Pyrococcus abyssi (strain GE5 / Orsay)  
MLREVTISRAIESYYRDLLNLELDVAIVGAGPSGLAAKYLAEKGKLVVYERRFSFGGGIPGGNMLPKIVVQEEAVPILRDFKVRYKPAEDGLY  
YVADAIEVATTIASKTVKAGVKIFNMIEVEDLVLVVKDNRVS GIVINWTPVLMTGLHVDPVTLV EAKYVIDSTGHGAQVAQFLLKRLIERIPGEGAMWA  
EQGERLTVENTREVFPGLYVTGMAANAIAGAPRMGPIFGGMILSGKKIAQIIIEKLN

>Rikenella microfusus  
MEKLVSLGIVENYFEKLKNNLSDAAIVGGGPGSLVAAYYLAKAGRKVVLAYERKLAGGGMWGGAMMFNDIIVQQEALPILDELGVCYKPYREGACV  
VDSVHATSALVYAAATKAGATIFNCYSVEDVIFRDEAVAGLVNVAPVMREGMHVDPLMTAKTVLEGTGHDCMIARLVARNNVRNLTPTGEVAGER  
SLNVEQGERLTVENTKEIYPGLFVSGMAANGVSGSFRMGPIFGGMLMSGKAAELMIAKING

>Saccharolobus solfataricus (strain 98/2)  
MEVKIKQVDEVKISRYIIKETMEDWYQFVESDVIVGAGPSGLSAAYYLAKAGLKTFLVFERRLSFGGGIGGGAMLFHKLIIEKPADEILREVNIRLK  
EEVEGVYVVDSAEFMAKLATAAIDAGAKIIHGVTDDVIFRENPLRVAGVAVEWTATQMASLHVDPFISAKAVVDATGHDAEVISVAARKIPELGI  
VIPGEKSAYSERAAELTVINTGKVAEGLYATGMATVEVKGLPRMGPIFGAMVLSGRKAVAGEITKDLLKSEIRA

>Smithella sp. SDB  
MLNETTISRRAILDAYFKKLDSCLELDVAIVGGGPGSLVAGYYLAKAGRKVALFERRLSIGGGIWGGGMMFNIAIVQEAGRQLLEFDLKGSEYAPGY  
YVLDADVDTATLIHKAVRAGLQVFNLNIAMEDVVIKNERVAGLVINWGAVIDTLKHVDPLTIHARYVIDGTGH PANVTEVLVRKMGRVLSNTPTGKMMG  
EKSMEAEGQELQTVENTREVYVPGLYVSGMAANAVFGGYRMGPVFGGMLLSGRKAAEELLAHL

>Spirochaetes bacterium ADurb.Bin215  
MLDETVISRAIETYMKKLTDLNSDVAIVGAGPSGFVAGYFLAKAGRKVVI FERALAVGGGMWGGGMGFNEIVVQEEGKAVLDEFDLPAVRYVQGY  
YTLDHSVRAVSALALRAVEAGVTVFNLGVVEDVVLHDERVSGLVLNWGTVMKAGIPVDPVTHSRCVLDTGHPAHVAEVLCRKMGVSLNTPTGKMMG  
EMMSDAEKGEKQTVENTREAYPGLFVSGMAANAVFGGYRMGPVFGGMLLSGRKAAEELLAHL

>Staphylothermus hellenicus (strain DSM 12710 / JCM 10830 / BK20S6-10-b1 / P8)  
MKFFPQONLYELSEGDSLTKLIDALYKKLSEIVKVDVAIVGAGPSGLTAAWKLGEKGYKVLVLERMLGVGGGMRGGSMLLPVGLIEDGEAAEIAREAG  
ARINKRNGLFVVDPSELAVRLASKAIAENGAIIPWGVLEDLITRGRGEDLTVKGVLINWTPPIYEAGWHVDPFYIEANAVV DATGDGSLLRLVLAKR  
HPELKINIPGMSSQNVWIGEMVVEKTSMMVKGLFTGMSVAELYNTNRMGAIFGGMLVSGRKVADLIDDYFGKTRTLREQ

>Sulfolobales archaeon SCGC AB-777\_J03  
MASVRVPESKISR FIVEETMKDWDMDIVESDVIVGAGPSGMATAAYYLAKAGLKTFFRLGFGGGIGGGAMQFHRLVIEEPADEVLRFGVRLKK  
VDEGVYVVDAEAEFMALKASAKAIDAGAKIILGVTDDVIFREDPPRVAVGVAWEWTATQMSGHLHVDPFLFISAKAVV DATGHDAEVISVAARKIPELNIS  
VPGEKSAYSEVAEQLVVNDTGPVAPGLYAAAGMAVCVEKSLPRMGPIFGAMVLSGRKVAELIIQDLRK

>Sulfolobus acidocaldarius  
MSDSIKIKAIDEVKISRYIIKQTMEDWMNFVENDVVIVGAGPAGMSAAYYLAKHGLKTFLVFERRLSFGGGIGGGAMLFHKLVIESPADEVILKEMNIR  
LEKVEDGVYIVDSAEMAKLAASAIIDAGAKIIHGVTDDVIFRENPLRVAGVAVEWTATQ MAGHLHVDPFISAKAVV DATGHDAEVVAVASRKIPEL  
GIVIPGERSAYSEMAEKLTVQTVGVPAGLYVAGMSVTEVRGLPRMGPIFGSMVLSGRKVAEDIIKDLRNS

>Synergistaceae bacterium  
MRLDEVTTKAIMERYFDKFMMNLELDVAIVGGGPGSLVAGYFLAKAGRVALYERKLSVGGMWGGMLFNEIVVQEDAKRLLDELDVPTLPYKEA  
GYYTADSVEVTSTITSKAVKAGLVVFNCSVEDVVVKDDRISGLVLINWTA VPMANLHVDPPLSIRSRYVIDATGDTEVVAMVAKAPGRLLTPSKNI  
EGEKFMNPEEAERLTLNTKEVFPGLYVAGMACNATFGGPRMGPIFGGMMLSGEKVARLILKELSK

>Syntrophaceae bacterium PtaB.Bin038  
MALNEVTISRAIETYTKKLAHLDVDVAVVGGGPAGLVAAYFLAGAGRKVALYERKLSIGGGMWGGMMFNEIVVQAEAKGILDHFGVRTQEYAPG  
YYTADAIEAVTICSRQAGAKVFNCITVEDVVIRDNRMGLVITWSPVEMTGLHVDPPLTIHAKAVIDATGDTEVLHVIER  
KADVTLNTPGKLMGERSMWSEKAERLTIDNTREICPGVYVAGMSANA AFGGPRMGPIFGGMMLSGRKVAEQILAGG

>Syntrophobacter sp. DG\_60  
MALDELIITQAIVERFSEKLKGCLMDVAIVGAGPSGLVAGYYLAKNGHRVAIFEKKLSIGGGMWGGMMFNQIVVQTEGKRILDEFI K TAPFSEG  
YYTADAIEAITTICSKACQAGVN FNTISVEDV L VREGRVIGLVLINWSTVEMAKL DVPLTIRAQYVIEATGHATEVVK VIEKKMGE SLLPTGKII  
GEKSLWAEIGEADTIKNTKEA PGLFVCGMAANATFGSYRMGPVFGGMLLSGRKVAEQIL

>Syntrophobacteraceae bacterium  
MELNEITITQAIVDRFLEKFRNSLET DVAIVGGGPAGLVAAYFLAGAGRKVALYERKLSVGGMWGGMLFNEIVVQEEAKRLLDLFGVGSHHYRD  
YYTADSVEAISTLTSQAVKAGVMIFNCISVEDVV MRPERVIGLVLINWTA VEMAGLHVDPPLAIRAKFVVDATGDV EVV VVVKR VPGKLLTPSGEIE  
GEKSMWSEVAEKLTLNTREVFPGLYVAGMAANATFGGPRMGPIFGGMMLSGEKVAQTLIDQ LKK

>Syntrophobacterales bacterium CG\_4\_8\_14\_3\_um\_filter\_49\_14  
MELNEITITKAIIERFSEKLI ACTEVDTAIVGGGPAGLVAAYFLAKVGKKVAIFEKKLSIGGGMWGGMMFNEIIIVQPEARELLDLFDVTRKYEAG  
YY SADAIEAVSTICSYATKAGARVFNCITVEDVMIREGRVIGLVLINWTPVSM TGLHVDPPLTIAKSTIDATGHATEVLRVIERKADVRLFTETGKLM  
GERSMWADRAERLTLENTREICPGVYVAGMSANA AFGGPRMGPIFGGMMLSGRKVAELIANG

>Syntrophus sp. (in: Bacteria)  
MELNEITITKAIIERFTEKFLACTEVDTAIVGGGPAGLVAAYFLAKVGKKVAIFEKKLSIGGGMWGGMMFNEIIIVQPEARELLDLFDVTRKYEAG  
YHTADAIEAVSTICSYATKAGARVFNCITVEDVMIREGRVIGLVLINWTPVSM TGLHVDPPLTIAKSTIDATGHATEVLRVIERKADVRLFTETGKLM  
GERSMWAEKAERMLTLENTKEICPGVYVAGMSANA AFGGPRMGPIFGGMMLSGRKVAEQILSR

>Thermaaerovibrio acidaminovorans (strain ATCC 49978 / DSM 6589 / Su883)  
MELDERRISAVI RRFMDRLLSDMDLDVAIVGGGPAGLVA GHNLAREGFKVAMFERKLSLGGGMWGGMMFNQIVVQEEGAQVLREFGVRVLD EGEG  
YY SADSVEAVSTL ISSATRAGL RFNCVTAEDVTMREDRVVGLVITWTPVEMAGLHVDPPLAIRSRF VIDATGDH DINVVRVVERK VPGKLMPTG RAE  
GEKSLWSHRAEELTLENTREVFPGLYVAGMSANATFGGPRMGPIFGGMMLSGRKAAQ LVS RALRGQGGRG

>Thermincola ferriacetica  
MHLDETVISRGIVQKYMEELMMDYMNTEVAIVGGGPGSMVAAYYLVKRGCKVALFDRKLAVGGGMWGGAMMFNKIVVQSAGKRILDEFAISCEEYERG  
YYVADAVESVTIATASMTVKAGCKIFNLIGAEDVMVEDGRVTGLVNWTPVQVNYYHVDPLVRAKYVIDGTGHPAEVTLTRKMGVRLNPTGGVA  
GEKPMNALKGEELDVVENTREVFPGLYVTGMAANAAFGSHRMGPVFGMLLSGEKAAMEIAARLGK

>Thermococcales archaeon 44\_46  
MLRDVTISRAIIETYFKELLEHNLDAIVGAGPSGMVAAYYLAKGGAKVAIFEKLSIGGGIWGGGMGFNKIVVVEEAEKIIEFGVRHEEFEEGY  
YVADAIEVATTIASKNIKAGAKIFNMVEVEDLVLVKENRVAGIVINWTPVKMTDLHVDPLTVEAKFVIDSTGHGAQVTQLLERKGLIERVPGESAMWA  
EMGEKLTVEHTKEIYPGLYVTGMAANAVAGAPRMGPIFGGMFLSGRKAAFEILEKLKK

>Thermococcus celer Vu 13 = JCM 8558  
MLKDVEVSRAIEAYTKDILDLSKLKDVAIVVGAGPSGMVAAYYLARGGAKVAIFEKLSVGGSIGGGIWGGAMGFNRIVVEESAREILDEFGVDYEEFKPGL  
YVADAIEVATTMASKTVKAGVKVFNMVEVEDLVLVKGDRVAGVVNWTPVKMTGLHVDPLTVEAEFVIDSTGHGAQITGHLLKRLIEELPGECPMWA  
EMERLTVEHTKEVFPGLYVTGMAANAVAGAPRMGPIFGGMFLSGRKAAFDILERLG

>Thermodesulfatator autotrophicus  
MALEDEIKISRAIITYFKKLTDYLEMDDAVVGAGPSGLMAAYKLASEGFKVAFVFERRLSIGGGIWGGMMFNEIVVQEEGARLLKEIGVRTEPWNGG  
EYYTADAVEVACILAALKSVQAGAKIFNLIMVEDVMVRDNRVVGVLVNWNSATEIAGLHVDPЛАVKAKYVVEATGHETAVLQVMQKLGAKLNTEGKV  
MGEKSMWAEVANLTVDTYDREVYPGVFVAGMAANATFGAYRMGPIFGGMLLSGERAQLIAERLRQ

>Thermodesulfobacterium geofontis (strain OPF15)  
MELKIQRAIVKFGMEDLYEYSDVDLIVGAGPSGLTSAYKLADKGFKLVYEKRLSFGGGIGGGGNMIPKIVVQEEALPILKDFKIKYKEAEKNLYT  
IDPAELIAKLAvgALDAGAKIILGVHVEDDVIRDNPPRTGVWRWTIAIEISGLHVDPLOYTQSKALIDATGHGAEIVQIAAEKNPELNIIKGEKSN  
WSEVSEKLVVDYTGKVAEGLYVTGIAVCEVFGLPRMGPIFGGMILMSGKKIAEIIIEKDLRG

>Thermodesulfobium acidiphilum  
MTKHFLNPVTDNVSKLILKHYFESITDALTSVIIIVGGPSGLTAARELGNSGYKVVIMERKLSPGGTWGGSMSFNKVIQKDLKDYLNELEIPF  
VEDLDALVVDSCLFASQLIAKALKTONVLFNLMVTDLEYTNNAITGVVNNNTGIETAGLHVDPMVFTQKAVALDATGHDAIAANIYSKRVQLPLRK  
EHFMNAVQGEEDTVNNNTKMLANGLFVSGMAANNVDDGSSRMGPIFGGMILMSGKKIAAKLIMEYIKTV

>Thermodesulfobacter norvegica  
MAELNEIIITRAIIDRYHAKITGNLDVDAIVGAGPSGLVAGYHLAQKGYRVTIFERKLSVGGGIWGGMLFNEIVVQDEARRILEEFGVRVNRYEE  
NYYTADAIETVSLLAARAIQKGVTILNGITVEDVVMRPNRIVGLVILWSAVEIAGLHVDPЛАIRAKYIVDATGHDTENVKVVHKVPGRMLPTGNI  
EGEKSWMSEEAKLTLENTREVFPGLFVAGMAANATFGPPRMGPIFGGMLLSGEKVAHLIDERLKQGS

>Thermoplasma archaeon HGW-Thermoplasma-1  
MTVIDEVVTRKIFDRFSREFLDHLDVDVALVGGGPANMVAAHYLAKAGKKVVLFERKLAPGGMWGGGMNFVIVIQEALPIMNEFGIKVEGSND  
GYYTADSVECVAKKLAKSIDSGARVNPNSMTVQDMVIRDDDNKVGVVINWTGVDIRKHMVDPITRAKFVVDGTGHPCEVNVAKKAGRRLRTPSGK  
VEGERSMWAEVGEETTVNNNTVEIYPGLYVAGMAANVGMAPRMGPIFGGMILSGKKVAEMILQKLGA

>Thermoprotei archaeon ex4572\_64  
MSIVRKVSEDEISKAIINEALKESLVDVDAVVVGSPGSLTCSYYLAKYGLKTVLIERRLSFGGGIGGGMLLPSIAIESPAEELIHDEFVLNIK  
KVRDGLYVVMNPAAFIKLAASKAINAGVKVLLGVSEDVIFRSNPLRIAGVVINWSAVHISQLWDPLFIKAKAVIDATGHDAEVVNIVSKIPDFKL  
AIKGEKSACSIEADELIISYSGKVEGLYVTGMATAKVYGLPRMGPIFGGMVLSGKKTAEEVYRDLSELNR

>Thermoproteus uzonensis (strain 768-20)  
MRKINLIWKLRGAMELKIGRAIIRHGAEDLYEYSDVDAIVGAGPSGLTAARYLAEGKLKVIVLERRFSFGGGIGPGGNMYPKIIIVQEEALPILRDF  
KVRYKPGADGLYAVDPAELIAKLAAGAIDAGAKILLGVHVDVIFRGPRTITGLWIWTPIQMSGMHVDPLYIQTКАVVDATGHDAEVVSVAARKV  
PELGIQLQGEKSAWSEVSEKLVVEHTGKVAAPGLYVAGMAAAVFGLPRMGPIFGGMILMSGKKVAEIVAKDLAAEVHAV

>Thermosiphon africanus  
MWDEVSKIIIVERFFEKLNNDLNVDVAIVGGGPGSALSASYYLSKKGLKVAIFEAKEPNEPGGGTWGGMMFNELVVENDIKSFLDELGMNYLIKDNFIS  
VDSVHFASSLLYNTKAGAVLFNNVIVEDIAFYENKVNNGIVINWAPVIRQKLHVDPITIMAKFVVDGTGH PANVNMVLDRGIDIDLPIKGIREYPM  
NAKEGEKFVVENTKEVFPGLYVGMMAAVSVGGGPRMGPIFGGMILSGKKVAEIVEKLSVEVSK

>Thermosiphon melanesiensis (strain DSM 12029 / CIP 104789 / BI429)  
MWDELEISKIIIVNGFFEKFNALDWDVAIVGGGPGSALTASYFLTKNGFKVVIIFEEKNDPGGGTWGGGMFLNELVVVEELEWMLKEFGMNYKRLNGFIS  
IDSVHFASSLLYNTKVGTKIFNNNIVEDILMEEVNLRCGVVINWAPVIKQRLHVDPITVKAKYVVDGTGH PASVVQMIIDRNLEVELPLDKIREFPM  
NAKEGENFVLKNTKEVFPGLFVGMMAAVSVGGGPRMGPIFGGMILSGKKVAEIVEKLSVEVSK

>Thermosulfidibacter takaii (strain DSM 17441 / JCM 13301 / NBRC 103674 / ABI70S6)  
MLDEKIITKAIIESYTQNLLDYIDMDVAIVGAGPAGLTCAYYLAKEGFKVGFERKLSIGGGIWGGAMFNEIVLQEEALPIVQEMEVSYKPYKEKG  
YYVINAVERFACALGLKAIRAGAKIFNLWSAIDVKVKGEDERVNGLWLLWTPVDTAGLHVDPITVEAKYVVDGTGHDAEIANVVKKLKKLATPTGD  
VAGERPMWAEGEKEATEFTGEVYVGLFVIGMAAVACYGKHRMGPIFGGMILSGKKAAKMILECLK

>Thermosulfurimonas dismutans  
MALDEVKITQAIVERFTEKLEALELDVAVVGAGPSGLMAAYKLAKEGFKVAIFERKLSIGGGIWGGMMFNEIVVQEEGARLLKEIGVEARFWQED  
YYTADSVETVCAVGLYAAKAGAMIFNLISVEDVMVRKDRVGLVINVWATVEMGGHLHVDPLAIRSKYVVESTGHELSVLHIMQKLGVLMTPSGKIE  
GEKSLWADVAETTLENTREVFPGVFVAGMAANATFGSYRMGPIFGGMILSGEKVAQEIAARLK

>Thermosyntropha lipolytica DSM 11003  
MVINDIKITRSIIIEYYYAFTRDFLDCDVIVGGGPAGMTAAAYTAQQGLRTVVLESRLSPGGMWGGGMFFNQIVFQPEAGEILQELGISYTANREG  
YLVVP SYRAVASLILAADRAGARILNGITAEDIMRENRCGVVINWTAAVKLGMDPLCIGGKVVIDATGHDAGIVRTYLDKSGGSLPEDEEERI  
RTSSMWAAKGEEMVVEYTRFITEGLIACGMSVSSLFNTPRMGPIFGGMLFSGRKAAELALDYIRKVKA

>Thermovirga lienii (strain ATCC BAA-1197 / DSM 17291 / Cas60314)  
MKLDEKIISKAITTRYQKILSHIQVDVAIVGGGPSGLVAGYYLAKEGRVALFERKLISVGGMWGGMLFNEIVVQEDAKEYILEDGFVRVQFWEDA  
GYYTADAIIESVCISITSKAIQAGLTVFNCISVEDVSVEGDRITGLVINWTPVEMSGLHVDPPLSIGASFVIATGDTEVVHMVAKAPGKLMTPSGDI  
EGEKFMCPDEAEKKVENTKEVFPGLYVAGMACNATFGGPRMGPIFGGMLLSGRKVAALISQRLK

>Treponema sp. CETP13  
MLEYNVSKGILDSYHTKLKSALDSAAIVGSGPSGLVAGYFLAKAGKKVVMFERELAPGGGIWGGMFNDVMQEEAATILSEIGVELPEVKDNFY  
TIDSVYLASTLISKAVEAGVTLLNMISIEDIIFAKDESIGGVVLNWAPVHKEHMHVDPPLMAISRCVLDATGHPSEIVNLTRKNEITLNTKTGKVMG  
ERSLKCKKAEALATAENTCEIYPRLFVSGMAANGVAGAYRMGPVFGGMIRSGKKVAEQMLQCIDTEAPIYD

>Vulcanisaeta distributa (strain DSM 14429 / JCM 11212 / NBRC 100878 / IC-017)  
MAGIYISESSITRAIMRSALKMLDEYSSVDVAIVGAGPSGMAAYYLAKAGLKTIVLERRFSFGGGIGGAASHLPSIVVEYPASDILSKDFGVRLQD  
MGDGLFAVDPAEMIAKLA VRAIDAGAKFLLGHVDDVIIRDNPPRVA GLAVY WSTVQ MAGVHTDPFFIEAKAVV DATGHDAEV AAVTTRKNPDGLA  
IHGEKSAHASVAE DLVVKYTG RVM EGLYVTG MAVA AVYGL PRMGP IFGSMIMSGKRVAELI INDLRR

>Zestosphaera tikiterensis  
MEPLEAKISKIWIWKETLN DWL KLSNV DVVVVGAGPSGMVTAKYLADSGIKTLVLERRLSFGGGIGGGMLMHKV VVDSKALNI LDDFKIKYSRS DYE  
GLYVVDASELMAKLAAGAIDSGAKIVNGITVEDLIVRDNPFRVEGVVIQWSAVNL SGLHVDPPLFIYSKAVV DATGHDAEV LKVLSRKNPEVNLKIPG  
EKSAYAELSEELVVVKHSGKVLPGLYVSGMAVAALYGIYRMGPIFTGMLLSGKKVAEEIAKDLRGGSQ

**Supplementary Table 4 Occurrence of ROS defense genes in genomes of prokaryotes whose non-Cys THI4s have crystal structures**

The predicted proteomes in NCBI were searched by BlastP with the indicated query sequences. Significant hits (e-value < 1e-05) are indicated with a plus sign.

Enzyme	Query <sup>a</sup>	Organism			
		<i>Thermovibrio ammonificans</i>	<i>Methanococcus igneus</i>	<i>Methanococcus jannaschii</i>	<i>Methanothermococcus thermolithotrophicus</i>
Catalase-peroxidase	WP_013536945	+	-	-	-
Cytochrome c peroxidase	WP_013537906	+	-	-	-
Cytochrome bd complex	WP_013537252 WP_013537253	+	-	-	-
Heme-catalase	WP_000077872	-	-	-	-
Mn-catalase	WP_000488336	-	-	-	-

<sup>a</sup> GenBank identifier.