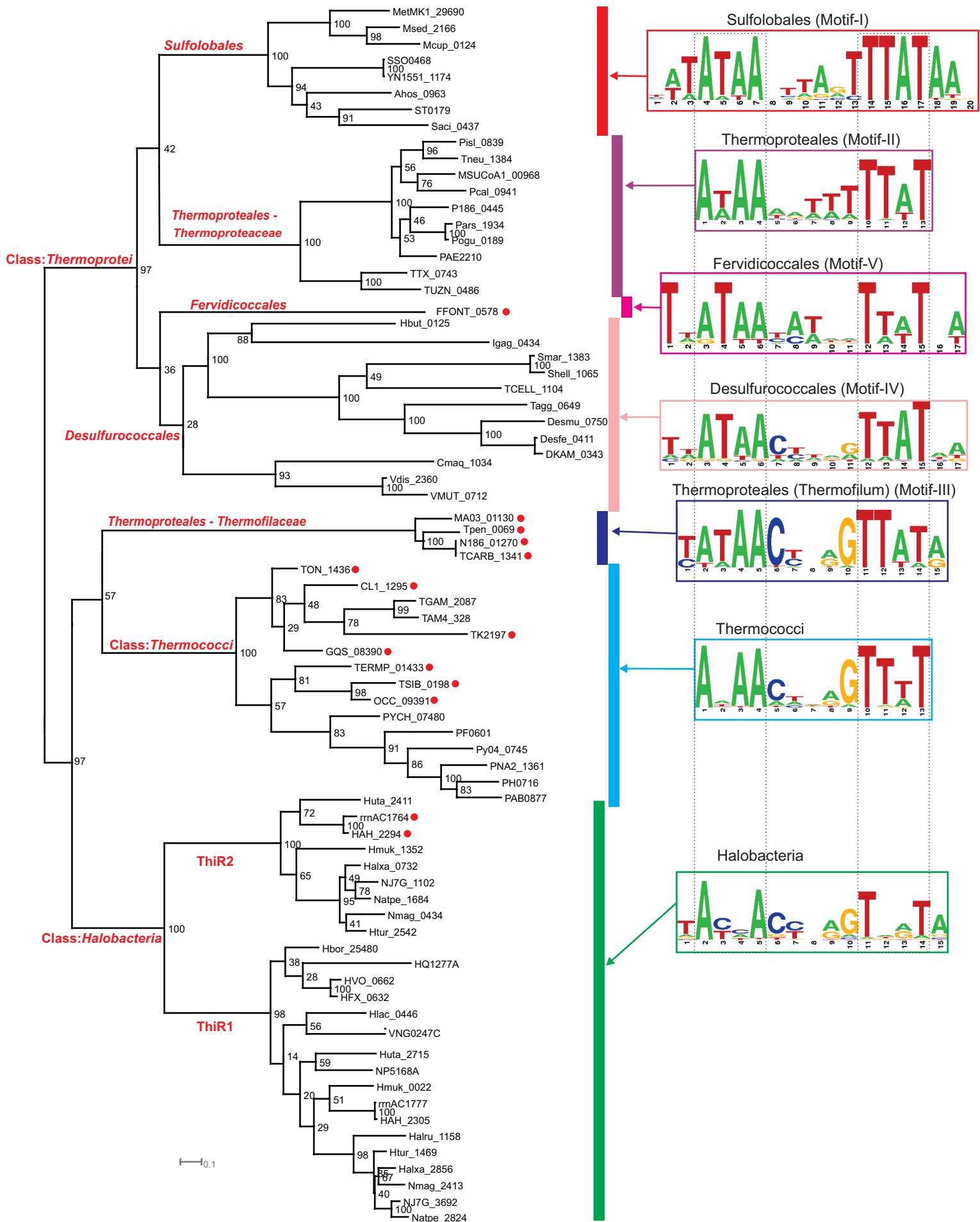


**Figure S1. Distribution of thiamine biosynthesis pathways, transporters and regulators in Archaea.**

The phylogenetic species tree was constructed using the concatenated alignment of 78 universal bacterial proteins. The tree includes the analyzed representatives of three archaeal lineages possessing *thiR* orthologs.



**Figure S2. Phylogenetic tree of the ThiR proteins and their predicted lineage-specific DNA motifs.**

Red dots show ThiRs that are predicted to be autoregulated, as the *thiR* genes belong to ThiR regulons.

**Figure S3. Phylogenetic footprinting of upstream regions of thiamine biosynthesis/transport genes and identification of candidate ThiR-binding sites in diverse archaeal lineages.**

**A. *Sulfolobales* (Motif I)**

**Upstream promoter regions of *thi4***

	wTATAAnnnnnnTTATAw	Site scores
MetMK1_23630	CCAT <u><b>GGTTGGAT</b></u> ----- <u><b>TATATAACTAGTTTATAAGCCACTAA</b></u> TATTACGATGTA <b>ACTATG</b> 5.89	
Mcup_0086	TTGTGTCGTTTTTAA <u><b>TATATAACTAGCTTATAAG</b></u> TCGGTTAAATGTTGATGATAT <b>ATG</b> 5.82	
Msed_2221	TTCTCATCGT-CACCTA <u><b>TTTATAACTTAGTTTATAAG</b></u> TGAGTTAAATGAAGATTTGTC <b>ATG</b> 5.98	
Saci_0854	TGATATAAAAATA--- <u><b>CATATAAGTGAGTTTATAAC</b></u> CTATTTAATTGATAAGAC--- <b>ATG</b> 5.60	
YN1551_1140	TTGTCCACGTTTCATGAT <u><b>TATAAAAAGTTAATTTATAAG</b></u> TGAGTTATATCAGTTT <b>AGATT-ATG</b> 5.66	
SSO0436	TTGTCTATGTATAACAT <u><b>TATAAAAAGTTAATTTATAAG</b></u> CAAGTTATATCATT <b>TTAGATT-ATG</b> 5.66	
Ahos_0692	TCTTAAATGA <b>ACTC</b> --- <u><b>TATATAACTTAGTTTATAAG</b></u> TAAGTTAT <b>TATGAAGTAAACACATG</b> 5.91	
	* * * * * * * * * * * * * *	

**Upstream promoter regions of *thi7***

	wTATAAnnnnnnTTATAw	Site scores
MetMK1_19590	TCGCAATAT <u><b>GTTAT</b></u> ----- <u><b>GCTAAATAAGTAAGTTTATAACCGACTTA</b></u> TCATAAATCTGAT-C <b>ATG</b> 5.27	
Saci_2110	TTGAAA---GTTAT-----AGGATTTGA <u><b>ATATAACTAAGTTTATAACT</b></u> TTGCTTATATGAAACTTAATTT <b>ATG</b> 5.21	
YN1551_2514	CTCAGGGCTGATTTGGTGATGGTCTATTA <b>AAAAATTC</b> <u><b>AAAATAAGTGACTTTATAAG</b></u> TACTAATTATCATT <b>TTTACATATG</b> 5.25	
SSO1593	CTCAGGGCTGATTTGGTGATGGTCTATTA <b>AAAAATTC</b> <u><b>AAAATAAGTGACTTTATAAG</b></u> TACTAATTATCATT <b>TTTACATATG</b> 5.25	
SI1132	TTTAAATCTTTTT-----CAGTGGAA <b>TTT</b> <u><b>CATATAAGTAGGTTTATAAG</b></u> TAGTAATATACAATATCA <b>ATATG</b> 5.52	
Ahos_0145	--GAAATCTGATAT-----GCTTT <u><b>CTTATAAGTAAGTTTATAAT</b></u> TATCTTATAACAAGATT <b>TACACATG</b> 5.80	
	* * * * * * * * * * * * * * * * * *	

**Upstream promoter regions of *tenA***

	wTATAAnnnnnnTTATAw	Site scores
MetMK1_24350	---AGCTGACGTCCT <u><b>ATAACTAGATTATAAAATGATTTATATC</b></u> CCGCCACTCTAG-TGGGATGT <b>ATG</b> 5.09	
SSO2089	CATAGATAG----ATATAAGTAA <u><b>GTTATAAAATAATTTATAAT</b></u> CTAGCTTTGTAATTTATA <b>AAATG</b> 5.51	
YN1551_2613	TATGGCTAG----ATATAAGCAA <u><b>CTTATAAAATGATTTATAAT</b></u> CTAGCTTTGTAATTTATAG <b>TTATG</b> 5.40	
	* * * * * * * * * * * * * * * * * *	

Candidate ThiR-binding sites are highlighted in yellow. 18-nt consensus site is given on top of each alignment. Translation start sites are in red. Individual site scores are given at right side. Experimentally determined transcription start sites (TSSs) in *Sulfolobus solfataricus* (SSO0436, SSO1593, SSO2089) (Wurtzel et al., 2010) are underlined/in bold (all three transcripts are leaderless). A TATA-like conserved core promoter motifs, centered ~26 nt upstream of TSSs are highlighted in green. 35-bp DNA fragments from *Metallosphaera yellowstonensis* MK1 that were tested in DNA binding assays are underlined and in bold.

**Figure S3. Phylogenetic footprinting of upstream regions of thiamine biosynthesis/transport genes and identification of candidate ThiR-binding sites in diverse archaeal lineages.**

**B. Thermococcales**

**Upstream promoter regions of *thi4***

	AWAACnnnGTTWT	AWAACnnnGTTWT	Site scores	
Py04_1231	AATACCTTGTATAACAAAGTTTTAAAT	TGATAAACTTAGTTTTGGAAT	-----GGTGAAGAGCATG 5.13/ 5.46	
PH1357	AAAATAATCCTATAACAAAGTTTTAAAT	TATAAAATGAAGTTTTAGGC	-----GGTGGGGAAGATG 5.13/ 4.51	
PAB0536	AAAAGTTCCTATAACAAAGTTTTTAAT	TGCTAAAACCTAAGTTTTTGC	-----GGGGGTAAAGATG 5.10/ 5.40	
PNA2_1941	CAGGGTTCCTATAACAAAGTTTTTAAT	TGCTAAAACCTGAGTTTTCGAT	-----GGTGATGGTGATG 5.13/ 5.30	
PF1530	CCATGATGAAAAAAACAAAGTTTTTAT	TATATAAAACCTCAGTTTTGAGT	-----GGTGAATAAGATG 5.29/ 5.30	
OCC_06556	ATAACTTGGTTATAATAAAGTTTTTAA	ATTCAAAACCTCAGTTTTTACTGG	AGTTTTAAAAGATTTACGGGTGAGGGG	AAAATG 4.73/ 5.30
PYCH_06240	-----GTATAACAAAGTTTTTAAGT	CTCAAAAACCTTAGTTTTGCCTGGT	CTTATAAAATTTTG---GGGGTGCGAAA	ATG 5.46/ 5.13
CL1_0036	AAAACCCCTTTATAACAAAGTTTTTAA	ATTCAAAACCTCAGTTTTTACCGCG	GCTTAAAAGGTTTT---GGGTGGTGGC	ATG 5.13/ 5.30
TGAM_1270	AAAACCCGTTTATAACAAAGTTTTTAA	ATTCAAAACCTTAGTTTTGAGCGGA	ACTTAAAAGGTTTT---GGGTGGTCGG	GATG 5.55/ 5.17
GQS_02015	AAAACCCGTTTATAACAAAGTTTTTAA	ATTCAAAACCTTAGTTTTGAGCGGG	ACTTAAAAGGTTTT---GGGTGGTGG	GATG 5.13/ 5.46
TK0434	TAACCATTTATAACAAAGTTTTTAA	ATCTCAAAAACCTTAGTTTTGAGCG	CGGTTTTAAAAGGTTTT---AGGTGGT	GGGCATG 5.13/ 5.46
	**** *	**** *	* ***	

**Upstream promoter regions of *thiB***

	AWAACnnnGTTWT	AWAACnnnGTTWT	Site scores	
TSIB_0369	-TTGCGAT--GGGGAAGCTTTTTAAAT	CACATTCAGAACACAGTTTATGTATAT	ATAAAATTAGTTTTAGGTGAAGGAG	ATGAAGAAGCT 4.49 / 4.86
PF1751	-----ATCAGGGAAGG-TTTTTAAAC	CTCCCTAGAACACAGTTTTGAAAGA	ATAGAAAATGGGTTTTAGGTGAT	GAGCATGAAGAGTT 4.65 / 3.94
OCC_11612	GCCCCGATGAGGGAAGATTTTTAAAC	CTCCCTAGAACACAGTTTTGGAAGA	ATAGAAAATGGGTTTTAGGTGAT	GAGCATGAAGAGTT 4.65 / 3.94
	** ** * *	***** ** * *	** * ** * ** * ** * *	

Candidate ThiR-binding sites are highlighted in yellow. 15-nt consensus site is given on top of each alignment. Most of the identified ThiR sites in Thermococcales are organized in tandems and have a fixed 21-nt distance between the site centers. Translation start sites are in red. Experimentally determined transcription start sites (TSSs) in *Pyrococcus furiosus* (PF1751, PF1530) (Yoon et al., 2011) are underlined/in bold. ). A TATA-like conserved core promoter motifs, centered ~26 nt upstream of TSSs are highlighted in green. Individual site scores are given at right side.

**Figure S3. Phylogenetic footprinting of upstream regions of thiamine biosynthesis/transport genes and identification of candidate ThiR-binding sites in diverse archaeal lineages.**

**C. Halobacteria**

**Upstream promoter regions of *thi4***

		<b>TAYnACnnnGTnRTAYnACnnnGTnRTA</b>	Site scores
HFX_0634	TCCCGTCAGAAGTCGCC---ATCAGTAGA---TTT	TACTACTGGTGGT <b>ACTAAGTAGT</b> ATGTC	4.69/ 4.65
HVO_0665	CTAGGTCGCCGCGCGCG---ACCACCAGAT---TTT	TACTACTAGGTAGTAT <b>ACTAAGTAGT</b> ATGTC	4.57/ 4.65
Hlac_2980	TTACCCTCCGAGAAACCGACCGACCACAACCCT- <u>TG</u>	TACCACGGGAGAT <b>ACTAAGTGGT</b> ATGAC	3.90/ 4.77
Nmag_2419	TGCGAATC-----ATCGCGACAGACTTTT	CTACTGGATGAT <b>ACCCTAGTTGT</b> ATGAG	3.90/ 4.68
Htur_1462	TATCCGTGGCT-----C-ATCAGTAGACTTTT	CACCACGGGTGCT <b>ACCACCAAGTTGT</b> ATGAG	3.53/ 4.68
Halxa_2863	TATCAGCAGCGCAAC-----ATCAGTAGAGTTT	CACCCTGAGTCT <b>ACCACCTAGTTGT</b> ATGAG	3.37/ 4.68
NJ7G_3680	TTCCGATCGCG-----GCTGCGATAGGCTTTT	CTACTACCAAGT <b>ACCACCTAGTTGT</b> ATGAG	3.90/ 4.68
Natpe_2830	TTCGGGCATCG-----GCTGCGATAGACTTTT	TACTAGGTCA <b>ACCACCTAGTTGT</b> ATGAG	4.08/ 4.68
Hmuk_0002	GTGCAATCGCGTCAGCGAGTCCAGACCGGAACCCTAT- <u>IG</u>	CAACGGTGGT <b>ACCCTAGTTGT</b> ATGAG	3.48/ 4.68
rrnAC1782	TGCGAGCAGCGACTGGTCTCTGCATCAGAACG- <u>ACTATG</u>	TACCACCAAGAGAT <b>ACTAGTTGT</b> ATGAG	3.90/ 4.63
HAH_2313	ATTCTAGCAACAGACGAGGCC--ATCAGAACG- <u>ACTATG</u>	TACCACCAAGAA <b>ACTACTAGTTGGT</b> ATGAG	- / 4.75

**Upstream promoter regions of *thiC***

		<b>TAYnACnnnGTnRTAYnACnnnGTnRTA</b>	Site scores
Hbor_14920	GGTCGGCAACGC-----T	TATTTCGGTAGT <b>ACTAGT</b> AGTTCGATGACG---ACGCAGTTCACACGC	3.96/ 4.41
Halru_0933	TCACCGCTCCACAGGTAT <b>TATGACT</b> TAGTAGT <b>ACTAGT</b> GATC- <u>ATG</u>	CCCGGTACCCAACGCCAGGCAGC	4.10/ 4.41
HVO_2154	CCCGCCCGTCACAAGT <b>TATTA</b> ACTAGT <b>TGT</b> AT <b>ACT</b> AGCTAA <b>TCA</b> <u>ATG</u>	CGCTCGACACAGCTCCAACACGC	3.66/ 3.84
VNG_0715G	CGTCGACCAGTGATT- <u>TTATTA</u> TCTAGTGGT <b>ATC</b> ACTAAGTA <b>ATT</b> GATGTC	---ACACAGCTCCAACACGC	4.06/ 4.40
NP2210A	ACCCGGCAGTAG---TTTT <b>TATTAG</b> CTGGT <b>TGT</b> AT <b>ACT</b> AGATT <b>GTT</b> GATGACG	---ACGCAACTACAGCACGC	3.97/ 3.74
HFX_2195	GCCGTTGTCTATAAAGAT <b>TATCAG</b> TAGTGGT <b>ATA</b> ACTAGTGGT <b>TAT</b> ATG	CGCTCAACTCAAATGCGGTGC	4.24/ 4.43
Hmuk_2297	GACGAGACTGGGAAGTT <b>TATAA</b> TCTAGT <b>GAT</b> ATA <b>ACT</b> AGCTT <b>ATC</b> TATGACG	---ACCCAGCTCCAGCGTGC	3.99/ 3.70
rrnAC1182	GTTCGACCATAGTGTT- <u>TTATAA</u> CTTGGT <b>AGT</b> ATA <b>ACT</b> AGT <b>GAT</b> - <u>ATG</u>	-----ACGCAGATGGCAGCAGC	4.50/ 4.37
HAH_1784	ACCGATCACCGCAGCT- <u>TTATAA</u> CCAAGT <b>AGT</b> ATA <b>ACT</b> AGT <b>GAT</b> - <u>ATG</u>	-----ACGCAGTAGCAGCGGC	4.50/ 4.37
Nmag_2593	GATCGGTGCCACGAAT- <u>TTATAA</u> TCTAGT <b>TAATA</b> CA <b>ACT</b> AGT <b>TATC</b> ATG	GGTCAACCCAGCTAGTGTGC	3.87/ 4.27
Halxa_3896	GTGACGGGACAACAT- <u>TTATAA</u> CTAGAT <b>ATA</b> CA <b>ACT</b> AGT <b>TATC</b> GATG	CGCAACCCAAATGCGCGCG	3.44/ 4.27
Htur_1332	TCCAGCGCAAGCGCT- <u>ATATAA</u> CTCGAT <b>GATA</b> CA <b>ACT</b> AGT <b>TATC</b> GATG	CGCAACCCAGATCCAAGCGC	4.01/ 4.27
NJ7G_3772	CGGAGCCGATAGTGAT- <u>TTATAA</u> TTCAGT <b>TAATA</b> CA <b>ACT</b> AGT <b>TATC</b> GATG	CCGACAGCTCAGATCCAAGCGC	3.91/ 4.27
Natpe_3081	CGGGAGCGGGAGAGTT- <u>TTATAA</u> CTAGT <b>ATA</b> CA <b>ACT</b> AGT <b>TATC</b> GATG	TCGAAAACGCAGATTCAGCGCG	4.50/ 4.27

**Upstream promoter regions of *thiBPQ***

		<b>TAYnACnnnGTnRTAYnACnnnGTnRTA</b>	Site scores
Huta_0388	AACGCGAGAGGTTTAA <b>TACC</b> ACT <b>AGTCA</b> AT <b>CCACCAAGTGGT</b> AAACGACG	AAACGACG	3.84/ 4.44
Htur_2516	ATCAGA-GGGATTAA- <b>GACC</b> AC <b>CCAGTGG</b> AA <b>CCACCGTGAAT</b> GAGACGACG	GAGACGACG	4.21/ 3.87
rrnAC0147	AACAGC-GGGATTAA <b>TACC</b> ACT <b>GGGTGT</b> GA <b>CCACGGTGAATG</b> GATAGACG	GATAGACG	3.86/ 3.49
HAH_0897	AACACC-GGCGTTAAG <b>TACC</b> ACT <b>GGGTGT</b> GA <b>CCACGGAGTAATG</b> GATAGACG	GATAGACG	3.86/ 3.90
HVO_0022	ATCAGC-CCCGTTAAG <b>TACC</b> AC <b>AGTAC</b> AA <b>CAACCGAGTAATG</b> AGACGACG	AGACGACG	3.23/ 4.15
HFX_0021	ATCAGC-GGCGTTAAG <b>TACC</b> AC <b>AGTAC</b> AA <b>CAACCGAGTAATG</b> AGACGACG	AGACGACG	3.23/ 4.13
Nmag_0460	ATCA----ACTTAC <b>AACC</b> AC <b>AGAGTGT</b> AA <b>CCACTGGGTGATG</b> AGACGACG	AGACGACG	3.53/ 4.40
NJ7G_0450	ATCAGC-GGTCTTAA- <b>GACT</b> AC <b>CTAGTGC</b> AA <b>CCACCGGTAATG</b> AGACGACG	AGACGACG	3.32/ 4.26
Natpe_3863	ATCAC--GCCGTTTAA <b>GACC</b> ACT <b>AGTGT</b> AA <b>CCACCGAGTAATG</b> GATCGACG	GATCGACG	3.91/ 4.28
Halxa_0707	ATCACA-GCGTTAA- <b>GACC</b> ACT <b>CACTAG</b> AA <b>CAACCGCGTAATG</b> AGACGACG	AGACGACG	4.11/ 3.78
Hbor_02000	ATGACG-CAGGTTAAG <b>TATG</b> AC <b>CAAGTGG</b> GA <b>CTACTCAGTAATG</b> AGACGACG	AGACGACG	3.55/ 4.15
Hlac_2697	CGCGGATCCTTTTGT <b>AACC</b> AC <b>CGGTGT</b> AA <b>CAACCGGTAATG</b> ACTAACGA	ACTAACGA	3.90/ 4.13
Halru_0146	AAATGCTGGTTCTTAA <b>GACC</b> AC <b>CCAGTGG</b> GA <b>CCATCCAGTAATG</b> AGACGACG	AGACGACG	3.96/ 3.68

Candidate ThiR-binding sites are highlighted in yellow. 15-nt consensus site is given on top of each alignment. Most of the identified ThiR sites in *Halobacteria* are organized in tandems that overlap by 2 nt (the overlapping nucleotides are highlighted in red). Translation start sites are in red. Experimentally determined transcription start sites in *Haloferax volcanii* (*thi4*, *thiC*) (J. Maupin-Furlow, personal communication) are underlined. A TATA-like conserved core promoter motifs, centered ~26 nt upstream of TSSs are highlighted in green. Individual site scores are given at right side. Weaker sites with scores that are 15% less than the threshold (3.85 for *Halobacteria*) are underlined. Consensus AC/GT positions are in bold.

**Figure S3. Phylogenetic footprinting of upstream regions of thiamine biosynthesis/transport genes and identification of candidate ThiR-binding sites in diverse archaeal lineages.**

**D. Thermoproteales – Pyrobaculum spp. (Motif II)**

**Upstream promoter regions of thi4**

		AWAAWnnnWTTWT		Site scores
Pisl_1998	CCCGTCTCTTAGGGCTG-TTCAGCGGTG	AAAAGATATTTAT	CTATACCACATTTATTACTTATGGAGTTAAAAATCG	4.44
Tneu_1052	GCGGCTGTGGTGGGTAGCTTCTGCGTCC	AAAATCTATTTAT	TTGGATTATATTTGTTGCTTATGGAGCTGAAGATCG	4.46
MSUCoA1_02190	GCCGTC-----GTCGCCTTCGAG---	ATAATATATTTAT	AGTCATATCATTTAGTACTTATGGAGCTGAAGATCG	4.96
	* * * * *	* * * * * * * * * * *	* * * * * * * * * * *	

**Common upstream promoter regions of divergently transcribed genes thiT and thi4**

		AWAAWnnnWTTWT		Site scores
PAE0176	<i>thiT</i> <- CAT	AAGCCAATATTC	AAATCTAAATAAAAATATTTATCTCGGCGATTACGCCG-GAGAAAAATTTATTATGAATTCGGATAATTACTTTATG -> <i>thi4</i>	5.16 / 4.79
P186_0937	CAT	AAGCAACTACTTGATTC	CAATAAAAACATTTATCCCCACGGATACGCTGAGAGAAAAATTTATTATAGAAATCGGATAATGTGGTTATG	4.88 / 4.79
	* * * * *	* * * * * * * * * * *	* * * * * * * * * * *	

**Common upstream promoter regions of divergently transcribed genes tenA and thiT**

		AWAAWnnnWTTWT		Site scores
MSUCoA1_02193	<i>tenA</i> <- TCTCCGATGC	CATAAACAGCTAATGAAGTGGTATAAAAAGATTTT	TGGAAAAATATATTTATATGGATTAGATAAAGCTTTATGG -> <i>thiT</i>	4.30 / 4.79
Tneu_1075	TC	CGGTGGCCATATCCCCCTCTGTCTGCACAT-----	AAAAATTTTTTATATGGAAGTCATCAGCCTTTTATGG	5.12
Pisl_1973	TTCCGTTGC	CATATACGTCAGATCTCCGCCTAT-----	AAAAATTTTTTATATGGAAAAGATCAATTAGTTATGG	5.12
	* * * * *	* * * * * * * * * * *	* * * * * * * * * * *	

**Upstream promoter regions of tenA**

		AWAAWnnnWTTWT		Site scores
P186_0930	TAAA-----TCAGTGTA-----	ATAACGTTTTTAT	TATTCAGGGGATAAATGTCCTATGTCGTCGGCTGAGTTGAG	4.62
Pogu_2702	TAAAGCCGCTCTGCGCCTCTGTATGT	CGGCAAAGGATTTTTCATAAAGTTTTTAT	AGGTCCTAGATAAATGTCCTATGCTACGGAGGAGTTGAG	5.01
Pars_2186	TAAAGCCGCTCTGCGCCTCTGTTTGT	CGGCAAAGGATTTTTCATAAAGTTTTTAT	AGGTCCTAGATAAATGTCCTATGCTACGGAGGAGTTGAG	5.01
PAE0170	TA-----CTCCCCCTTTTAT-----	ATAAAAATTTTTAT	CCCGATTAGATAAATTCACTTATGGTAACCTGGGAGTTGAG	5.29
	* * * * *	* * * * * * * * * * *	* * * * * * * * * * *	

**Upstream promoter regions of thiT**

		AWAAWnnnWTTWT		Site scores
Pcal_0348	GAAAAGCCCGCTCTGTTTTCCAAAGCGG	AAAACCCGCGATAAATGTCCTATGTCGTCGGCTGAGTTGAG	---GACAAGTGGACTAGG	4.68
Pars_2195	GAACACACCGATATTTA-----	CGCAGATAAATATTTAAT	AGTAA-TTGGATTAGTTTGTATGCCTGAGAGTTGGAAGAGG	4.36
Pogu_2693	GGCCACGCCATTATCTA-----	CGCAGATAAATATTTAAT	AGTAAATTTGGATTAGTTTATATGCCTGAGAGTTGGAAGAGG	4.16
	* * * * *	* * * * * * * * * * *	* * * * * * * * * * *	

Candidate ThiR-binding sites are highlighted in yellow. 13-nt consensus site is given on top of each alignment. Translation start sites are in red. Individual site scores are given at right side.

**E. Thermoproteales – Thermofilum spp. (Motif III)**

**Upstream promoter regions of ykoEDC**

		TATAACnnnGTTATA		TATAACnnnGTTATA		Site scores
Tpen_0068	ATATCCTTTGAATGCATTAAACCGCTAGAGG----	ATAA	CTTGGTTAAATTGAAAGATTCTTCTCTATAACCGAGTTATA-----TAAAAATA	ATAA		5.02 / 5.43
TCARB_1340	--ATCTTACGACGAGA--AACTTTTAAAGGTAATC	ATAA	CTCTGTTATGTTCTCA---TTTACCTCAAAACTCGGTTATAAAGAAATGAAATA	ATAA		5.18 / 5.16
N186_01265	--ATCTTACGACGAGA--AACTTTTAAAGGTAATC	ATAA	CTCTGTTATGTTCTCA---TTTACCTCAAAACTCGGTTATAAAGAAATGAAATA	ATAA		5.18 / 5.16
	* * * * *	* * * * * * * * * * *	* * * * * * * * * * *	* * * * * * * * * * *		
		TATAACnnnGTTATA				
Tpen_0068	TTGIGCTTAACTTAAAG-----GCTACGGGAAAATT	ATAA	CCAGGTTTGTAGTTTTCAACGCGATGAGGGAAGAAAACTCAGAGTAGAA	ATAA		5.06
TCARB_1340	----CTTCTTCTTAAATGGTGGAAAGTCAAACGTTAAGCA	ATAA	ATCAGTTTAAAGAAATAAAA---AATGATGAACGAAAACTAAGAATTGAA	ATAA		5.22
N186_01265	----CTTCTTCTTAAATGGTGGAAAGTCAAACGTTAAGCA	ATAA	ATCAGTTTAAAGAAATAAAA---AATGATGAACGAAAACTAAGAATTGAA	ATAA		5.22
	* * * * *	* * * * * * * * * * *	* * * * * * * * * * *	* * * * * * * * * * *		

Candidate ThiR-binding sites are highlighted in yellow. 15-nt consensus site is given on top of each alignment. All identified ThiR sites in *Thermofilum* spp. are organized in tandems of three sites with variable distance between sites. Translation start sites are in red. Individual site scores are given at right side.



DNA-binding domain, XRE family (COG2522)



**Figure S4. Alignment of the thiamine regulators ThiR in Archaea.** MK1\_29690 denotes the characterized in this work ThiR protein from *Metallosphaera yellowstonensis*. Mk1. C-terminal domain in ThiR is aligned with thiamine phosphate synthase domain ThiN from *Pyrococcus furiosus*. (PDB ID 2PB9). Conserved residues in ThiR and ThiN proteins are marked with asterisk. Red asterisks indicate conserved residues that are presumably involved in thiamine phosphate binding, they were also bound phosphoric acid in 2PB9.

N298

R320

S339 H341

ThiN/2PB9	IPEVGTNFVYSLPL/RSTKDVAGVGRIVKYGNSVAVGPEVFCASDHLARVLTMYRFYEVRSAINIRYSREIIEEIEIAQERGFKVSFYDRREPEEIKAKEGATIPWG/ERPDIYIHLGDVKGEMILVFGNRPVLEKIKMLIE	452
MK1_29690	VPEVLSNLAFAKSRNQSVDVAVAGRIAVIGGIPTPASRPTWGGSRHLATILLRLALKVCGKWRVAVMNIKFDRLVEQALKTAGLVYSKVGPSQDKDDRIAQMVASSLGDC----	286
Msed_2166	IPEILSNLAYSKRDAEILDVLAIPGRIAVIGGVPTPASRPTWGGSRHLATILLEIRKRCERWRSVMNIKYDEKVEAILKSGRLRVKVPSSDRRDDSIANMVASVSSC----	298
Mcup_0124	IPEVLSNLAISKRDADKADVLAIPGRIAVIGGLPTPASKPTWGGSRHLATILLEVRKCSKWRVSMNIKLYDCIEKALKQANLSKVVGPSDDKRDVSIARIVAAVMND----	286
Ahos_0963	IPEVLSNLAIAKRNPKSLEDIIVAGRIIKIRNLPPTASKPMWAGSKHLNILLVSRKYPIRSVMNIKYDEKIKKCMDELKMKYAVIAPHDYATDIIETIASVMVFRPC----	298
SS00468	IPEVLSNLAFAKANAKREEDVLAIPGRIIKIRNLPPTASKPMWAGSKHLAKVLLSVMKNHPAVRSVMNIKYDEKVEAILKSGRLRVKVPSSDRRDDSIANMVASVSSC----	306
YN1551_1174	IPEVLSNLAFAKANAKREEDVLAIPGRIIKIRNLPPTASKPMWAGSKHLAKVLLSVMKNHPAVRSVMNIKYDEKVEAILKSGRLRVKVPSSDRRDDSIANMVASVSSC----	306
ST0179	IPEVLSNLAFAKANPKGINDVLAIPGRIIKVKGIPPTASKPMWAGSKHLAKVLLSVMKNHPAVRSVMNIKYDEKVEAILKSGRLRVKVPSSDRRDDSIANMVASVSSC----	313
Saci_0437	IQVLSNLAIAEQAKNEEDVLAIPGRIIRVLTGVPPTASKPMRAGSKHLRILLIKIKRRIRAVMNIKYDETKTFTVSOQLKVVVYVPHDYATNDEIAKIEVDAFDFTD----	298
N186_01270	IPEVMSNLAIAVPAKSIYDVAAPGRIISRKTNVEIYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	309
Tpen_0069	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	301
Vdis_2360	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	301
VMUT_0712	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	305
Cmaq_1034	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	313
Pars_1934	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	287
Pogu_0189	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	287
P186_0445	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	287
PAE2210	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	287
Pis1_0839	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	287
Tneu_1384	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	287
MSUCoA1_00968	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	287
Pca1_0941	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	287
TTX_0743	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	290
TUZN_0486	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	291
Igag_0434	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	300
FBUT_0125	VPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	300
FFont_0578	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	300
Desfe_0411	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	287
DKAM_0343	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	287
Desmu_0750	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	288
Tagg_0649	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	288
TCELL_1104	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	288
Smar_1383	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	288
Shell_1065	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	288
Htur_2542	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	303
Halxa_0732	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	303
NJ7G_1102	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	303
Natpe_1684	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	303
Nmag_0434	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	303
rrnA1_7757	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	321
HAH_2294	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	303
Huta_2411	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	301
VNG0247C	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	302
OE13857	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	302
rrnA1_1775	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	298
HAH_2305	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	298
Hmuk_0022	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	298
Huta_2715	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	297
NP51684	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	298
Htur_1469	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	299
Halxa_2856	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	299
Nmag_2413	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	299
NJ7G_3692	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	299
Natpe_2824	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	299
Halru_1158	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	299
HVO_0662	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	299
HFX_0632	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	299
Hbor_25480	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	301
PH0716	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	294
PAB0877	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	294
PNA2_1361	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	294
Py04_0745	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	295
PF0601	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	293
PYCH_07480	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	294
TS1B_0198	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	296
OCC_09391	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	296
TERMP_01433	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	299
TOÑ_1436	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	294
GQS_08390	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	294
CL1_1295	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	294
TGAM_2087	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	293
TAM4_328	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	293
TKZ197	IPEVLSNLAIAVPAKSIYDVAAPGRIISRKSGCELRYRPFEGASRHLGGILLISNG---KFRAVMNIKYDEVIRSELISMGISFKFESSEDPANPAASAAADLLEKC----	293



**Table S1. (A) Distribution of thiamin biosynthesis / salvage and ThiR genes in archaeal genomes.**

Genomes / Lineage (number of genomes per lineage)	Pheno-type <sup>1</sup>	Thiamine Repressor ThiR <sup>2</sup>	Thiamine biosynthesis and salvage enzymes <sup>2</sup>							
			ThiC	Thi4	ThiM	ThiE	ThiD	ThiDN	TenA	
<b>Crenarchaeota / Thermoprotei Sulfolobales (8)</b>			<b>ThiR</b>	<b>ThiC</b>	<b>Thi4</b>			<b>ThiDN</b>	<b>TenA</b>	
<i>Metallosphaera yellowstonensis</i> MK1*	Ah&T	MetMK1_29690	-	&MetMK1_23630	-	-	-	MetMK1_11590	&MetMK1_24350-40	
<i>Metallosphaera cuprina</i> Ar-4	P&T	Mcup_0124	Mcup_1215	&Mcup_0086	-	-	-	Mcup_1524	Mcup_0881	
<i>Metallosphaera sedula</i> DSM 5348	P&T	Msed_2166	&Msed_0906	&Msed_2221	-	-	-	Msed_0646	Msed_1351	
<i>Sulfolobus acidocaldarius</i> DSM 639	P&T	Saci_0437	&Saci_0137	&Saci_0854	-	-	-	Saci_0953	Saci_2283; Saci_1230	
<i>Sulfolobus islandicus</i> Y.N.15.51	P&T	YN1551_1174	YN1551_2484	&YN1551_1140	-	-	-	YN1551_2791	&YN1551_2613; YN1551_3080	
<i>Sulfolobus solfataricus</i> P2	P&T	SSO0468	SSO1941; SSO1324	&SSO0436	-	-	-	SSO2393	&SSO2089; SSO2598	
<i>Sulfolobus tokodaii</i> str. 7	P&T	ST0179	ST1864	&ST0361	-	-	-	ST0526	ST1041; ST1007	
<i>Acidianus hospitalis</i> W1	P&T	Ahos_0963	&Ahos_1208	&Ahos_0692	-	-	-	Ahos_2188	Ahos_2099	
<b>Thermoproteales (13)</b>			<b>ThiR</b>	<b>ThiC</b>	<b>Thi4</b>	<b>ThiM</b>	<b>ThiE</b>	<b>ThiDN</b>	<b>TenA</b>	
<i>Caldivirga maquilingensis</i> IC-167	Ah&T	Cmaq_1034	-	&Cmaq_1420	&Cmaq_0062	Cmaq_0061	-	&Cmaq_0452	Cmaq_0480	
<i>Pyrobaculum aerophilum</i> str. IM2	P&T	PAE2210	&PAE0333	&PAE0175	-	-	-	&PAE2535	&PAE0170-1	
<i>Pyrobaculum arsenaticum</i> DSM 13514	Ah&T	Pars_1934	-	Pars_1498	-	-	-	Pars_0839	&Pars_2186-7	
<i>Pyrobaculum caldifontis</i> JCM 11548	Ah&T	Pcal_0941	-	Pcal_1555	-	-	-	Pcal_0240	&Pcal_1316; &Pcal_1890	
<i>Pyrobaculum islandicum</i> DSM 4184	P&T	Pisl_0839	&Pisl_1450	&Pisl_1998	-	-	-	&Pisl_1419	&Pisl_1975-4	
<i>Pyrobaculum oguniense</i> TE7	P&T	Pogu_0189	Pogu_2748	Pogu_0708	-	-	-	Pogu_1494	&Pogu_2702-01	
<i>Pyrobaculum</i> sp. 1860	Ah&T	P186_0445	-	&P186_0936	-	-	-	P186_2074	&P186_0930-1	
<i>Pyrobaculum yellowstonensis</i> WP30	P&T	MSUCoA1_00968	&MSUCoA1_00516	&MSUCoA1_02190	-	-	-	MSUCoA1_02233	&MSUCoA1_02191-2	
<i>Thermoproteus neutrophilus</i> V24Sta	P&T	Tneu_1384	&Tneu_0530	&Tneu_1052	-	-	-	&Tneu_1881	&Tneu_1073-4	
<i>Thermoproteus tenax</i> Kra 1	P&T	TTX_0743	&TTX_0287	&TTX_0666	-	-	-	TTX_0857	TTX_0325-24	
<i>Thermoproteus uzoniensis</i> 768-20	Ah&T	TUZN_0486	-	&TUZN_0079	-	-	-	TUZN_0613	TUZN_0831-30	
<i>Vulcanisaeta distributa</i> DSM 14429	P&T	Vdis_2360	Vdis_0190	&Vdis_2373	Vdis_0394	Vdis_1205	-	&Vdis_2402	Vdis_2464; Vdis_0560	
<i>Vulcanisaeta mouthovskia</i> 768-28	Ah&T	VMUT_0712	-	&VMUT_0723	VMUT_1232	VMUT_2053	-	&VMUT_0747	VMUT_0803; VMUT_1030	
<b>Thermoproteales - Thermofilum (4)</b>			<b>ThiR</b>							
<i>Thermofilum carboxyditrophus</i> 1505	A&T	&TCARB_1341	-	-	-	-	-	-	-	
<i>Thermofilum pendens</i> Hrk5	A&T	&Tpen_0069	-	-	-	-	-	-	-	
<i>Thermofilum</i> sp. 1807-2	A&T	&MA03_01130	-	-	-	-	-	-	-	
<i>Thermofilum</i> sp. 1910b	A&T	&N186_01270	-	-	-	-	-	-	-	
<b>Desulfurococcales (13)</b>			<b>ThiR</b>	<b>ThiC</b>	<b>Thi4</b>	<b>ThiM</b>	<b>ThiE</b>	<b>ThiD</b>	<b>ThiDN</b>	<b>TenA</b>
<i>Aeropyrum camini</i> SY1	Ah&T	-	-	ACAM_1344	-	-	-	ACAM_1502	ACAM_0495	
<i>Aeropyrum pernix</i>	Ah&T	-	-	APE_2149	-	-	-	APE_2400	APE_0669	
<i>Desulfurococcus fermentans</i> DSM 16532	P&T	Desfe_0411	&Desfe_0198	&Desfe_0795	-	-	-	&Desfe_0794	-	
<i>Desulfurococcus kamchatkensis</i> 1221n	P&T	DKAM_0343	&DKAM_0122	&DKAM_0713	-	-	-	&DKAM_0714	-	
<i>Desulfurococcus mucosus</i> DSM 2162	A&T	Desmu_0750	-	-	-	-	-	Desmu_0544	-	
<i>Hyperthermus butylicus</i> DSM 5456	P	Hbut_0125	&Hbut_0289	&Hbut_0253	-	-	-	Hbut_1459	-	
<i>Ignicoccus hospitalis</i> KIN4/I	P	-	Igni_0104	Igni_0560	-	-	-	Igni_0888	-	
<i>Ignisphaera aggregans</i> DSM 17230	P	Igag_0434	&Igag_1298	Igag_0295	-	-	-	&Igag_1669	-	
<i>Pyrolobus fumarii</i> 1	P	-	Pyrfu_1448	Pyrfu_0366	-	-	-	Pyrfu_0689	-	
<i>Staphylothermus hellenicus</i> DSM 12710	Ah&T	Shell_1065	-	&Shell_0764	-	-	-	&Shell_0767	Shell_0762	
<i>Staphylothermus marinus</i> F1	Ah&T	Smar_1383	-	&Smar_0052	-	-	-	&Smar_0049	Smar_0054	

<i>Thermogladius cellulolyticus</i> 1633	P&T	TCELL_1104	&TCELL_1013	&TCELL_1023	-	-	-	TCELL_1024	-
<i>Thermosphaera aggregans</i> DSM 11486	As&T	Tagg_0649	-	-	Tagg_0775	Tagg_0776	&Tagg_0777	-	-
<b>Acidilobales (2)</b>								<b>ThiDN</b>	<b>TenA</b>
<i>Acidilobus saccharovorans</i> 345-15	A	-	-	-	-	-	-	ASAC_1274	ASAC_0493
<i>Caldisphaera lagunensis</i> DSM 15908	A&T	-	-	-	-	-	-	Calag_0858	Calag_0587
<b>Fervidicoccales (1)</b>		<b>ThiR</b>		<b>Thi4</b>				<b>ThiDN</b>	<b>TenA</b>
<i>Fervidococcus fontis</i> Kam940	Ah&T	&FFONT_0578	-	&FFONT_0392	-	-	-	FFONT_0261	FFONT_0756
<b>Euryarchaeota</b>									
<b>Thermococci (15)</b>		<b>ThiR</b>	<b>ThiC</b>	<b>Thi4</b>	<b>ThiM</b>	<b>ThiE</b>	<b>YjbQ</b>	<b>ThiDN</b>	<b>TenA</b>
<i>Pyrococcus abyssi</i> GE5	P&T	PAB0877	&PAB1930	&PAB0536	PAB2432	PAB1645	PAB0878	PAB1646	&PAB1641-2
<i>Pyrococcus furiosus</i> DSM 3638	P&T	PF0601	&PF1531	&PF1530	PF1335	PF1334	PF0680	PF1333	&PF1338-7
<i>Pyrococcus horikoshii</i> OT3	Ah&T	PH0716	-	&PH1357	PH1157	PH1156	PH0713	PH1155	&PH1161-0
<i>Pyrococcus</i> sp. NA2	P	PNA2_1361	&PNA2_0094	&PNA2_1941; &PNA2_1583	PNA2_1766	PNA2_1765	PNA2_1360	PNA2_1764	&PNA2_1769-8
<i>Pyrococcus</i> sp. ST04	Ah&T	Py04_0745	-	&PNA2_1941; &PNA2_1583	Py04_0937	Py04_0936	Py04_0747	Py04_0935	&Py04_0940-39
<i>Pyrococcus yayanosii</i> CH1	P&T	PYCH_07480	&PYCH_06250	&PYCH_06240; &PYCH_08650	-	-	PYCH_07470	PYCH_04500	-
<i>Thermococcus barophilus</i> MP	As&T	&TERMP_01433	-	-	TERMP_00506	TERMP_00507	TERMP_01431	TERMP_00508	-
<i>Thermococcus gammatolerans</i> EJ3	P&T	TGAM_2087	&TGAM_1269	&TGAM_1270	-	-	TGAM_2086	TGAM_1271	-
<i>Thermococcus kodakaraensis</i> DSM 3638	P&T	&TK2197	&TK0433	&TK0434	-	-	TK2198	TK0435	-
<i>Thermococcus litoralis</i> DSM 5473	P&T	&OCC_09391	&OCC_06561	&OCC_06556	OCC_04979	OCC_04974	OCC_09401	OCC_04964	-
<i>Thermococcus onnurineus</i> NA1	A&T	&TON_1436	-	-	-	-	TON_1434	-	-
<i>Thermococcus sibiricus</i> MM 739	As&T	&TSIB_0198	-	-	TSIB_0465	TSIB_0464	TSIB_0200	TSIB_0462	TSIB_0466
<i>Thermococcus</i> sp. 4557	P&T	&GQS_08390	&GQS_02010	&GQS_02015	-	-	GQS_08395	GQS_02020	-
<i>Thermococcus</i> sp. AM4	A&T	TAM4_328	-	-	-	-	TAM4_136	-	-
<i>Thermococcus</i> sp. CL1	P&T	&CL1_1295	&CL1_0038	&CL1_0036	-	-	CL1_1296	CL1_0035	-
<b>Halobacteria (17)</b>		<b>ThiR</b>	<b>ThiC</b>	<b>Thi4</b>	<b>ThiM</b>	<b>ThiE</b>	<b>ThiD</b>	<b>ThiDN</b>	<b>TenA</b>
<i>Haloarcula hispanica</i> ATCC 33960	P&T	HAH_2305; HAH_2294	&HAH_1784	&HAH_2313	-	HAH_2625	-	HAH_2882	&HAH_2295-6
<i>Haloarcula marismortui</i> ATCC 43049	P&T	rrnAC1777; rrnAC1764	&rrnAC1182	&rrnAC1782	-	rrnAC2143	-	rrnAC2449	&rrnAC1765-7
<i>Halobacterium</i> sp. NRC-1	P&T	VNG0247C	&VNG0715G	&VNG2604G	-	-	-	VNG2606G	-
<i>Haloferax mediterranei</i> ATCC 33500	P&T	HFX_0632	&HFX_2195	&HFX_0634	HFX_2684	HFX_2685	HFX_2683	-	&HFX_5296-5
<i>Haloferax volcanii</i> DS2	P&T	HVO_0662	&HVO_2154	&HVO_0665	HVO_2667	HVO_2668	HVO_2666	-	HVO_B0381-80
<i>Halo geometricum borinquense</i> DSM 11551	P&T	Hbor_25480	&Hbor_14920	&Hbor_25460	Hbor_34630	Hbor_34620	Hbor_34650	-	&Hbor_30560-70
<i>Halomicrobium mukohataei</i> DSM 12286	P&T	Hmuk_1352; Hmuk_0022	&Hmuk_2297	&Hmuk_0002	-	Hmuk_0403	-	Hmuk_0364	-
<i>Halopiger xanaduensis</i> SH-6	P&T	Halxa_0732; Halxa_2856	&Halxa_3896	&Halxa_2863	Halxa_1353	Halxa_1354	-	Halxa_2106	&Halxa_2795
<i>Haloquadratum walsbyi</i> DSM 16790	Ah&T	HQ1277A	-	&HQ1276A	HQ2656A	HQ2655A	HQ2657A	-	&HQ2369A
<i>Halorhabdus utahensis</i> DSM 12940	P&T	Huta_2715; Huta_2411	&Huta_2472	&Huta_2718	-	Huta_1463	-	Huta_0491	&Huta_0684-5
<i>Halorubrum lacusprofundi</i> ATCC 49239	Ah&T	Hlac_0446	-	&Hlac_2980	-	-	-	Hlac_2988	Hlac_2987-6
<i>Haloterrigena turkmenica</i> DSM 5511	P&T	Htur_1469; Htur_2542	&Htur_1332	&Htur_1462	Htur_0247	Htur_0246	-	Htur_3232	Htur_2707
<i>Halovivax ruber</i> XH-70	P&T	Halru_1158	&Halru_0933	&Halru_1150	Halru_1513	Halru_1512	Halru_0514	-	Halru_0039
<i>Natrialba magadii</i> ATCC 43099	P&T	Nmag_0434; Nmag_2413	&Nmag_2593	&Nmag_2419	Nmag_1810	Nmag_1811	-	Nmag_1282	Nmag_0869
<i>Natrinema pellirubrum</i> DSM 15624	P&T	Natpe_1684; Natpe_2824	&Natpe_3081	&Natpe_2830	Natpe_3728	&Natpe_3729	-	Natpe_0691	-
<i>Natrinema</i> sp. J7-2	P&T	NJ7G_1102; NJ7G_3692	&NJ7G_3772	&NJ7G_3680	NJ7G_0298	&NJ7G_0299	-	NJ7G_2232	&NJ7G_3269
<i>Natronomonas pharaonis</i> DSM 2160	P	NP5168A	&NP2210A	&NP5174A	NP4052A	&NP4054A	-	NP0546A	NP4082A-80A

<sup>1</sup> Predicted TBP phenotypes: P, B1 prototroph; A, B1 auxotroph, As, B1 auxotroph with HMP/HEP precursor salvage; Ah, HMP auxotroph; T, B1/precursor transporter is present.

<sup>2</sup> Predicted ThiR-regulated genes are marked in red, genes that are preceded by candidate ThiR-sites are marked with '&'

**Table S1. (B) Distribution of thiamin-related transporters, ThiL and ThiR genes in archaeal genomes.**

Genomes / Lineage (number of genomes per lineage)	Pheno-type <sup>1</sup>	Thiamine Repressor ThiR <sup>2</sup>	Thiamine phosphate kinase ThiL	Uptake transporters for thiamine and its precursors <sup>2</sup>		
				B1 / HMP transporters	HMP transporter	HET transporter
<b>Crenarchaeota / Thermoprotei</b>		<b>ThiR</b>	<b>ThiL</b>	<b>ThiT</b>		
<b>Sulfolobales (8)</b>						
<i>Metallosphaera yellowstonensis</i> MK1*	Ah&T	MetMK1_29690	MetMK1_22900	&MetMK1_19590	-	-
<i>Metallosphaera cuprina</i> Ar-4	P&T	Mcup_0124	Mcup_0019	Mcup_1817	-	-
<i>Metallosphaera sedula</i> DSM 5348	P&T	Msed_2166	Msed_0019	Msed_0254	-	-
<i>Sulfolobus acidocaldarius</i> DSM 639	P&T	Saci_0437	Saci_0704	&Saci_2110	-	-
<i>Sulfolobus islandicus</i> Y.N.15.51	P&T	YN1551_1174	YN1551_0944	&YN1551_2514	-	-
<i>Sulfolobus solfataricus</i> P2	P&T	SSO0468	SSO0237	&SSO1593	-	-
<i>Sulfolobus tokodaii</i> str. 7	P&T	ST0179	ST0287	&ST1132	-	-
<i>Acidianus hospitalis</i> W1	P&T	Ahos_0963	Ahos_0762	&Ahos_0145	-	-
<b>Thermoproteales (13)</b>		<b>ThiR</b>	<b>ThiL</b>	<b>ThiT</b>		
<i>Caldivirga maquilingensis</i> IC-167	Ah&T	Cmaq_1034	Cmaq_1646	Cmaq_1199 , &Cmaq_0759	-	-
<i>Pyrobaculum aerophilum</i> str. IM2	P&T	PAE2210	PAE1022	&PAE0176	-	-
<i>Pyrobaculum arsenaticum</i> DSM 13514	Ah&T	Pars_1934	Pars_0185	&Pars_2195	-	-
<i>Pyrobaculum caldifontis</i> JCM 11548	Ah&T	Pcal_0941	Pcal_0657*	&Pcal_0348	-	-
<i>Pyrobaculum islandicum</i> DSM 4184	P&T	Pisl_0839	Pisl_1202	&Pisl_1973	-	-
<i>Pyrobaculum oguniense</i> TE7	P&T	Pogu_0189	Pogu_2294	&Pogu_2693	-	-
<i>Pyrobaculum</i> sp. 1860	Ah&T	P186_0445	P186_2443	&P186_0937	-	-
<i>Pyrobaculum yellowstonensis</i> WP30	P&T	MSUCoA1_00968	MSUCoA1_00352	&MSUCoA1_02193	-	-
<i>Thermoproteus neutrophilus</i> V24Sta	P&T	Tneu_1384	Tneu_0170	&Tneu_1075	-	-
<i>Thermoproteus tenax</i> Kra 1	P&T	TTX_0743	TTX_2042	TTX_1146	-	-
<i>Thermoproteus uzoniensis</i> 768-20	Ah&T	TUZN_0486	TUZN_1441	TUZN_0889	-	-
<i>Vulcanisaeta distributa</i> DSM 14429	P&T	Vdis_2360	Vdis_0764	&Vdis_2303, &Vdis_2247	-	-
<i>Vulcanisaeta mouthovskia</i> 768-28	Ah&T	VMUT_0712	VMUT_1612	&VMUT_0613	-	-
<b>Thermoproteales - Thermofilum (4)</b>		<b>ThiR</b>	<b>ThiL</b>	<b>YkoEDC</b>		
<i>Thermofilum carboxyditrophus</i> 1505	A&T	&TCARB_1341	TCARB_1738	&TCARB_1340-38	-	-
<i>Thermofilum pendens</i> Hrk5	A&T	&Tpen_0069	Tpen_0476	&Tpen_0068-66	-	-
<i>Thermofilum</i> sp. 1807-2	A&T	&MA03_01130	MA03_06765	&MA03_01135-45	-	-
<i>Thermofilum</i> sp. 1910b	A&T	&N186_01270	N186_03355	&N186_01265-55	-	-
<b>Desulfurococcales (13)</b>		<b>ThiR</b>	<b>ThiL</b>	<b>YkoEDC</b>	<b>CytX</b>	
<i>Aeropyrum camini</i> SY1	Ah&T	-	ACAM_0179	ACAM_1349; ACAM_1350-1	-	-
<i>Aeropyrum pernix</i>	Ah&T	-	APE_0235	APE_2154-5^	-	-
<i>Desulfurococcus fermentans</i> DSM 16532	P&T	Desfe_0411	Desfe_0567	&Desfe_0090; &Desfe_0187-6	-	-

<i>Desulfurococcus kamchatkensis</i> 1221n	P&T	DKAM_0343	DKAM_0574	&DKAM_1465; &DKAM_0109-8	-	-
<i>Desulfurococcus mucosus</i> DSM 2162	A&T	Desmu_0750	Desmu_1220	&Desmu_0176; &Desmu_0257-6	&Desmu_0542	-
<i>Hyperthermus butylicus</i> DSM 5456	P	Hbut_0125	Hbut_0793	-	-	-
<i>Ignicoccus hospitalis</i> KIN4/I	P	-	Igni_0219	-	-	-
<i>Ignisphaera aggregans</i> DSM 17230	P	Igag_0434	Igag_1192	-	-	-
<i>Pyrolobus fumarii</i> 1	P	-	Pyrfu_1719	-	-	-
<i>Staphylothermus hellenicus</i> DSM 12710	Ah&T	Shell_1065	Shell_0827	&Shell_0763; &Shell_0766-5	&Shell_0907	-
<i>Staphylothermus marinus</i> F1	Ah&T	Smar_1383	Smar_1576	&Smar_0053; &Smar_0050-1	-	-
<i>Thermogladius cellulolyticus</i> 1633	P&T	TCELL_1104	TCELL_1258	&TCELL_0952; &TCELL_0944-5	-	-
<i>Thermosphaera aggregans</i> DSM 11486	As&T	Tagg_0649	Tagg_0467	&Tagg_0461; &Tagg_0459-60	-	-
<b>Acidilobales (2)</b>			<b>ThiI</b>	<b>ThiT</b>		
<i>Acidilobus saccharovorans</i> 345-15	A	-	ASAC_1266	-	-	-
<i>Caldisphaera lagunensis</i> DSM 15908	A&T	-	Calag_0867	Calag_0174	-	-
<b>Fervidicoccales (1)</b>		<b>ThiR</b>	<b>ThiI</b>	<b>YkoEDC</b>		
<i>Fervidococcus fontis</i> Kam940	Ah&T	&FFONT_0578	FFONT_0963	&FFONT_0579; &FFONT_0581-0	-	-
<b>Euryarchaeota</b>						
<b>Thermococci (15)</b>		<b>ThiR</b>	<b>ThiI</b>	<b>ThiBPQ</b>	<b>CytX</b>	<b>ThiW</b>
<i>Pyrococcus abyssi</i> GE5	P&T	PAB0877	PAB2358	&PAB1835; PAB0543-45	PAB1643	-
<i>Pyrococcus furiosus</i> DSM 3638	P&T	PF0601	PF1877	&PF1751; &PF1748-50. PF1518; PF1520-19	PF1336	-
<i>Pyrococcus horikoshii</i> OT3	Ah&T	PH0716	PH1833	&PH1349; PH1352-50	PH1159	-
<i>Pyrococcus</i> sp. NA2	P	PNA2_1361	PNA2_0432	-	PNA2_1767	-
<i>Pyrococcus</i> sp. ST04	Ah&T	Py04_0745	Py04_1764	Py04_1237; Py04_1235-6	Py04_0938	-
<i>Pyrococcus yayanosii</i> CH1	P&T	PYCH_07480	PYCH_00820	PYCH_10760; PYCH_10780-70	-	-
<i>Thermococcus barophilus</i> MP	As&T	&TERMP_01433	TERMP_00381	&TERMP_01742; &TERMP_01606-07	&TERMP_00505	-
<i>Thermococcus gammatolerans</i> EJ3	P&T	TGAM_2087	TGAM_0436	&TGAM_1706; TGAM_1704-5	-	-
<i>Thermococcus kodakaraensis</i> DSM 3638	P&T	&TK2197	TK0890	TK0757; &TK0159-8	-	-
<i>Thermococcus litoralis</i> DSM 5473	P&T	&OCC_09391	OCC_04465	&OCC_11612; OCC_11622-17	&OCC_04984	OCC_04969
<i>Thermococcus onnurineus</i> NA1	A&T	&TON_1436	TON_0419	&TON_0963; &TON_1959-8	-	-
<i>Thermococcus sibiricus</i> MM 739	As&T	&TSIB_0198	TSIB_0584	&TSIB_0369; TSIB_0367-8	&TSIB_0467	TSIB_0463
<i>Thermococcus</i> sp. 4557	P&T	&GQS_08390	GQS_03115	&GQS_10685; &GQS_05315-20	-	-
<i>Thermococcus</i> sp. AM4	A&T	TAM4_328	TAM4_510	&TAM4_1850; TAM4_1694-35	-	-
<i>Thermococcus</i> sp. CL1	P&T	&CL1_1295	CL1_0270	CL1_1711; &CL1_0708-9	-	-
<b>Halobacteria (17)</b>		<b>ThiR</b>	<b>ThiI</b>	<b>ThiBPQ</b>	<b>ThiV</b>	
<i>Haloarcula hispanica</i> ATCC 33960	P&T	HAH_2305; HAH_2294	HAH_0441	&HAH_0897-95	HAH_2297	-
<i>Haloarcula marismortui</i> ATCC 43049	P&T	rrnAC1777; rrnAC1764	rrnAC3177	&rrnAC0147-45	rrnAC1769	-
<i>Halobacterium</i> sp. NRC-1	P&T	VNG0247C	VNG2011G	&VNG2395-97C; VNG1879G	-	-



<i>Haloferax mediterranei</i> ATCC 33500	P&T	HFX_0632	HFX_1943	&HFX_0021-19	HFX_5294	-
<i>Haloferax volcanii</i> DS2	P&T	HVO_0662	HVO_1861	&HVO_0022-20	HVO_B0379	-
<i>Halogeometricum borinquense</i> DSM 11551	P&T	Hbor_25480	Hbor_04620	&Hbor_02000-10; Hbor_18130	Hbor_30580	-
<i>Halomicrobium mukohataei</i> DSM 12286	P&T	Hmuk_1352; Hmuk_0022	Hmuk_1297	Hmuk_1933-35	-	-
<i>Halopiger xanaduensis</i> SH-6	P&T	Halxa_0732; Halxa_2856	Halxa_3467	&Halxa_0707-05; Halxa_1431	-	-
<i>Haloquadratum walsbyi</i> DSM 16790	Ah&T	HQ1277A	HQ1102A	-	HQ2370A	-
<i>Halorhabdus utahensis</i> DSM 12940	P&T	Huta_2715; Huta_2411	Huta_1979	&Huta_0388-90	Huta_0686	-
<i>Halorubrum lacusprofundi</i> ATCC 49239	Ah&T	Hlac_0446	Hlac_2151	&Hlac_2697-95	-	-
<i>Haloterrigena turkmenica</i> DSM 5511	P&T	Htur_1469; Htur_2542	Htur_3484	&Htur_2516-14	-	-
<i>Halovivax ruber</i> XH-70	P&T	Halru_1158	Halru_0115	&Halru_0146-48	-	-
<i>Natrialba magadii</i> ATCC 43099	P&T	Nmag_0434; Nmag_2413	Nmag_1515	&Nmag_0460-62; Nmag_1940	-	-
<i>Natrinema pellirubrum</i> DSM 15624	P&T	Natpe_1684; Natpe_2824	Natpe_0446	&Natpe_3863-65	-	-
<i>Natrinema</i> sp. J7-2	P&T	NJ7G_1102; NJ7G_3692	NJ7G_2637	&NJ7G_0450-52	-	-
<i>Natronomonas pharaonis</i> DSM 2160	P	NP5168A	NP2280A	-	-	-

<sup>1</sup> Predicted TBP phenotypes: P, B1 prototroph; A, B1 auxotroph, As, B1 auxotroph with HMP/HEP precursor salvage; Ah, HMP auxotroph; T, B1/precursor tr

<sup>2</sup> Predicted ThiR-regulated genes are marked in red, genes that are preceded by candidate ThiR-sites are marked with '&'

**Table S2. Predicted ThiR-binding sites and reconstructed regulons in archaeal genomes.**

Genome / Lineage (Motif type)	thiR gene locus_tag	Target operon <sup>1</sup>	First gene locus tag	ThiR-binding site and consensus <sup>2</sup>			Distance between tandem ThiR boxes	TSS position <sup>5</sup>
				Position <sup>3</sup>	Score	Sequence		
<b>Sulfolobales (Motif-I)</b>				<b>Threshold=5.1</b>			<b>nwTATAAnnnnnTTATAwn</b>	
<i>Metallosphaera yellowstonensis</i> MK1	MetMK1DRAFT_00029690	<i>thi4</i>	MetMK1DRAFT_00023630	-41	5.89	taTATAAcTAgTTTATAAG	validated	
		<i>thiT</i>	MetMK1DRAFT_00019590	-42	5.27	taaATAAgTaAgTTTATAAc	validated	
		<i>tenA1-tenA2</i>	MetMK1DRAFT_00024350	-39	<u>5.09</u>	aTTATAAaTgATTTATAc	validated	
<i>Metallosphaera sedula</i> DSM 5348	Msed_2166	<i>thi4</i>	Msed_2221	-43	5.98	tTTATAAcTTAgTTTATAAG		
		<i>thiC</i>	Msed_0906	-17	5.75	CTTATAAcTTAcTTTATAAT		
<i>Metallosphaera cuprina</i> Ar-4	Mcup_0124	<i>thi4</i>	Mcup_0086	-41	5.82	taTATAAcTAgCTTATAAG		
<i>Sulfolobus acidocaldarius</i> DSM 639	Saci_0437	<i>thi4</i>	Saci_0854	-37	5.6	CaTATAAgTgAgTTTATAAc		
		<i>thiT</i>	Saci_2110	-41	5.21	aaTATAAcTaAcTTTATAAc		
		<i>thiC</i>	Saci_0137	-31	5.18	tTTATAAcTAgTTTATtga		
<i>Sulfolobus solfataricus</i> P2	SSO0468	<i>thi4</i>	SSO0436	-40	5.66	taTaaAaGTTAATTATAAG		0 bp
		<i>thiT</i>	SSO1593	-41	5.25	aaaATAAgTgAcTTTATAAG		0 bp
<i>Sulfolobus tenA</i>	SSO2089	<i>thi4</i>	SSO2089	-40	5.51	gTTATAAaTAATTATAAT		0 bp
		<i>thiT</i>	ST0361	-32	5.39	taTATtAgcTAgTTTATAAT		
<i>Sulfolobus tokodaii</i> str. 7	ST0179	<i>thi4</i>	ST0361	-32	5.39	taTATtAgcTAgTTTATAAT		
		<i>thiT</i>	ST1132	-41	5.52	CaTATAAgTaggTTTATAAG		
<i>Sulfolobus islandicus</i> Y.N.15.51	YN1551_1174	<i>thi4</i>	YN1551_1140	-40	5.66	taTaaAaGTTAATTATAAG		
		<i>thiT</i>	YN1551_2514	-41	5.25	aaaATAAgTgAcTTTATAAG		
		<i>tenA</i>	YN1551_2613	-40	5.4	CTTATAAaTgATTTATAAT		
<i>Acidianus hospitalis</i> W1	Ahos_0963	<i>thi4</i>	Ahos_0692	-43	5.91	taTATAAcTTAgTTTATAAG		
		<i>thiT</i>	Ahos_0145	-41	5.8	CTTATAAgTaAgTTTATAAT		
		<i>thiC</i>	Ahos_1208	-20	5.14	CTTATtAaGtAcTTTATAa		
<b>Thermoproteales (Motif-II)</b>				<b>Threshold=4.5</b>			<b>AWAAnnnWTWT</b>	
<i>Pyrobaculum aerophilum</i> str. IM2	PAE2210	<i>tenA1-tenA2</i>	PAE0170	-34	5.29	ATAAAATTTTTAT	33 nt	
		<i>thi4</i>	PAE0175	-65	4.79	ATAAtAaTTtT		
		<i>thiT</i>	PAE0176	-32	5.16	ATAATTTTTAT		
		<i>thi(DN)</i>	PAE2535	-37	5.12	ATAAAATTTTTT		
		<i>thiC</i>	PAE0333	-17	<u>4.33</u>	AaAActAaTTTAT		
<i>Pyrobaculum</i> sp. 1860	P186_0445	<i>tenA1-tenA2</i>	P186_0930	-34	4.62	ATAAcgTTTTAT	34 nt	
		<i>thi4</i>	P186_0936	-66	4.79	ATAAtAaTTtT		
		<i>thiT</i>	P186_0937	-62	4.88	ATAAAACaTTAT		
				-28	4.79	AaAAATTTATTAT		
<i>Pyrobaculum oguniense</i> TE7	Pogu_0189	<i>tenA1-tenA2</i>	Pogu_2702	-32	5.01	ATAAAGTTTTAT		
<i>Pyrobaculum arsenaticum</i> DSM 13514	Pars_1934	<i>thiT</i>	Pars_2693	-33	<u>4.16</u>	ATAAATTaTTaAT		
		<i>tenA1-tenA2</i>	Pars_2186	-34	5.01	ATAAAGTTTTAT		
<i>Pyrobaculum calidifontis</i> JCM 11548	Pcal_0941	<i>thiT</i>	Pars_2195	-34	<u>4.36</u>	ATAAAATaTTaAT		
		<i>tenA2</i>	Pcal_0348	-32	4.68	ATAAtgTaTTTAT		
<i>Pyrobaculum islandicum</i> DSM 4184	Pisl_0839	<i>tenA1</i>	Pcal_1316	-34	<u>4.15</u>	tTAAtgTTTTAT		
		<i>thi(DN)</i>	Pcal_1890	-32	<u>4.17</u>	AaAAAcTTTTAT		
		<i>thiC</i>	Pisl_1419	-10	4.73	AaAAATTgTTAT		
<i>Pyrobaculum yellowstonensis</i> WP30	MSUCoA1_00968	<i>thiC</i>	Pisl_1450	-32	4.66	AaAAtcTTTTAT		
		<i>thiT</i>	Pisl_1973	-34	5.12	ATAAAATTTTTT		
		<i>tenA1-tenA2</i>	Pisl_1974	-34	5.12	AaAAATTTTTAT		
		<i>thi4</i>	Pisl_1998	2	<u>4.44</u>	AaAAGaTTTAT		
		<i>thiC</i>	MSUCoA1_00516	-32	5.16	ATAAtTTTTAT		
<i>Thermoproteus neutrophilus</i> V24Sta	Tneu_1384	<i>thi4</i>	MSUCoA1_02190	-32	4.96	ATAAtaTTTAT	16 nt	
		<i>tenA1-tenA2</i>	MSUCoA1_02192	-48	4.79	ATAAAATaTTtT		
		<i>thiT</i>	MSUCoA1_02193	-50	<u>4.3</u>	ATAAAgaTTTtT		
				-34	4.79	AaAAtaTTTAT		
<i>Thermoproteus tenax</i> Kra 1	TTX_0743	<i>thiC</i>	Tneu_0530	5	4.4	AaAaccTTTTAT		
		<i>thi4</i>	Tneu_1052	-32	<u>4.46</u>	AaAAtcTaTTTTAT		
		<i>tenA1-tenA2</i>	Tneu_1074	-32	5.12	AaAAATTTTTAT		
<i>Thermoproteus uzoniensis</i> 768-20	TUZN_0486	<i>thiT</i>	Tneu_1075	-36	5.12	ATAAAATTTTTT		
		<i>thi(DN)</i>	Tneu_1881	-50	<u>4.45</u>	AaAacgTTTTAT		
<i>Vulcanisaeta distributa</i> DSM 14429	Vdis_2360	<i>thiC</i>	TTX_0287	-32	4.89	ATAAATTaTTTTAT		
<i>Vulcanisaeta moutnovskia</i> 768-28	VMUT_0712	<i>thi4</i>	TTX_0666	-15	4.5	ATAAAcTgaTTAT		
		<i>thiT2</i>	TUZN_0079	8	<u>4.23</u>	ATAAAATTaTAT		
		<i>thiT</i>	Vdis_2247	-32	4.99	ATAAAATaTTtT		
		<i>thi4</i>	Vdis_2303	-38	<u>4.08</u>	ATAAAtgTgTTAT		
<i>Caldivirga maquilingensis</i> IC-167	Cmaq_1034	<i>thi4</i>	Vdis_2373	-37	<u>4.37</u>	ATAActTgTTTTAT		
		<i>thi(DN)</i>	Vdis_2402	-39	<u>4.39</u>	ATAAAAgcTTTTAT		
		<i>thiT</i>	VMUT_0613	-7	4.7	ATAAATTTgTTAT		
<i>Caldivirga maquilingensis</i> IC-167	Cmaq_1034	<i>thi4</i>	VMUT_0723	-37	<u>4.37</u>	ATAActTgTTTTAT		
		<i>thi(DN)</i>	VMUT_0747	-38	<u>4.47</u>	ATAAAgaTTTTAT		
<i>Caldivirga maquilingensis</i> IC-167	Cmaq_1034	<i>thiM-thiE</i>	Cmaq_0062	-66	4.76	ATAAttAaTTTTAT		

<i>thi(DN)</i>	Cmaq_0452	-62	<u>4.18</u>	ATAAccATgTTAT
<i>thiT</i>	Cmaq_0759	-25	<u>4.37</u>	ATAAATgTgTTAT
<i>thi4</i>	Cmaq_1420	-64	<u>4.43</u>	ATAAAAATagTAT

**Thermoproteales (Thermofilum) (Motif-III) Threshold=5.0 TATAACnnnGTTATA**

<i>Thermofilum pendens</i> Hrk5	Tpen_0069	<i>ykoEDC</i>	Tpen_0068	-113	5.02	TATAACTTgGTTAaa	31 nt
				-82	5.43	cATAACcGAGTTATA	54 nt
				-28	5.06	TATAACcagGTTtTg	
		<i>thiR</i>	Tpen_0069	-170	5.06	cAaaCctgGTTATA	
				-116	5.43	TATAACTCgGTTATg	
				-85	5.02	TtTAACCaAGTTATA	
<i>Thermofilum</i> sp. 1910b	N186_01270	<i>ykoEDC</i>	N186_01265	-169	5.18	TtTAACCTcGTTATA	47 nt
				-122	5.16	TATAACTCtGTTATg	28 nt
				-94	5.22	cAaaACTCgGTTATA	
		<i>thiR</i>	N186_01270	-90	5.22	TATAACcGAGTTtTg	
				-62	5.16	cATAACCaGAGTTATA	
				-15	5.18	TATAACTGAGTTAaa	
<i>Thermofilum carboxyditrophus</i> 1505	TCARB_1341	<i>ykoEDC</i>	TCARB_1340	-171	5.18	TtTAACCTcGTTATA	47 nt
				-124	5.16	TATAACTCtGTTATg	28 nt
				-96	5.22	cAaaACTCgGTTATA	
		<i>thiR</i>	TCARB_1341	-112	5.22	TATAACcGAGTTtTg	
				-84	5.16	cATAACCaGAGTTATA	
				-37	5.18	TATAACTGAGTTAaa	
<i>Thermofilum</i> sp. 1807-2	TCARB_1341	<i>ykoEDC</i>	MA03_01135*	-161	5.23	TaaAACTagGTTATA	75 nt
				-86	5.31	cATAACTagGTTATA	53 nt
				-33	<u>4.75</u>	TATAACcagaTTATA	
		<i>thiR</i>	MA03_01130	-165	<u>4.75</u>	TATAAtctgGTTATA	
				-112	5.31	TATAACcctAGTTATg	
				-37	5.23	TATAACcctAGTTtTA	

**Desulfurococcales (Motif-IV-a) Threshold=4.5 wwATAACnnnGTTATww**

<i>Desulfurococcus fermentans</i> DSM 16532	Desfe_0411	<i>X-thiC</i>	Desfe_0199	-62	4.99	TAATAACaTtaTTATcA	
		<i>thi(DN)</i>	Desfe_0794	-53	4.7	TtATAAatTtGTTATtG	
		<i>thi4</i>	Desfe_0795	-51	4.7	cAATAACCaAatTTATaA	
		<i>ykoE</i>	Desfe_0090	-58	4.63	TAATAACCaTgATATag	
		<i>ykoDC</i>	Desfe_0187	-35	<u>4</u>	aAgTAgCtAaaTTATTA	
		<i>Desulfurococcus kamchatkensis</i> 1221n	DKAM_0343	<i>X-thiC</i>	DKAM_0123	-92	5.12
<i>thi4</i>	DKAM_0713			-49	4.7	cAATAACCaAatTTATaA	
<i>thi(DN)</i>	DKAM_0714			-40	4.7	TtATAAatTtGTTATtG	
<i>ykoE</i>	DKAM_1465			-46	4.92	TAATAACCaTgATATaA	
<i>ykoDC</i>	DKAM_0109			-21	<u>4</u>	aAgTAgCtAaaTTATTA	
<i>Desulfurococcus mucosus</i> DSM 2162	Desmu_0750			<i>cytX-X-thi(DN)</i>	Desmu_0542	-67	4.82
		<i>ykoE</i>	Desmu_0176	-65	<u>4.34</u>	TgATAACaTGGTgATTt	
		<i>ykoDC</i>	Desmu_0257	-43	<u>4.31</u>	TAATAAaCtGTTATct	
		<i>Staphylothermus marinus</i> F1	Smar_1383	<i>thi(DN)</i>	Smar_0049	-73	<u>4.49</u>
-50	4.77					TgATAACtAaaTTATat	
<i>ykoDC</i>	Smar_0050			-58	4.77	atATAAttTaGTTATcA	23 nt
				-35	<u>4.49</u>	atATAACtTgATATcg	
<i>thi4</i>	Smar_0052			-60	5.19	TtATAACCTGATATTA	23 nt
				-37	<u>4.29</u>	TAATAACcgaGTaATaA	
<i>ykoE-tenA</i>	Smar_0053	-64	<u>4.29</u>	TtAtAcTcGTTATTA	23 nt		
		-41	5.19	TAATAAtCAGGTTATaA			
<i>Staphylothermus hellenicus</i> DSM 12710	Shell_1065	<i>ykoE-tenA</i>	Shell_0763	-63	<u>4.28</u>	TtAtAcTcagTTATTA	23 nt
				-40	5.18	TAATAAtAGGTTATaA	
		<i>thi4</i>	Shell_0764	-63	5.18	TtATAACCTaaTTATTA	23 nt
				-40	<u>4.28</u>	TAATAACtgaGTaATaA	
		<i>ykoDC</i>	Shell_0766	-56	4.77	atATAAttTaGTTATcA	23 nt
				-33	<u>4.49</u>	atATAACtTgATATcg	
<i>thi(DN)</i>	Shell_0767	-79	<u>4.49</u>	cgATAACCaGTTATat	23 nt		
<i>cytX</i>	Shell_0907	-56	4.77	TgATAACtAaaTTATat			
<i>Thermosphaera aggregans</i> DSM 11486	Tagg_0649	<i>ykoE</i>	Tagg_0461	-56	4.97	TtATAACtTaGaTATTA	
		<i>ykoDC</i>	Tagg_0459	-63	<u>4.3</u>	TgATAAttcGtTTATcA	
		<i>thiD</i>	Tagg_0777	-48	<u>4.13</u>	TAATcACCTaGaTATcA	
<i>Thermogladius cellulolyticus</i> 1633	TCELL_1104	<i>thiC</i>	TCELL_1013	-62	4.57	TtATAACcCGTTATtG	23 nt
				-39	4.81	TtATAtCCIGaTTATaA	
				-64	<u>4.14</u>	TgATAAaCgtGTTATag	23 nt
		<i>thi4-thi(DN)</i>	TCELL_1023	-41	4.63	TAtTAACCTGaTTATaA	
				-41	4.85	TAATAACcAGGTTtTaA	
		<i>ykoE</i>	TCELL_0952	-41	4.85	TAATAACcAGGTTtTaA	
<i>ykoDC</i>	TCELL_0944	-26	<u>4.06</u>	cgATAACtTGGTtCtag			

**Desulfurococcales (Motif-IV-b) Threshold=5.0 wwAwACnnnGTWATww**

<i>Hyperthermus butylicus</i> DSM 5456	Hbut_0125	<i>thi4</i>	Hbut_0253	-53	5.01	TTAgacACACTGTAATcT
--	-----------	-------------	-----------	-----	------	--------------------

		<i>thiC</i>	Hbut_0289	-70	5.19	TTATTACAGTGTtAtgT		
<i>Ignisphaera aggregans</i> DSM 17230	lgag_0434	<i>thi(DN)</i>	lgag_1669	-30	5.05	TTAaTACTtTGTtATAT		
		<i>thiC</i>	lgag_1298	-64	4.6	ATAgTACTActTAATAA		
<b>Fervidicoccales (Motif-V)</b>							<b>Threshold=4.8</b>	<b>TnATAAnnnnnTTATnA</b>
<i>Fervidococcus fontis</i> Kam940	FFONT_0578	<i>thi4</i>	FFONT_0392	-51	4.9	TTATAACAaAaATTtTtA	23 nt	
				-28	4.41	TgATtATATaGTTATAt		
		<i>thiR</i>	FFONT_0578	-120	4.91	TaATAATAAgtTtTtTAA	87 nt	
				-33	4.22	TaATAACcTgGTaATgt		
		<i>ykoE</i>	FFONT_0579	-74	4.89	TTgTAAaATtTtTTATAA	34 nt	
				-40	4.84	TTATAtCcTTATTATtA		
		<i>ykoDC</i>	FFONT_0581	-48	5	TaATAATATatTTATcA		
<b>Euryarchaeota</b>							<b>Threshold=4.4</b>	<b>AWAACnnnnGTTWT</b>
<b>Thermococci (Thermococcales)</b>								
<i>Pyrococcus abyssi</i> GE5	PAB0877	<i>thi4</i>	PAB0536	-48	5.1	AtAACaAAGTTTT	21 nt	
				-27	5.4	AAAACtAAGTTTT		
		<i>thiC</i>	PAB1930	-37	4.31	AttACTTgGTTTT	21 nt	
				-16	3.84	AtAACaAgGTgaT		
		<i>thiB</i>	PAB1835	-36	4.07	AAAAtcTtaTTTT		
		<i>tenAA-cytX-thiME(DN)</i>	PAB1641	-69	4.89	AtAACcAAGTTTT		
<i>Pyrococcus furiosus</i> DSM 3638	PF0601	<i>thi4</i>	PF1530	-113	4.29	AtAAtaggGTTTT	21 nt -6 bp	
				-92	5.18	AAAACtTgGTTTT		
				-47	5.29	AAAACaAAGTTTT	21 nt	
				-26	5.3	AAAACtCAGTTTT		
		<i>thiC</i>	PF1531	-115	5.3	AAAACtAgGTTTT	21 nt -33 bp	
				-94	5.29	AAAACtTtGTTTT		
				-49	5.18	AAAACcAAGTTTT	21 nt	
				-28	4.29	AAAACcctaTTaT		
		<i>tenAA-cytX-thiME(DN)</i>	PF1338	-67	5.02	AtAACTTgGTTTT	21 nt -21 bp	
				-46	4.05	AcAAaaAAGTTaT		
		<i>thiB</i>	PF1751	-44	4.65	AgAAcAcAGTTTT	21 nt -20 bp	
				-23	3.94	AgAAaTggGTTTT		
<i>Pyrococcus horikoshii</i> OT3	PH0716	<i>thi4</i>	PH1357	-38	5.13	AtAACaAAGTTTT	21 nt	
				-17	4.51	AAAAtgAAGTTTT		
		<i>thiB</i>	PH1349	-64	4.85	AAAtACTTAGTTTT	21 nt	
				-43	4.23	AtAgCTTgGTTTT		
		<i>tenAA-cytX-thiME(DN)</i>	PH1161	-60	4.69	AtAACaAgTTaT	21 nt	
				-39	4.05	AcAAaaAAGTTaT		
<i>Pyrococcus</i> sp. NA2	PNA2_1361	<i>thiC</i>	PNA2_0094	-39	4.41	AttACTTgGTTTT		
		<i>thi4-2</i>	PNA2_1583	-38	4.62	AAAAtcAAGTTaT	21 nt	
				-17	4.08	AtAACTTgGTgaT		
		<i>tenAA-cytX-thiME(DN)</i>	PNA2_1769	-67	4.98	AtAACTcAGTTaT	21 nt	
				-46	3.77	AcAAaaAgGTTaT		
		<i>thi4-1</i>	PNA2_1941	-47	5.13	AtAACaAAGTTTT	21 nt	
				-26	5.3	AAAACtAgGTTTT		
<i>Pyrococcus</i> sp. ST04	Py04_0745	<i>thi4</i>	Py04_1231	-39	5.13	AtAACaAAGTTTT	21 nt	
				-18	5.46	AAAACtTAGTTTT		
		<i>tenAA-cytX-thiME(DN)</i>	Py04_0940	-68	4.69	AtAACcAtGTTaT	21 nt	
				-47	3.89	AcAAaagGTTaT		
<i>Pyrococcus yayanosii</i> CH1	PYCH_07480	<i>thi4-1</i>	PYCH_06240	-65	5.13	AtAACaAAGTTTT	21 nt	
				-44	5.46	AAAACtTAGTTTT		
		<i>thiC</i>	PYCH_06250	-49	5.46	AAAACtAAGTTTT	21 nt	
				-28	5.13	AAAACtTtGTTaT		
		<i>thi4-2</i>	PYCH_08650	-36	4.62	AAAAtcAAGTTaT	21 nt	
				-15	4.08	AtAACTTgGTgaT		
<i>Thermococcus litoralis</i> DSM 5473	OCC_09391	<i>thiC</i>	OCC_06561	-56	5.3	AAAACtAgGTTTT	21 nt	
				-35	4.73	AAAACtTtaTTaT		
				-24	4.86	AtAACcAAGTTaT		
		<i>thi4</i>	OCC_06556	-80	4.86	AtAACTTgGTTaT	21 nt	
				-69	4.73	AtAAtaAAGTTTT		
				-48	5.3	AAAACtCAGTTTT		
		<i>cytX-thiMEW(DN)</i>	OCC_04984	-68	4.3	AcAACTTgaTTTT	21 nt	
				-47	3.77	AcAAaaAgGTTaT		
		<i>thiB</i>	OCC_11612	-46	4.65	AgAAcAcAGTTTT	21 nt	
				-25	3.94	AgAAaTggGTTTT		
		<i>thiR-X-yjbQ</i>	OCC_09391	-38	4.74	AtAACcTgGTTTT		
<i>Thermococcus barophilus</i> MP	TERMP_01433	<i>cytX-thiM-thiE-thi(DN)</i>	TERMP_00505	-70	5.14	AtAACTTAGTTaT	21 nt	
				-49	3.89	AcAAaagGTTaT		
		<i>thiPQ</i>	TERMP_01606	-18	5.46	AAAACtAAGTTTT		
		<i>thiB</i>	TERMP_01742	-70	4.41	AcAAtaAAGTTTT		
		<i>thiR-X-yjbQ</i>	TERMP_01433	-23	4.18	AtAAcccGTTTT		
<i>Thermococcus sibiricus</i> MM 739	TSIB_0198	<i>cytX-tenA-thiMEW(DN)</i>	TSIB_0467	-69	5.02	AtAACTTgGTTTT	21 nt	



					-48	4.05	AcAaaaAGTTaT	
		<i>thiB</i>	TSIB_0369		-35	4.49	AgAACagAGTTaT	21 nt
					-14	4.86	AAAAaTTAGTTTT	
		<i>thiR-X-yjbQ</i>	TSIB_0198		-38	5.14	AtAACTgAGTTTT	
<i>Thermococcus onnurineus</i> NA1	TON_1436	<i>thiPQ</i>	TON_1959		-16	4.54	AAAACcgAGTTgT	
		<i>thiB</i>	TON_0963		-25	3.93	AcAACcgAGTaT	
		<i>thiR-X-yjbQ</i>	TON_1436		-23	4.58	AtAAACcgGTTTT	
<i>Thermococcus kodakaraensis</i> DSM 3638	TK2197	<i>thiC</i>	TK0433		-62	5.46	AAAACCTAAGTTTT	21 nt
					-41	5.13	AAAACCTtGTTaT	
		<i>thi4-thi(DN)</i>	TK0434		-67	5.13	AtAACaAAGTTTT	
					-46	5.46	AAAACCTAGTTTT	
		<i>thiPQ</i>	TK0159		56	4.38	AgAACcgAGTTaT	
		<i>thiR-yjbQ</i>	TK2197		-26	4.26	AtAAACcgTgTgT	
<i>Thermococcus gammatolerans</i> EJ3	TGAM_2087	<i>thiC</i>	TGAM_1269		-61	5.55	AAAACCTAAGTTTT	21 nt
					-40	5.17	AAAACCTTGTAT	
		<i>thi4-thi(DN)</i>	TGAM_1270		-45	5.55	AAAACCTAGTTTT	
					-66	5.17	ATAACAAAGTTTT	
		<i>thiB</i>	TGAM_1706		-16	4.47	AAAACCGGTTTT	
<i>Thermococcus</i> sp. 4557	GQS_08390	<i>thiC</i>	GQS_02010		-61	5.46	AAAACCTAAGTTTT	21 nt
					-40	5.13	AAAACCTtGTTaT	
		<i>thi4-thi(DN)</i>	GQS_02015		-67	5.13	AtAACaAAGTTTT	
					-46	5.46	AAAACCTAGTTTT	
		<i>thiPQ</i>	GQS_05315		-18	4.54	AAAACcgAGTTgT	
		<i>thiB</i>	GQS_10685		-26	4.13	AAAACcgGTaT	
		<i>thiR-X-yjbQ</i>	GQS_08390		-25	4.74	AtAACcTgGTTTT	
<i>Thermococcus</i> sp. AM4	TAM4_328	<i>thiB</i>	TAM4_1850		-24	5.01	AAAACcAtGTTTT	
<i>Thermococcus</i> sp. CL1	CL1_1295	<i>thi4-thi(DN)</i>	CL1_0036		-65	5.13	AtAACaAAGTTTT	21 nt
					-44	5.3	AAAACtAGTTTT	
		<i>thiC</i>	CL1_0038		-63	5.46	AAAACCTAAGTTTT	21 nt
					-42	5.13	AAAACCTtGTTaT	
		<i>thiPQ</i>	CL1_0708		-18	5.02	AAAACcgAGTTTT	
		<i>thiR-yjbQ</i>	CL1_1295		-25	4.58	AtAAACcgGTTTT	
<b>Halobacteria (Halobacteriales / Haloferacales / Natrialbales)</b>					<b>Threshold=3.8</b>		<b>TAYnACnnnGtnRTA</b>	
<i>Natronomonas pharaonis</i> DSM 2160	NP5168A	<i>thi4</i>	NP5174A		-13	4.53	TAtaACtAGtGTA	
		<i>thiC</i>	NP2210A		-28	3.97	TAttAgctgGtGTA	12 nt
					-15	3.74	TAttACTagaTtGtT	
		<i>thiE-thiM</i>	NP4054A		-13	4.17	TACCAGcGgGtGTA	
<i>Halorhabdus utahensis</i> DSM 12940	Huta_2411; Huta_2715	<i>thi4</i>	Huta_2718		-13	4.83	TACCACtAGTGGTA	
		<i>thiC</i>	Huta_2472		-41	4.56	TtCACtAgGtTaTA	12 nt
					-28	3.93	TAtaAtctAGTaGTA	
		<i>tenA1-tenA2-thiV</i>	Huta_0684		-55	3.93	TAtaAtTtgTaGTA	12 nt
					-42	3.97	TAttAgTGAGtTaTA	
		<i>thiBPQ</i>	Huta_0388		-24	3.84	TACCACtAGTcaaA	12 nt
					-11	4.44	aACCACcaAGTGGTg	
<i>Haloarcula marismortui</i> ATCC 43049	rrnAC1777; rrnAC1764	<i>thi4</i>	rrnAC1782		(-26)	3.9	TACCACcaAGaGaTA	12 nt
					(-13)	4.63	TAttACTagTGGTA	
		<i>thiC</i>	rrnAC1182		-27	4.5	TAtaACTtgTaGTA	12 nt
					-14	4.37	TAtaACctAGTgTt	
		<i>tenA1-tenA2-thiV</i>	rrnAC1765		-59	4.55	TACaACcaAGTgGTA	
		<i>thiBPQ</i>	rrnAC0147		-24	3.86	TACCACtGgGTgA	12 nt
					-11	3.49	gACCACgGtGtaaTg	
<i>Haloarcula hispanica</i> ATCC 33960	HAH_2305; HAH_2294	<i>thi4</i>	HAH_2313		(-13)	4.75	TAtCACtAgTGGTA	
		<i>thiC</i>	HAH_1784		-27	4.5	TAtaACcaAGTaGTA	12 nt
					-14	4.37	TAtaACctAGTgTt	
		<i>tenA1-tenA2-thiV</i>	HAH_2295		-59	4.55	TACaACTtgGtGTA	
		<i>thiBPQ</i>	HAH_0897		-24	3.86	TACCACtGgGTgA	12 nt
					-11	3.9	gACCACgGAGTaaTg	
<i>Halomicrobium mukohataei</i> DSM 12286	Hmuk_1352; Hmuk_0022	<i>thi4</i>	Hmuk_0002		-26	3.48	TgCaACgtgGTGGTA	12 nt
					-13	4.68	TACCACtAgGtGTA	
		<i>thiC</i>	Hmuk_2297		-28	3.99	TAtaAtctAGTgTaTA	12 nt
					-15	3.7	TAtaACTagcTtaTc	
<i>Haloterrigena turkmenica</i> DSM 5511	Htur_2542; Htur_1469	<i>thi4</i>	Htur_1462		-26	3.53	cACCACgCgGTGcTA	12 nt
					-13	4.68	TACCACcaAGTgGTA	
		<i>thiC</i>	Htur_1332		-28	4.01	TAtaACTCgaTgTaTA	12 nt
					-15	4.27	TACaACctAGTtaTc	
		<i>thiBPQ</i>	Htur_2516		-24	4.21	gACCACcCAGTGGaA	12 nt
					-11	3.87	aACCACcCgTaaTg	
<i>Natrialba magadii</i> ATCC 43099	Nmag_0434; Nmag_2413	<i>thi4</i>	Nmag_2419		-26	3.9	cActACTGgaTgTaTA	12 nt
					-13	4.68	TACCACctAGTgGTA	
		<i>thiC</i>	Nmag_2593		-28	3.87	TAtaAtctAGTaaTA	12 nt
					-15	4.27	TACaACctAGTtaTc	

		<i>thiBPQ</i>	Nmag_0460	-24	<u>3.53</u>	aACCACaGAGTgtaA	12 nt	
				-11	4.4	aACCACCTgGTGaTg		
<i>Halorubrum lacusprofundi</i> ATCC 49239	Hlac_0446	<i>thi4</i>	Hlac_2980	-26	3.9	TACCACcGgGaGaTA	12 nt	
				-13	4.77	TAtCACTaAGTGGTA		
		<i>thiBPQ</i>	Hlac_2697	-24	3.9	aACCACcGgGTgtaA	12 nt	
				-11	4.13	aACaACcCgGTaaTg		
<i>Haloquadratum walsbyi</i> DSM 16790	HQ1277A	<i>thi4</i>	HQ1276A	-26	<u>3.25</u>	TAgtAgTtgGTaaTA	12 nt	
				-13	4.31	TaaCACCaAGTaaTA		
		<i>tenA-thiV</i>	HQ2369A	-61	4.01	TAttAgTtAGTtGTA		
<i>Haloferax volcanii</i> DS2	HVO_0662	<i>thi4</i>	HVO_0665	-26	4.57	TACTACTagGTaGTA	12 nt	0 bp
				-13	4.65	TAtCACTaAGTaGTA		
		<i>thiC</i>	HVO_2154	-28	<u>3.66</u>	TAttAactAGTtGTA	12 nt	0 bp
				-15	3.84	TAtCACTagcTaaTc		
		<i>thiBPQ</i>	HVO_0022	-24	3.23	TACCACaGAGTacaA	12 nt	
				-11	4.15	aACaACcGAGTaaTg		
<i>Halobacterium</i> sp. NRC-1	VNG0247C	<i>thi4-thi(DN)</i>	VNG2604G	-26	<u>3.74</u>	aAttAtctAGTtGTA	12 nt	
				-13	3.89	TACCAtgtAGTGGTA		
		<i>thiC</i>	VNG0715G	-30	4.06	TAttAtctAGTGGTA	12 nt	
				-17	4.4	TAtCACTaAGTaaTt		
		<i>thiBPQ</i>	VNG2395C	-11	4.13	aACaACcCgGTaaTg		
<i>Halogeometricum borinquense</i> DSM 11551	Hbor_25480	<i>thi4</i>	Hbor_25460	-26	<u>3.39</u>	TAgtAgcCAGTgaTA	12 nt	
				-13	4.49	TaaCACTagGTGGTA		
		<i>thiC</i>	Hbor_14920	-28	3.96	TAttAtTCgGTaGTA	12 nt	
				-15	4.41	TAtCACTagGTaGTc		
		<i>thiBP</i>	Hbor_02000	-24	<u>3.55</u>	TAtgACcaAGTGGgA	12 nt	
				-11	4.15	gACTACTCAGTaaTg		
		<i>tenA1-tenA2-thiV</i>	Hbor_30560	-55	<u>3.73</u>	aACaACcCgaTaaTA		
<i>Natrinema</i> sp. J7-2	NJ7G_1102; NJ7G_3692	<i>thi4</i>	NJ7G_3680	-26	3.9	cACTACcaAGTcaTA	12 nt	
				-13	4.68	TACCACctAGTtGTA		
		<i>thiC</i>	NJ7G_3772	-28	3.91	TAtaAtTCAGTaaTA	12 nt	
				-15	4.27	TACaACctAGTtaTc		
		<i>tenA</i>	NJ7G_3269	-44	3.94	TAttAgctAGTaaTA		
		<i>thiE-thiM</i>	NJ7G_0299	-20	4.2	TACTACcGAcTGGTA		
		<i>thiBPQ</i>	NJ7G_0450	-24	<u>3.32</u>	gACTACctAGTgcaA	12 nt	
				-11	4.26	aACCACcCgGTaaTg		
<i>Natrinema pellirubrum</i> DSM 15624	Natpe_1684; Natpe_2824	<i>thi4</i>	Natpe_2830	-26	4.08	TACTACTagGTcaTA	12 nt	
				-13	4.68	TACCACctAGTtGTA		
		<i>thiC</i>	Natpe_3081	-28	4.5	TAtaACctAGTaGTA	12 nt	
				-15	4.27	TACaACctAGTtaTc		
		<i>thiBPQ</i>	Natpe_3863	-24	3.91	gACCACtGAGTgtaA	12 nt	
				-11	4.28	aACCACcGAGTaaTg		
		<i>thiE-thiM</i>	Natpe_3729	-20	4.02	cACTACcGAcTGGTA		
<i>Halopiger xanaduensis</i> SH-6	Halxa_0732; Halxa_2856	<i>thi4</i>	Halxa_2863	-26	<u>3.37</u>	cACCACtGAGTccTA	12 nt	
				-13	4.68	TACCACctAGTtGTA		
		<i>thiC</i>	Halxa_3896	-28	<u>3.44</u>	TAtaACTagaTcaTA	12 nt	
				-15	4.27	TACaACctAGTtaTc		
		<i>tenA</i>	Halxa_2795	-37	<u>3.39</u>	gAtaACTaAtTtaTA	12 nt	
				-24	4.53	TAttACTaAGTaGTA		
		<i>thiBPQ</i>	Halxa_0707	-24	4.11	gACCACtCAGTaGaA	12 nt	
				-11	<u>3.78</u>	aACaACcCgGTaaTg		
<i>Haloferax mediterranei</i> ATCC 33500	HFX_0632	<i>thi4</i>	HFX_0634	-26	4.69	TACTACTtgGTGGTA	12 nt	
				-13	4.65	TAtCACTaAGTaGTA		
		<i>thiC</i>	HFX_2195	(-28)	4.24	TAtCagctAGTGGTA	12 nt	
				(-15)	4.43	TAtaACTagGTGGTt		
		<i>thiBPQ</i>	HFX_0021	(-24)	<u>3.23</u>	TACCACaCAGTacaA		
				(-11)	4.13	aACaACcCgGTaaTg		
		<i>tenA1-tenA2-thiV</i>	HFX_5296	-59	4.25	aAtaACcaAGTaaTA		
<i>Halovivax ruber</i> XH-70	Halru_1158	<i>thi4</i>	Halru_1150	-13	4.38	TaaCACcGAGTtGTA		
		<i>thiC</i>	Halru_0933	-27	4.1	TAtgACTtAGTaGTA	12 nt	
				-14	4.41	TACTACctAGTgaTc		
		<i>thiBPQ</i>	Halru_0146	-24	3.96	gACCACcCAGTGGgA	12 nt	
				-11	<u>3.68</u>	gACCAtcCAGTaaTg		

<sup>1</sup>Divergently located adjacent operons that share ThiR-site(s) in the same promoter region are boxed.

<sup>2</sup>Weaker sites with scores that are 15% less than the threshold are underlined. Tandem sites are boxed. Score thresholds are given for each motif.

<sup>3</sup>Position of the 5' end of a ThiR-site is given relative to the translation start site. Positions relative to the corrected translation start sites are in parenthesis.

<sup>4</sup>Positions of transcription start sites (TSSs) mapped to the analyzed genomes according to the following references:

*Sulfolobus solfataricus* (Wurtzel et al., 2010); *Pyrococcus furiosus* (Yoon et al., 2011);

*Haloferax volcanii* (Brenneis et al., 2007; J. Maupin-Furlow, personal communication)

**Table S3. ThiR-binding sites and their experimental validation in *Metallosphaera yellowstonensis* MK1.**

First gene locus tag	ThiR regulated operon	Predicted ThiR-binding sites and synthetic DNA fragments used in fluorescence polarization DNA-binding assays <sup>1</sup>
MetMK1DRAFT_00023630	<i>thi4</i>	<b>a</b> cccccggttggat <b>TATATAATCTAGTTTATAAG</b> <u>Gccacta</u> accccc
MetMK1DRAFT_00024350	<i>thiT</i>	<b>a</b> ccccggttatg <b>tcTAAATAAGTAAGTTTATAAC</b> <u>Ccgactt</u> accccc
MetMK1DRAFT_00019590	<i>tenA</i>	<b>a</b> cccc <b>ataactagATTATAAAATGATTTATAT</b> <u>Cccgcc</u> accccc
MetMK1DRAFT_00012900	<i>trxA(N.C.)</i> <sup>2</sup>	<b>a</b> ggggaatataaaactttttaaacggggg

<sup>1</sup> Predicted ThiR-binding sites are capitalized. DNA fragments from target gene upstream region are underlined.

Additional poly-cytosine oligonucleotides have been added on each side of DNA fragments to improve annealing of oligos.

For each DNA fragment, two complimentary single-stranded oligonucleotides were synthesized by IDT

at that one of these fragments (5'-3' shown here) was labeled by 6-carboxyfluorescein at 5' end (marked by red).

<sup>2</sup>NC, a negative control DNA fragments from promoter region of a gene that lack a ThiR-binding site.