H_2-dependent formate production by hyperthermophilic Thermococcales: An alternative to sulphur reduction for reducing-equivalents disposal

Sébastien Le Guellec[1], Elodie Leroy[1], Damien Courtine[1], Anne Godfroy[1], Erwan G. Roussel[1]*

[1] Ifremer, Univ Brest, CNRS, Laboratoire de Microbiologie des Environnements Extrêmes, F-29280 Plouzané, France.

*Corresponding author. Address: Erwan G. Roussel, Erwan.Roussel@ifremer.fr

Supplementary Methods . Thermodynamic calculations: reaction quotient calculation

Table S1. List of *Thermococcales* used in this study.

Table S2. List of 59 conserved single copy genes used to calculate the *Thermococcales* phylogenomic tree.

Table S3. Reversibility of H_2-dependent formate production by (hyper)thermophilic *Thermococcales*.

Table S4. H_2 and formate concentrations used for representation of environmental concentrations measured for seawater and end-member fluids from six deep-sea hydrothermal vent fields in figure 6.

Figure S1. Comparison of the gene organisations of the *fdh-nor*, *fhl1*, *fhl2* and *fhl3* clusters from 35 genomes of *Thermococcales*.

Figure S2. Phylogenetic trees based on translated partial amino acid sequences of functional genes: (a) formate dehydrogenase catalytic subunit (68 sequences) and (b) formate transporter (17 sequences).

Figure S3. Comparison of the gene organisations of the *codh-nor* and *codh* clusters from 35 genomes of *Thermococcales*.

Figure S4. Cross effects of H_2 and formate on CO_2 reduction rate in *T. onnurineus* NA1 after 96 hours.

Figure S5. The direction of the reaction of CO_2 reduction depends primarily on substrates concentration.

Figure S6. 2D-contour plot representing the thermodynamic modelling of the CO_2 reduction to formate for a range of concentration of formate and H_2 (assumed as dissolved).

Figure S7. Effect of temperature on formate production rates in *T. onnurineus* NA1.

Supplementary Methods

Thermodynamic calculations: reaction quotient calculation

The reaction quotient Qr was calculated according to :

$$Qr = \prod a_i^{V_{i,r}}$$

where a_i stands for the activity of the i^{th} species raised to its stoichiometric reaction coefficient $V_{i,r}$ in the r reaction. In the case of gases, activity is replaced by the fugacity f_i of the species.

The activity (a) of any species i can be determined according to:

$$a = \frac{Ci}{C_i^O} \times \gamma i$$

where Ci stands for the concentration, C_i^O represents the standard state concentration (usually 1 *m*), and γi represents the corresponding activity coefficient (which is unitless).

The fugacity (f) of any i^{th} gas can be determined with the relation:

$$f_i = P_i \times \gamma i$$

where Pi stands for the partial pressure (bar) and γi represents the corresponding fugacity coefficient (unitless).

Activity of H_2O (l) is taken to be unity ($a_{H2O(l)}$ = 1). Activity coefficients of bicarbonate ($\gamma_{HCO_3^-}$) and formate (γ_{HCOO^-}) are almost identical as they have the same species charge of -1 whatever the ionic strength and temperature as described by Amend et LaRowe (43). In our low-pressure condition (2 bar), regardless of the temperature or ionic strength, values of γi are 0.99–1.00. We therefore assumed that $fi \approx Pi$.

Hence, calculation of reaction quotient (Qr) for the reaction of CO_2 reduction to formate:

$$H_2\,(g) + HCO_3^-\,(aq) \leftrightarrow HCOO^-\,(aq) + H_2O\,(l)$$

was simplified using concentrations considering activities of reactants and products almost equal to concentrations:

$$Qr = \frac{a_{HCOO^-(aq)} \times a_{H_2O(l)}}{f_{H_2(g)} \times a_{HCO_3^-(aq)}} = \frac{([HCOO^-] \times \gamma_{HCOO^-})}{(P_{H_2} \times \gamma_{H_2}) \times ([HCO_3^-] \times \gamma_{HCO_3^-})} \approx \frac{[HCOO^-]}{P_{H_2} \times [HCO_3^-]}$$

where [i] represents the concentration of i in the aqueous solution (M) and Pi the partial pressure (bar).

Note that (Qr) could also be calculated considering H_2 as dissolved. In that case, for uncharged species H_2, $\gamma_{H_2(aq)} \approx 1$, regardless of temperature or ionic strength and thus, very little error is also introduced by equating $a_{H_2(aq)} \approx [H_2]_{(aq)}$.
Dissolved H_2 concentrations are estimated form the partial pressure of H_2 in the headspeace (P_{H_2}) as described in Materials and Methods section.

43. Amend et LaRowe Amend JP, LaRowe DE. Minireview: demystifying microbial reaction energetics. Environ Microbiol. 2019;21:3539-47.

Table S1. List of *Thermococcales* used in this study. Strains performing H_2-dependent formate production are indicated in red.

Genus	Specie	Strain	Culture collection number	Genome accession number (NCBI)
Thermococcus	*acidaminovorans*	AEDII10^T	DSM-11906	Not available
Thermococcus	*aegaeus*	P5^T	DSM-12767	Not available
Thermococcus	*aggregans*	TY^T	DSM-12819	Not available
Thermococcus	*alcaliphilus*	AEDII12^T	DSM-10322	Not available
Thermococcus	*atlanticus*	MA898^T	IFREMER-898	Not available
Thermococcus	*barophilus*	MP^T	DSM-11836	CP002372
Thermococcus	*barophilus*	Ch5	UBOCC-M-3206	CP013050
Thermococcus	*barossii*	SHCK-94^T	UBOCC-M-3299	CP015101
Thermococcus	*celer*	Vu13^T	DSM-2476	CP014854
Thermococcus	*celericrescens*	TS2^T	DSM-17994	LLYW00000000
Thermococcus	*chitonophagus*	GC 74^T	DSM-10152	CP015193
Thermococcus	*cleftensis*	CL1^T	DSM-27260	CP003651
Thermococcus	*eurythermalis*	A501^T	JCM-30233	CP008887
Thermococcus	*fumicolans*	ST557^T	IFREMER-557	Not available
Thermococcus	*gammatolerans*	EJ3^T	DSM-15229	CP001398
Thermococcus	*gorgonarius*	W-12^T	DSM-10395	CP014855
Thermococcus	*guaymasensis*	TYS^T	DSM-11113	CP007140
Thermococcus	*hydrothermalis*	AL662^T	IFREMER-662	Not available
Thermococcus	*kodakarensis*	KOD1^T	JCM-12380	AP006878
Thermococcus	*litoralis*	NS-C^T	DSM-5473	CP006670
'*Thermococcus*	*marinus*'	EJ1	UBOCC-M-3266	Not available
'*Thermococcus*	*mexicalis*'	GY869	IFREMER-869	Not available
Thermococcus	*nautili*	30-1^T	/	CP007264
'*Thermococcus*	*onnurineus*'	NA1	UBOCC-M-3201	CP000855
Thermococcus	*pacificus*	P4^T	DSM-10394	CP015102
Thermococcus	*paralvinellae*	ES1^T	DSM-27261	CP006965
Thermococcus	*peptonophilus*	OG-1^T	DSM-10343	CP014750
Thermococcus	*piezophilus*	CDGS^T	UBOCC-M-3296	CP015520
Thermococcus	*prieurii*	Bio-PL-0405IT2^T	UBOCC-M-3069	Not available
Thermococcus	*profundus*	DT5432^T	UBOCC-M-3193	CP014862
'*Thermococcus*	*radiotolerans*'	EJ2	DSM-15228	CP015106
Thermococcus	*sibiricus*	MM 739^T	DSM-12597	CP001463
Thermococcus	*siculi*	RG-20^T	DSM-12349	CP015103
Thermococcus	*stetteri*	K-3^T	DSM-5262	Not available
Thermococcus	*thioreducens*	OGL-20P^T	DSM-14981	CP015105
Thermococcus	sp.	AM4	UBOCC-M-3218	CP002952
Thermococcus	sp.	MF15	/	unpublished
Thermococcus	sp.	P6	Not tested	CP015104
Thermococcus	sp.	4557	Not tested	CP002920
'*Pyrococcus*	*abyssi*'	GE5	IFREMER-855	AL096836
Pyrococcus	*furiosus*	Vc 1^T	DSM-3638	AE009950
Pyrococcus	*glycovorans*	AL585^T	IFREMER-585	Not available
Pyrococcus	*horikoshii*	OT-3^T	DSM-12428	BA000001
Pyrococcus	*kukulkanii*	NCB100^T	DSM-101590	CP010835
Pyrococcus	*yayanosii*	Ch1^T	UBOCC-M-3110	CP002779
Pyrococcus	*woesei*	DSM-3773^T	DSM-3773	Not available
Pyrococcus	sp.	ES4	UBOCC-M-3230	Not available
Palaeococcus	*ferrophilus*	DMJ^T	DSM-13482	LANF00000000
Palaeococcus	*pacificus*	DY20341^T	DSM-24777	CP006019

Table S2. List of 59 conserved single copy genes used to calculate the *Thermococcales* phylogenomic tree.

PFAM ID	HMM name	Description of archaeal genes
PF00687	Ribosomal_L1	Ribosomal protein L1p/L10e family
PF00466	Ribosomal_L10	Ribosomal protein L10
PF00298	Ribosomal_L11	Ribosomal protein L11, RNA binding domain
PF03946	Ribosomal_L11_N	Ribosomal protein L11, N-terminal domain
PF00572	Ribosomal_L13	Ribosomal protein L13
PF00238	Ribosomal_L14	Ribosomal protein L14p/L23e
PF00827	Ribosomal_L15e	Ribosomal L15
PF00252	Ribosomal_L16	Ribosomal protein L16p/L10e
PF01280	Ribosomal_L19e	Ribosomal protein L19e
Pf00181	Ribosomal_L2	Ribosomal Proteins L2, RNA binding domain
PF01157	Ribosomal_L21e	Ribosomal protein L21e
PF00237	Ribosomal_L22	Ribosomal protein L22p/L17e
PF00276	Ribosomal_L23	Ribosomal protein L23
PF00831	Ribosomal_L29	Ribosomal L29 protein
PF03947	Ribosomal_L2_C	Ribosomal Proteins L2, C-terminal domain
FF00297	Ribosomal_L3	Ribosomal protein L3
PF00327	Ribosomal_L30	Ribosomal protein L30p/L7e
PF01198	Ribosomal_L31e	Ribosomal protein L31e
PF01655	Ribosomal_L32e	Ribosomal protein L32
PF01780	Ribosomal_L37ae	Ribosomal L37ae protein family
PF00832	Ribosomal_L39	Ribosomal L39 protein
PF00573	Ribosomal_L4	Ribosomal protein L4/L1 family
PF00935	Ribosomal_L44	Ribosomal protein L44
PF00281	Ribosomal_L5	Ribosomal protein L5
PF00673	Ribosomal_L5_C	ribosomal L5P family C-terminus
PF00347	Ribosomal_L6	Ribosomal protein L6
PF00411	Ribosomal_S11	Ribosomal protein S11
PF00164	Ribosomal_S12	Ribosomal protein S12
PF00416	Ribosomal_S13	Ribosomal protein S13/S18
PF08069	Ribosomal_S13_N	Ribosomal S13/S15 N-terminal domain
PF00312	Ribosomal_S15	Ribosomal protein S15
PF00366	Ribosomal_S17	Ribosomal protein S17
PF00833	Ribosomal_S17e	Ribosomal S17
PF00203	Ribosomal_S19	Ribosomal protein S19
PF01090	Ribosomal_S19e	Ribosomal protein S19e
PF00318	Ribosomal_S2	Ribosomal protein S2
PF01282	Ribosomal_S24e	Ribosomal protein S24e
PF01599	Ribosomal_S27	Ribosomal protein S27a
PF01200	Ribosomal_S28e	Ribosomal protein S28e
PF01015	Ribosomal_S3Ae	Ribosomal S3Ae family
PF00189	Ribosomal_S3_C	Ribosomal protein S3, C-terminal domain
PF00900	Ribosomal_S4e	Ribosomal family S4e
PF00333	Ribosomal_S5	Ribosomal protein S5, N-terminal domain
PF03719	Ribosomal_S5_C	Ribosomal protein S5, C-terminal domain
PF01092	Ribosomal_S6e	Ribosomal protein S6e
PF00177	Ribosomal_S7	Ribosomal protein S7p/S5e
PF00410	Ribosomal_S8	Ribosomal protein S8
PF01201	Ribosomal_S8e	Ribosomal protein S8e
PF00380	Ribosomal_S9	Ribosomal protein S9/S16
PF01984	dsDNA_bind	Double-stranded DNA-binding domain
PF01912	eIF-6	eIF-6 family
PF09173	eIF2_C	Initiation factor eIF2 gamma, C terminal
PF00749	tRNA-synt_1c	tRNA synthetases class I (E and Q), catalytic domain
PF03950	tRNA-synt_1c_C	tRNA synthetases class I (E and Q), anti-codon binding domain
PF00750	tRNA-synt_1d	tRNA synthetases class I (R)
PF09249	tRNA_NucTransf2	tRNA nucleotidyltransferase, second domain
PF04414	tRNA_deacylase	D-aminoacyl-tRNA deacylase
PF00368	HMG-CoA_red	Hydroxymethylglutaryl-coenzyme A reductase
PF01351	RNase_HII	Ribonuclease HII

Table S3. Reversibility of H_2-dependent formate production by (hyper)thermophilic *Thermococcales* depending on substrates-products concentration ratio in absence of S^0. *Thermococcus cleftensis* CL1^T possess the FHL1 complex, *Palaeococcus pacificus* DY20341^T possess the Fdh-Nor complex and *Thermococcus onnurineus* NA1 possess both. ASW medium contained: no S^0, 0.2 g L^{-1} yeast extract, headspace gas (200 kPa). Incubation temperature was 80 °C.

	Formate concentration (mM)						
	CO_2 reduction	**Formate oxidation**			**Successively during the incubation**		
Incubation time (h)	30 mM bicarbonate	2 mM bicarbonate			2 mM bicarbonate		
	H_2/CO_2 ; 80:20	N_2/CO_2 ; 80:20			H_2/CO_2 followed after 48h by N_2/CO_2 ; 80:20		
	no formate	10 to 12 mM formate			no formate		
	T. onnurineus NA1 (FHL1+Fdh-Nor)	*T. onnurineus* NA1 (FHL1+Fdh-Nor)	*T. cleftensis* CL1^T (FHL1)	*P. pacificus* DY20341^T (Fdh-Nor)	*T. onnurineus* NA1 (FHL1+Fdh-Nor)	*T. cleftensis* CL1^T (FHL1)	*P. pacificus* DY20341^T (Fdh-Nor)
0	0.0	12.0	9.6	9.4	0.0	0.0	0.0
24	4.4	7.5	9.3	8.8	1.4	0.3	0.8
48	6.3	4.6	9.0	8.9	2.0	0.5	0.9
72	6.8	3.3	8.6	8.9	0.7	0.5	0.6
96	6.9	2.6	8.3	8.9	0.4	0.4	0.4
192	7.0	2.2	8.0	8.8	0.2	0.2	0.2

Table S4. H_2 and formate concentrations used for representation of environmental concentrations measured for seawater and end-member fluids from six deep-sea hydrothermal vent fields in figure 6. At the pH used for the calculations (6.8), HCOOH was considered negligible as formic acid (pKa 3.75) does not significantly dissociate and remains at 99,999 % as $HCOO^-$, and HCO_3^- represent the dominant form (>80%) of CO_2. The diversity of total CO_2 concentrations in endmember fluids at different hydrothermal sites were not specifically used for the thermodynamic modelling. (*) This work.

Name	Localisation	H_2 (mM)	Ref.	Formate (mM)	Ref.	ΣCO_2 (mM)	Ref.
Ashadze 1	12°58'N - 44°51'W	8 - 19	(78)	0.023 - 0.028	(*)	3.7	(78)
Lost City	30°07'N - 42°07'W	1 - 15	(75)	0.036 - 0.158	(75)	0.0001 - 0.0026	(82)
Lucky Strike	37°17'N - 32°16'W	0.02 - 0.73	(78)	0.0004 - 0.0043	(*)	13 - 28	(78)
Rainbow	36°14'N - 33°54'W	13 - 16	(78)	0.007 - 0.028	(*)	16 - 17	(78)
Snake Pit (MARK)	23°22'N - 44°57'W	0.19 - 0.48	(78)	0.0018 - 0.0039	(*)	5.2 - 6.7	(78)
Von Damm (VDVF)	18°22'N - 18°47'W	9.96 - 18.3	(79)	0.082 - 0.669	(79)	2.8	(79)
Seawater		0.0004	(78)	0.001 - 0.002	(*)	2.3	(78)

75. Lang SQ, Butterfield DA, Schulte M, Kelley DS, Lilley MD. Elevated concentrations of formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field. Geochim Cosmochim Acta. 2010;74:941-52

78. Charlou JL, Donval JP, Konn C, Ondréas H, Fouquet Y, Jean-Baptiste P, et al. High production and fluxes of H_2 and CH_4 and evidence of abiotic hydrocarbon synthesis by serpentinization in ultramafic-hosted hydrothermal systems on the Mid-Atlantic Ridge. In: Rona PA, Devey CW, Dyment J, Murton BJ (eds). Diversity of hydrothermal systems on slow spreading ocean ridges. (John Wiley & Sons: Hoboken, 2010) pp 265-296

79. McDermott JM, Seewald JS, German CR, Sylva SP. Pathways for abiotic organic synthesis at submarine hydrothermal fields. Proc Natl Acad Sci USA. 2015;112:7668-72

82. Proskurowski G, Lilley MD, Seewald JS, Früh-Green GL, Olson EJ, Lupton JE, et al. Abiogenic hydrocarbon production at Lost City hydrothermal field. Science. 2008;319:604-7

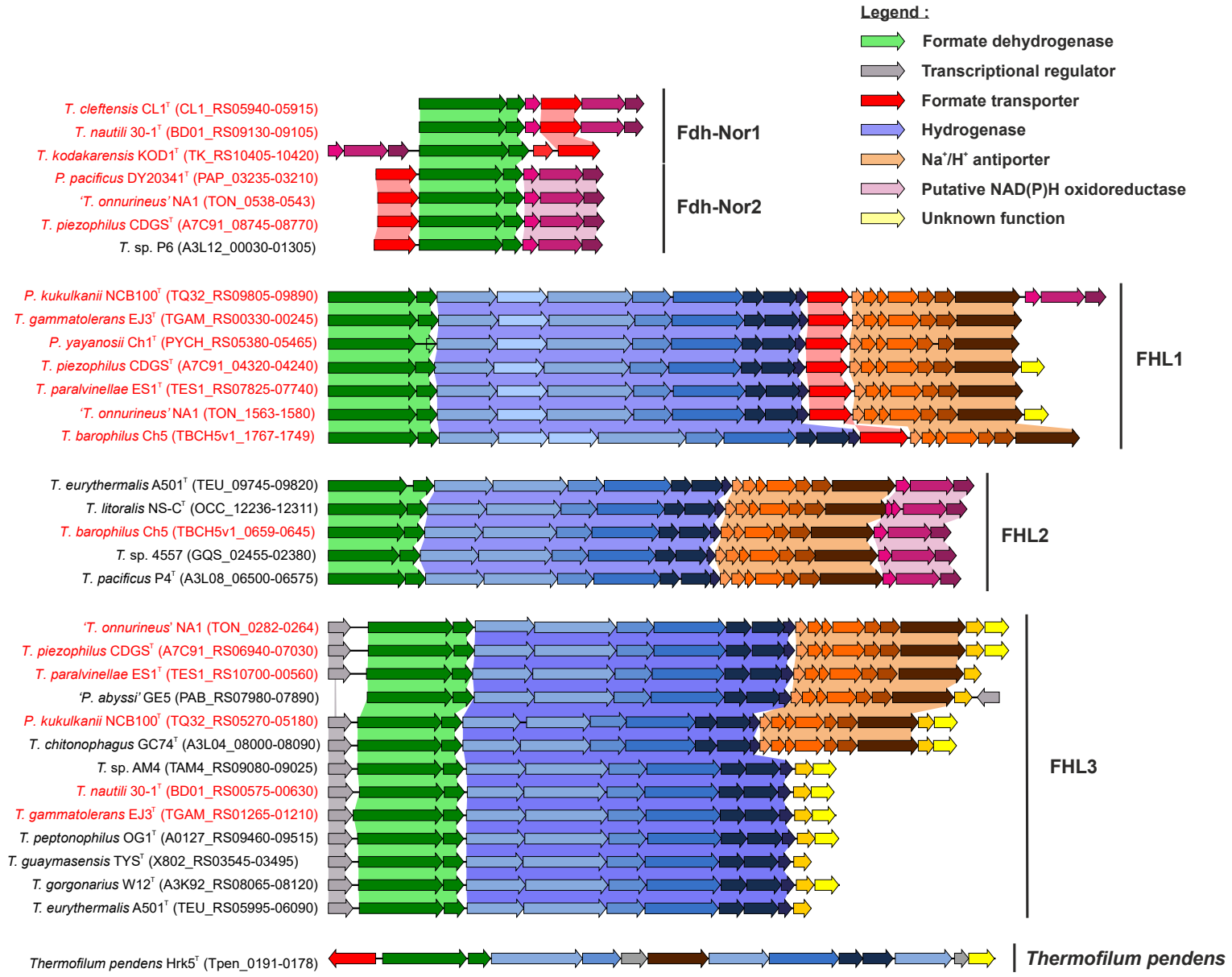

Figure S1. Comparison of the gene organisations of the *fdh-nor*, *fhl1*, *fhl2* and *fhl3* clusters from 35 genomes of *Thermococcales*. Gene clusters of *Thermococcales* that perform H_2-dependent formate production are indicated in red. GenBank accession numbers are given in brackets. Data were obtained from SyntTax software (http://archaea.u-psud.fr/synttax).

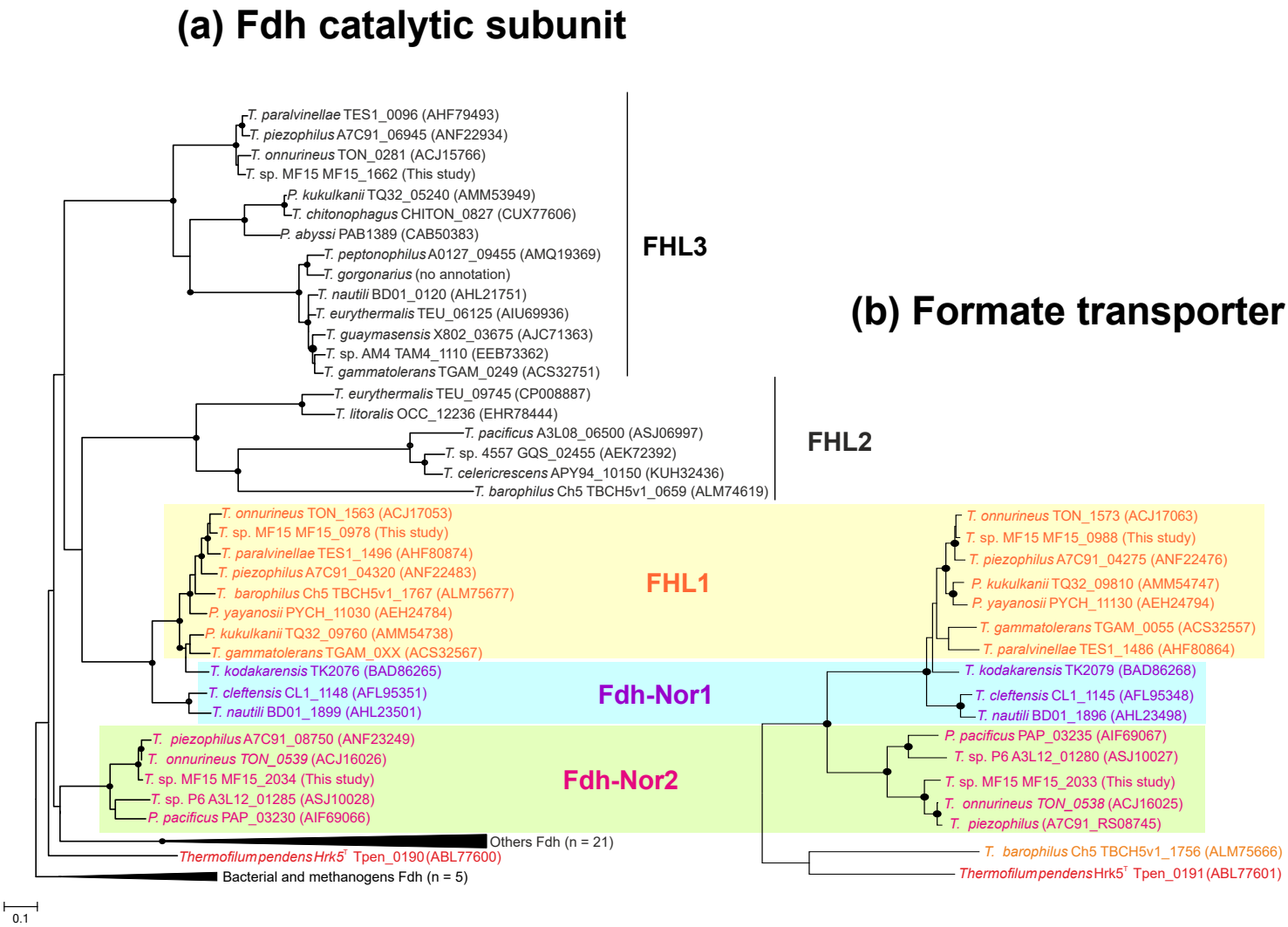

Figure S2. Phylogenetic trees based on translated partial amino acid sequences of functional genes: (a) formate dehydrogenase catalytic subunit (68 sequences) and (b) formate transporter (17 sequences). The trees were constructed using the neighbor-joining method (83) and the evolutionary distances were computed using the Dayhoff matrix-based method (84). Nodes with a bootstrap value > 80% (for 1 000 replicates) are marked in black solid dots. GenBank accession numbers are given in brackets. Scale bar represents 0.1 substitutions per amino acid position.

83. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular biology and evolution. 1987;4(4):406-25.

84. Dayhoff M, Schwartz R, Orcutt B. A model of evolutionary change in proteins. Atlas of protein sequence and structure. 1978;5:345-52.

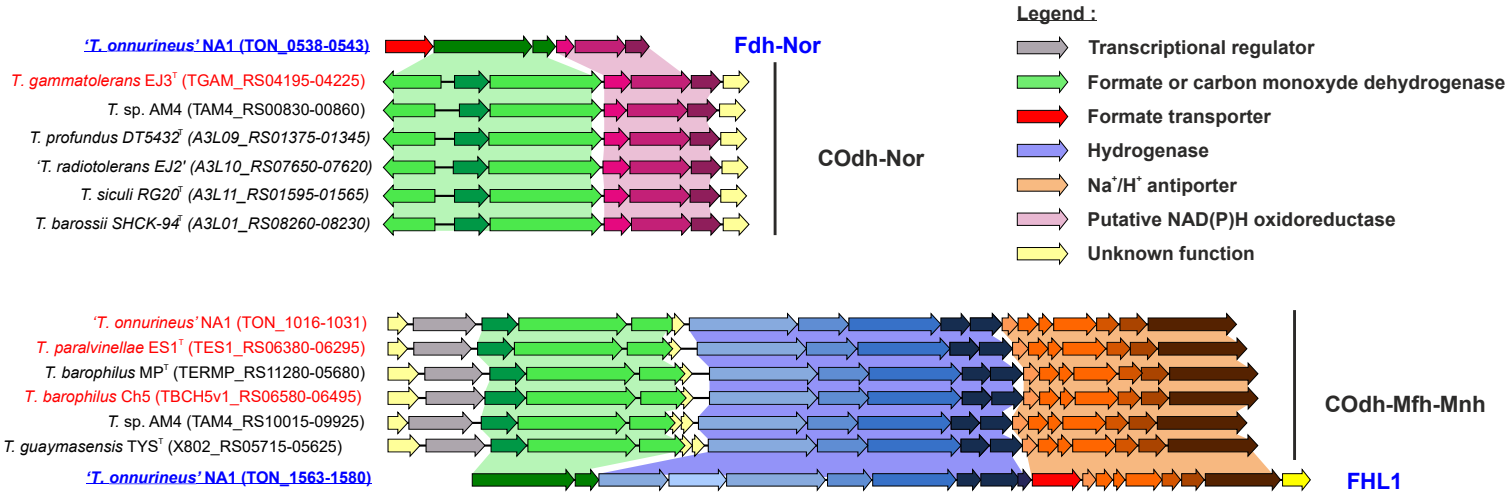

Figure S3. Comparison of the gene organisations of the *codh-nor* and *codh* clusters from 35 genomes of *Thermococcales*. Gene clusters of *Thermococcales* that perform H_2-dependent formate production are indicated in red. GenBank accession numbers are given in brackets. Data were obtained from SyntTax software (http://archaea.u-psud.fr/synttax).

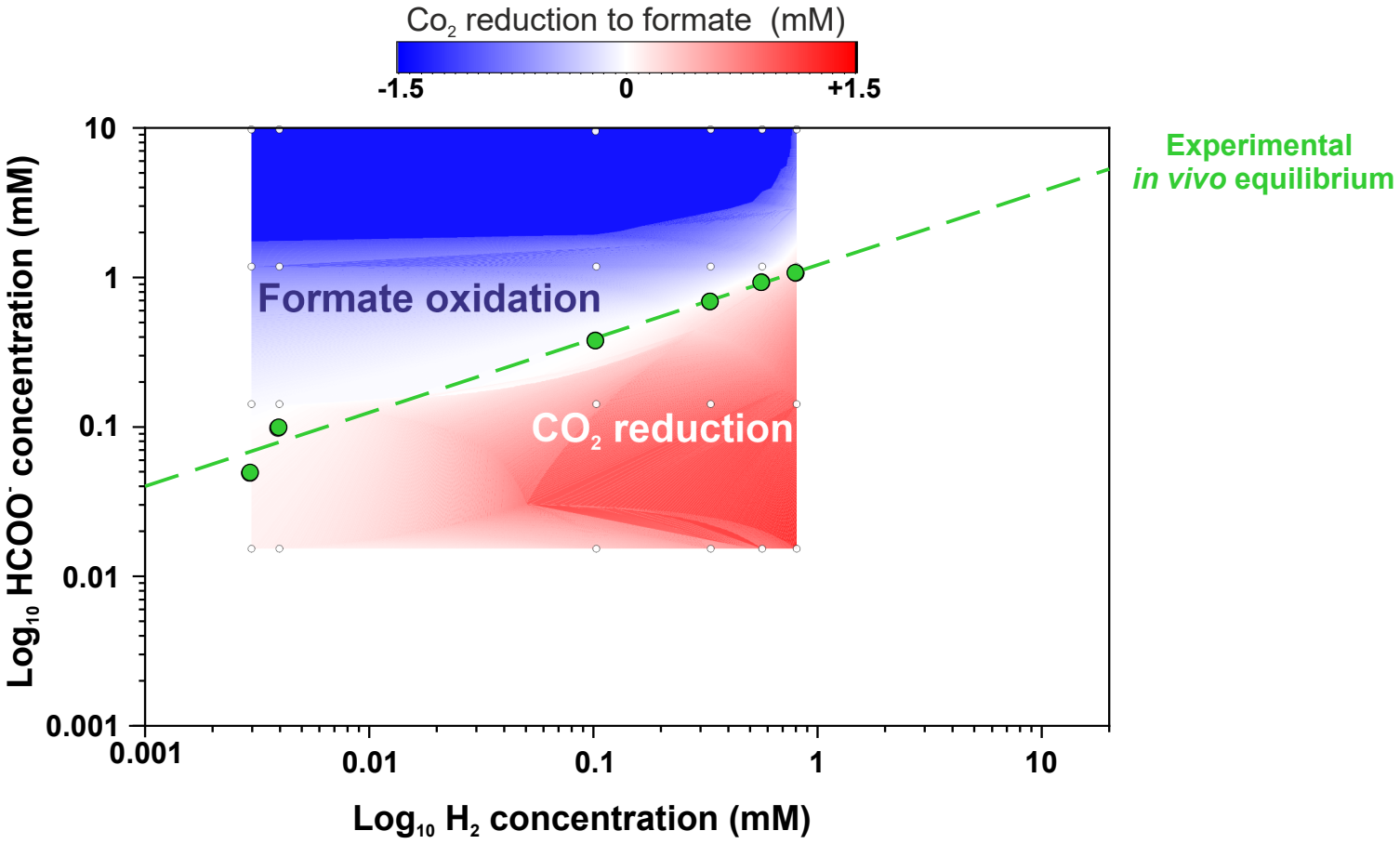

Figure S4. Cross effects of H_2 and formate on CO_2 reduction rate in *T. onnurineus* NA1 after 96 hours. The blue area indicates favourable zone for formate oxidation while the red area indicates favourable zone for CO_2 reduction and the white area when reaction is at equilibrium. Green dashed line defines the experimental *in vivo* point of chemical equilibrium of the reaction ($\Delta G' = 0$ kJ mol^{-1}). Effect of H_2 concentrations was examined in a range of 0-80 % (v/v; 200 kPa; 0-0.8 mM dissolved H_2) with CO_2 concentrations kept at 20 % (v/v). Formate concentrations were in a range of 0 to 1 mM. ASW medium contained: 2 mM bicarbonate, 0.2 g L^{-1} yeast extract, no S^0 was added, temperature 80 °C.

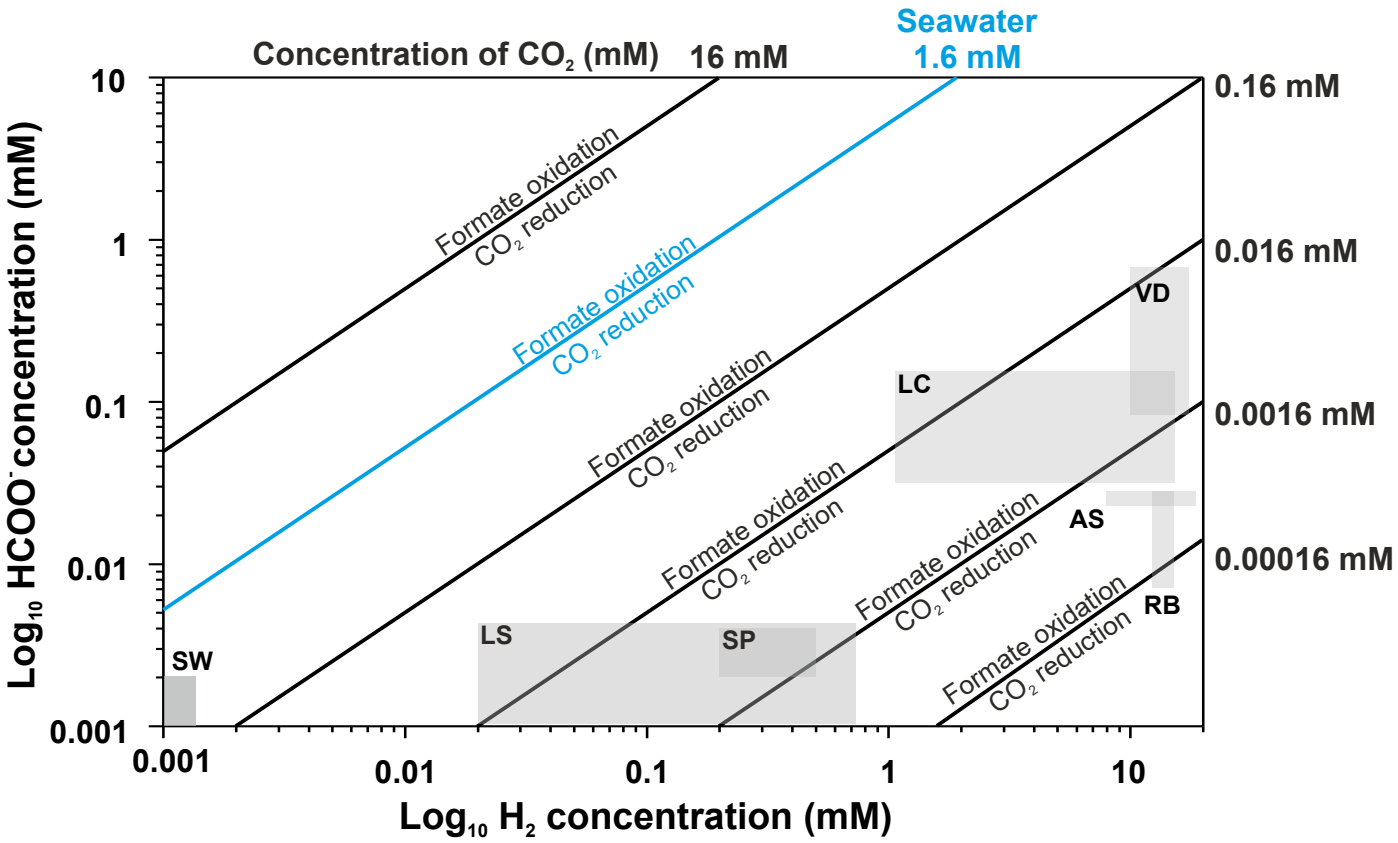

Figure S5. The direction of the reaction of CO_2 reduction depends primarily on substrates concentration. Plot representing the thermodynamic modelling of the CO_2 reduction to formate for a range of concentration of formate, H_2. and CO_2. Gibbs free energy of the reaction ($H_{2\,(g)} + HCO_3^-{}_{(aq)} \rightarrow HCOO^-{}_{(aq)} + H_2O_{(l)}$) was calculated for 80 °C and for a range of bicarbonate concentrations (0.00016 to 16 mM). 1.6 mM represents the bicarbonate concentration in seawater at a pH of 6.8 for a concentration of 2 mM total DIC. Black solid lines define points of chemical equilibrium of the reaction ($\Delta G' = 0$ kJ mol^{-1}) from thermodynamic modelling for 6 different CO_2 concentrations, 0.00016 mM, 0.0016 mM, 0.016 mM, 0.16 mM, 1.6 mM, and 16 mM. Grey squares represent environmental concentrations measured for seawater and six hydrothermal vent field (data available on Table S4).

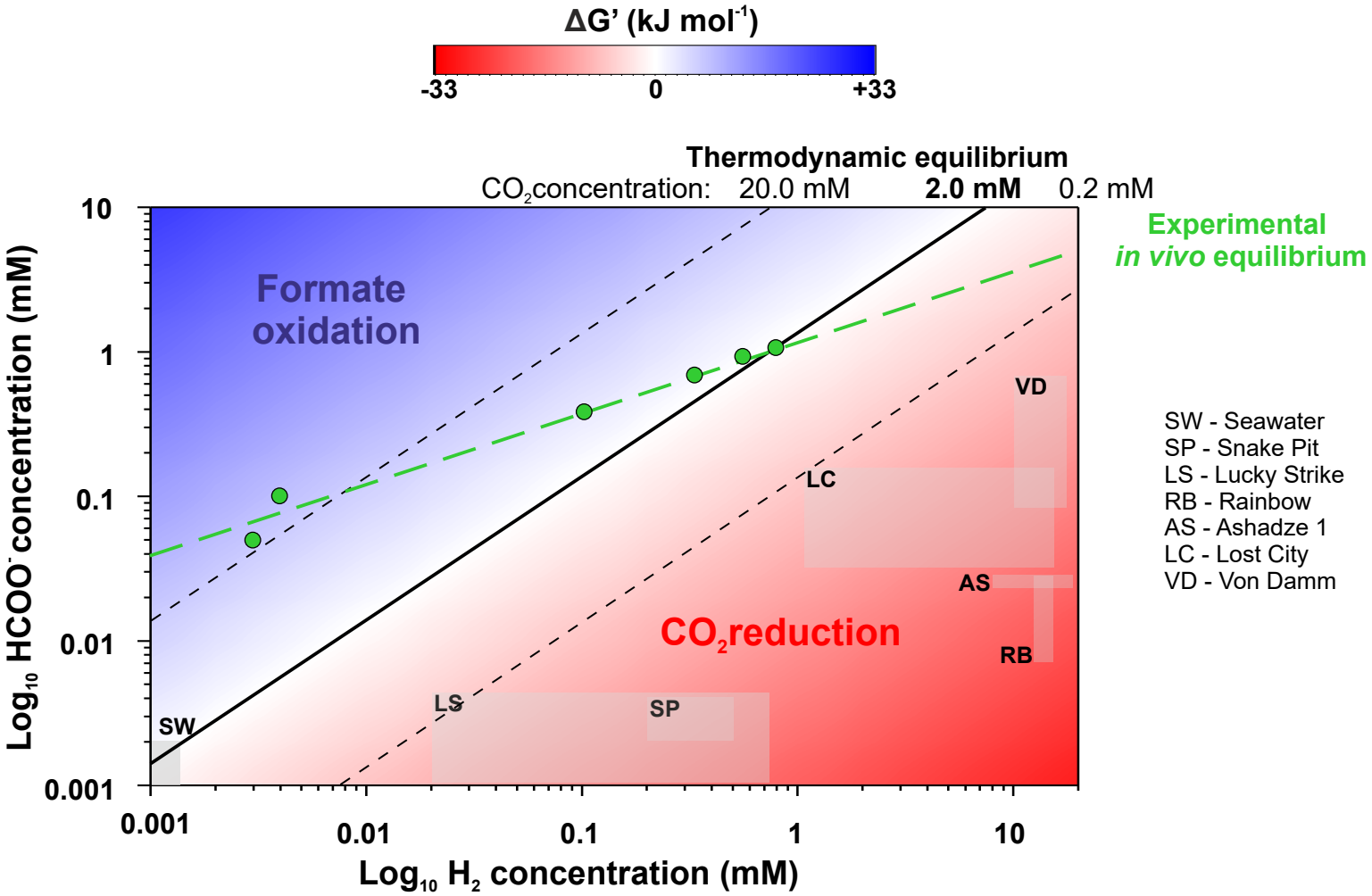

Figure S6. The direction of the reaction of CO_2 reduction depends primarily on substrates concentration. 2D-contour plot representing the thermodynamic modelling of the CO_2 reduction to formate for a range of concentration of formate and H_2. Gibbs free energy of the reaction ($H_{2(aq)} + HCO_3^-{}_{(aq)} \rightarrow HCOO^-{}_{(aq)} + H_2O_{(l)}$) was calculated for 80 °C and 2 mM of bicarbonate and H_2 was considered as dissolved. The blue area indicates conditions for which formate oxidation occurs while the red area indicates conditions for which CO_2 reduction to formate occurs. Black solid line and dashed lines define points of chemical equilibrium of the reaction ($\Delta G' = 0$ kJ mol^{-1}) from thermodynamic modelling for three different bicarbonate concentration, 2.0 mM, 0.2 mM, 20.0 mM respectively. Green dashed line defines the experimental *in vivo* point of the chemical equilibrium of the reaction measured for *T. onnurineus* NA1 in culture medium after 96 hours of incubation at 80 °C (see details Fig. S4). Effect of H_2 concentrations were examined in a range of 0-80 % (v/v; 200 kPa; 0-0.8 mM dissolved H_2) with CO_2 concentrations kept at 20 % (v/v). Effect of formate concentrations were examined in a range of 0 to 1 mM. ASW medium contained: 2 mM bicarbonate, 0.2 g L^{-1} yeast extract and no S^0 was added. Grey squares represent environmental concentrations measured for seawater and six hydrothermal vent field (data available on Table S4).

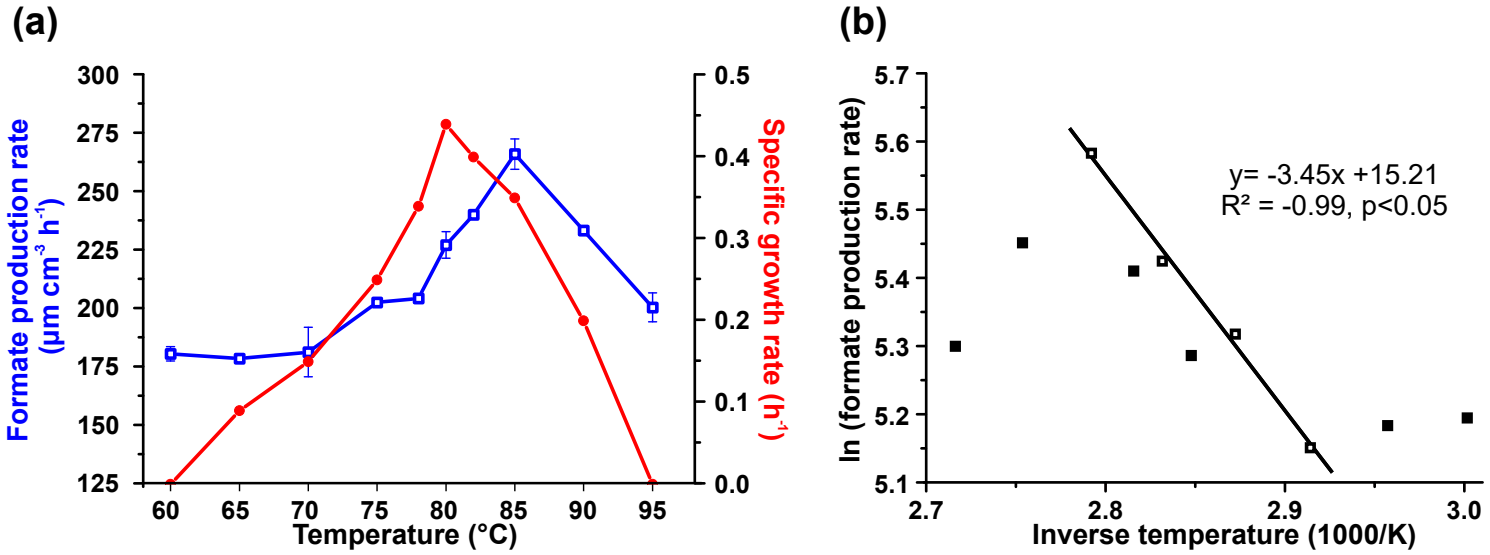

Figure S7. Effect of temperature on formate production rates in *T. onnurineus* NA1. **(a)** Formate production rates (blue squares) determined from the temperature-gradient incubation experiment after 24 hours of incubation and specific growth rate (red circles) obtained from (31). ASW medium contained: 30 mM bicarbonate, H_2/CO_2 (80:20, v/v; 200 kPa), and 0.2 g L^{-1} yeast extract. Error bars show standard error (n = 3). **(b)** Arrhenius plot of the data from panel (a).

31. Bae S-S, Kim Y-J, Yang S-H, Lim J-K, Jeon J-H, Lee H-S, et al. *Thermococcus onnurineus* sp. nov., a hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent area at the PACMANUS field. J. Microbiol. Biotechnol. 2006;16:1826-31.