## **Supporting Information**

## Francez-Charlot et al. 10.1073/pnas.0810291106

## **SI Experimental Procedures**

Strain Constructions. For complementation and overexpression studies in M. extorquens AM1, plasmid pCM62 and its derivative pCM80, which contains the strong  $p_{mxaF}$  promoter, were used (1). To construct the phyRD190A allele, an overlapping PCR strategy was used. Two PCR products were generated with pCM80\_regI\_PstI/phyRD190A primers R and pCM80\_regI\_KpnI/phyRD190A F. PCR products were mixed and reamplified with primers pCM80\_regI\_PstI/ pCM80\_regLKpnI. The resulting product and the wild-type phyR allele generated with the same primers were cloned into pCM62 to generate pCM62\_phyRD190A and pCM62\_phyR, respectively. In these vectors, phyR expression was driven from its native promoter. Plasmid pCM80\_PhyRNterm to overexpress the N-terminal domain of PhyR was derived from plasmid pBG11 (2) cut with SalI and self-circularized. To overexpress nepR (rmq12793), a PCR fragment containing the nepR ORF and 38 nucleotides upstream of the start codon was generated with primers 3115 -137 F and 3115R3 and was cloned into TOPO PCR2.1 (Invitrogen). The nepR ORF was then excised with HindIII and XbaI and cloned in pCM80 to generate pCM80\_nepR. To construct pCM80\_nepR\_phyR for coexpression studies, a PCR fragment containing phyR ORF and 19 nucleotides upstream of the start codon made with primers PhyR F1 PstI and PhyR R1 XbaI was cloned into pCM80\_nepR digested with PstI and XbaI. In the resulting plasmid pCM80\_nepR\_phyR, phyR was located downstream of nepR, with 40 nucleotides between the nepR stop codon and the phyR start codon.

To overproduce PhyR with a C-terminal hexahistidine tag, the corresponding gene was amplified with primers phyR EXP F1 and phyR EXP R1 and cloned into the pET24b vector using NdeI and NotI restriction sites to generate pET24b-phyR. To overoverproduce and purify the PhyR mutant D190A and the N-terminal ECF domain of PhyR, PCR products generated with primers phyR TP F/phyR TP R and phyR TP F/phyRNterm R were cloned into the Champion pET101 vector (Invitrogen), adding a C-terminal hexahistidine tag to the synthesized protein (plasmids pET101-phyRD190A and pET101-phyRNterm, respectively). NepR and  $\sigma^{\text{EcfG1}}\Delta 1-64$  were overexpressed with an N-terminal decahistidine tag. The corresponding genes were amplified with primers nepR for/nepR rev or 4906 wo Nterm F1/4906 R2 and cloned into the pET16b vector using BamHI and NdeI restriction sites to generate plasmids pET16b-nepR and pET16b-ecfG1 $\Delta$ 1–64, respectively.

For *in vivo* interaction studies, the *nepR* gene was amplified using the oligonucleotides nepRhis-f/nepRhis-r, adding a C-terminal hexahistidine tag to the construct, and cloned into pCM80 to generate pCM80-nepR-his.

To construct the unmarked ecfG1 (rmq08147) deletion mutant, an overlapping PCR was used: 2 PCR fragments of 600 bp

corresponding to upstream and downstream flanking regions of a rmg08147 ORF were generated with primers META1\_4906 Up F/META1\_4906 Up R and META1\_4906 Down F/META1\_4906 Down R; PCR products were then mixed and used in PCR with primers META1\_4906 Up F and META1\_4906 Down R. The 1.2-kb fragment obtained corresponding to merged 600-bp upstream and downstream regions was cloned into the broad-hostrange sacB-based vector for unmarked allelic exchange pCM433 (3) using NdeI and AgeI restriction sites. The resulting plasmid, pCM433-ecfG1, was transformed in M. extorquens by electroporation, and transformants were selected on tetracycline. For sucrose selection, colonies were resuspended in MM and plated on MM supplemented with methanol, succinate, and 5% sucrose. Single colonies were tested for tetracycline resistance, and tetracycline-sensitive strains were analyzed by PCR using primers located outside the region where recombination occurred. Mutant strains were further analyzed by sequencing the region encompassing the locus where recombination had occurred.

Overproduction and Purification of PhyR, NepR, and  $\sigma^{EcfG1}$ . For in vitro interaction studies, the different PhyR, NepR, and  $\sigma^{\text{EcfG1}}\Delta 1$ -64 constructs were produced in *E. coli* BL21 (DE3). Expression was induced at an  $OD_{600}$  of 0.8 with 1 mM isopropyl  $\beta$ -D-thiogalactoside. Four hours after induction, cells were harvested by centrifugation and stored at -20 °C. For cellextract preparation, thawed cells were resuspended in 10 mL of TNI5 buffer [20 mM Tris/HCl (pH 8.0), 150 mM NaCl, and 5 mM imidazol], and a protease inhibitor mixture was added (Complete, EDTA-free; Roche). Cell disruption was performed using Bugbuster reagent (Novagen). Intact cells and cell debris were removed by centrifugation (15 min, 8,000  $\times$  g, 4 °C). Afterward, the cell-free extract was subjected to ultracentrifugation (1 h,  $150,000 \times g, 4$  °C). Recombinant protein present in the supernatant was purified by nickel chelate affinity chromatography using Ni-NTA (HisTrap HP; GE Healthcare). After immobilization of the protein, the column was washed with 15 column volume TNI20 buffer (TNI buffer with 20 mM imidazol), and bound protein was eluted with TNI200 buffer (TNI buffer with 200 mM imidazol). Fractions containing the protein were pooled, and the elution buffer was exchanged against TG buffer [20 mM Tris/HCl (pH 8.0) and 10% (vol/vol) glycerol] and stored at −20 °C.

NepR and  $\sigma^{\text{EcfG1}}\Delta 1-64$  were purified using Ni-NTA as described for PhyR. Ni-NTA pooled fractions were then applied to a Superdex75 10/300 gel filtration column (Amersham Biosciences) equilibrated with buffer A [50 mM Tris/HCl (pH 8.0) and150 mM NaCl] at a flow rate of 0.5 mL/min. The major peak fractions were pooled and concentrated, and the buffer was exchanged against TG buffer. Because of stability problems with NepR, purified protein was used for *in vitro* interaction studies within 24 h. During the course of  $\sigma^{\text{EcfG1}}$  purification, it was noted that  $\sigma^{\text{EcfG1}}$  precipitated at concentrations >0.5 mg/mL.

Marx CJ, Lidstrom ME (2001) Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria. *Microbiology* 147:2065–2075.

Gourion B, Rossignol M, Vorholt JA (2006) A proteomic study of Methylobacterium extorquens reveals a response regulator essential for epiphytic growth. Proc Natl Acad Sci USA 103:13186–13191.

<sup>3.</sup> Marx CJ (2008) Development of a broad-host-range sacB-based vector for unmarked allelic exchange. *BMC Res Notes* 1:1.

	2.1	2.2	2.3	2.4	
Blr7797	5-DSLENDILAAVPSLEAFAISI	SCNADRADDLVOETLLRALAN	IDSFOPGSNLPAWLFTI	LRNLFRSDYRKRREVEI	AEGNYAKTLKTOPSONAHLEFEEFRTZ
RPA4225	5-DSLRDDILAAVPSLRAFAISI	SGNADRADDLVOETLLRALAN	IDSFOPGSNLPAWLFTI	LRNLFRSDYRKRRREVEI	ADGSYAKTLKSOPGOTAHLEFEEFRAZ
<sub>O</sub> EcfG1	-70-PEFRNGLLSAVPSLRAFALSI	TSNPARSDDLVODTLLKGWOH	RARFOPGTNLNAWLFTI	LRNIFYSDHRKRVREVEI	ODGSYAARLATAPHOGDRLDVEDLOTA
BME10371	-13-PPFKRELLASLPSLRAFAVSI	IGOHDRADDLVODTIMKAWAK	OESFEVGTNMKAWLFTI	LRNEFYTOMRKRGREVOI	TDGMFSEOLAVHPSOYGTLDLODFRAZ
SMc01506	6-QEFKREMLAALPSLRAFAMSI	IGRHDRADDLVQDTIMKAWAB	ODHFEIGTNMKAWLFTI	LRNELYSOMRKRGREVOI	SDGHLTETLAHHPEQYGSLDLQDFRR
M113697	6-QGFKTDLLGAIPSLRAFAVSI	TONADKADDLVQETLVKAWDF	HESFOPGTNLKAWLFTI	LRNEFYSOMRKRGREVOI	SDGIMTARLAVHPAQHGQLDLKDFRGA
CC3475	-17-NSFKRELVTLIPHLRAFARMI	TCDPTAADDLAQDAMMKAWDA	RASFQLGTNMKAWTEMI	LRNQFYSEKRRSWRQSQI	DQEAAERTLVAVDDPEAPVALDELRQ
B117795	6-MSRSQLVAEHLPLLRRYARAI	T <mark>G</mark> SQASG <mark>D</mark> AYVAAM <mark>LEAMLGE</mark>	PAVLDE SHGPRAGLERL	FTQIWNSV	SVNDDSEVTTLPMPPERR
RPA4223	1-MSRSQVVAEHLPLLRRYARAI	T <mark>G</mark> SQTSG <mark>D</mark> AY <mark>VGAM</mark> LEALLQI	PALLDERHGPRPGLFRL	FTQI <mark>WNS</mark> V	ALNDDADVRTLPMPSER
PhyR	1-MSTAQLVVQHLPYLRRYARAI	T <mark>G</mark> SQVAG <mark>D</mark> AYVAATLETLVNE	PETLGRSTNVKADLFRV	FTRIWNSL	S <mark>V</mark> NGHSDQVQHDLPAEVR
BME10372	1-MTLSTR <mark>IAPH</mark> LPYLRRFSRAI	T <mark>G</mark> SQSS <mark>GD</mark> AY <mark>VAAA</mark> LEALIAI	VGIFPEASS <mark>D</mark> RIGLYRL	FCNLYKNA	SIRMPETSSEFAWETHAARN
SMc01504	1-MTLSTRIAPFLPYLRRYSRAI	T <mark>G</mark> SQTSG <mark>D</mark> AY <mark>VAAVLEALIAI</mark>	TSIFPEASS <mark>D</mark> RVALFRL	FTSM <mark>F</mark> GSS	SVLVPEPVSPFAWEQRASVN
Mlr3700	1-MSLSAT <mark>IAPH</mark> LPFLRRFSRAT	7S <mark>G</mark> SQESG <mark>D</mark> AL <mark>VAAM</mark> LEAIIAL	VDIFPDASS <mark>D</mark> RIALYKV	FAKL <mark>F</mark> TSV	AIR <mark>V</mark> PQEQAQSAWEQRAAAN
CC3477	1-MSLLARLAPHLPYIRRYARAI	T <mark>G</mark> DQATG <mark>D</mark> HYVRVA <mark>LEALAAG</mark>	ELVLDANLSPRVALYRV	FHAIWLSSG	AQLE <mark>VGHDQGLHAGDD-AAQ</mark> R
	4.1	4.2			
B1r7797	IIEKIIPQDQRDALIIIVGASCISYD	DAASICGCAVGTIKSRVNRAR	SKLAALLYVDGAEDFGP	DETVRAVIGGSGG-	
RPA4225	IIDKIIPQDQREALIIIVGASCESYE	DAAAICGCAVGTIIKSRVNRAR	SKLSALLYVDGAEDFGP	DDTVRAVIGGNG	
GECIGI		BAATIMGCKVGTVKSRVSRAR	GRLAELLGYD-EEDLSS	DRLIQSAMPDA	
BME10371		BAAEICGCAVGTIKSRVSRAR	ARLOEILOIEGEADIGPI	DADSSRATPRSFA-	
SMC01506	LEOLDEDODEA TILLICA SCESSIO	EAATICGCALGTIKSRVINRAR	URLOEILOVRGENDIGP		
CC3475	LSMUDA FORFALLINGA COPAVE	FAFTCNCAVGTWKSDVSDAD	PALOATLESCAVDPDCA	AACDAMPATLADAEDI.SZ	T.P
B117795	LSNT TDL DPOADLLISLE CESE	FVA FILSTDVA FUDDLADA C	PEMA A E TA TOVI. T TEDE	TETAMDLESINKNIG-16	6-
BPA4223	USATTPLPROAFILLISLEGESE	EVA EVILOVDVA EVROLVETA G	KEMAEE TATDULTIEDE	TETAMDLESIJIKGLG-16	1-
PhyR	LOT TPLPROAFLISCLEGESE	DAGVILDVDVSKVRDLVDEAG	RELAADMATETLITEDE	PLIAMDLEALVEGLG-16	1-
BMEI0372	AHUAPLPROAFULWAVE GENDI	FAAFILEVDOAFFGRUISTAS	GEISROVATRIMITEDE	PLIAMDIEOMVESLG-16	3-
SMc01504	LAAVSPLAROAFLLVSVEADAP	FAAEVLDVDAAKTAALLDRAS	OEISROVATEIMIIEDE	PLIAIDIEOMVESLG-16	- 3-
Mlr3700	LNAIAPLPROAFLLVAVE GFSED	EAAEILDVGDOOPSELLAOAS	NEISROVATDVLIIEDE	PLIAMDIEEMVESLG-16	3-
CC3477	LMRIAPRSRQAFLLUALECFTPU	EAAQILDCDFGEVERLIGDAQ	AEIDAELATEVLIIEDE	PVIAADIEALVRELG-16	4 –

**Fig. S1.** Sequence alignment of PhyR and  $\sigma^{\text{EcfG1}}$  homologues of selected Alphaproteobacteria. The alignment shows that PhyR homologues lack a part of the region 2.4, involved in –10 binding. In addition, region 4.2, which is involved in –35 binding, shows degeneration in PhyR homologues, compared with the high degree of conservation observed among the different ECF sigma factors. The alignment was generated with Clustal\_W. Sequences are based on data deposited in GenBank, except for MIr3700, where a longer version containing the ECF sigma factor-like domain was taken. Protein labels indicate species as follows: *Bradyrhizobium japonicum USDA 110*, BIr/BII; *Rhodopseudomonas palustris CGA009*, RPA; *Brucella melitensis 16M*, BMEI; *Sinorhizobium meliloti 1021*, SMc; *Mesorhizobium loti MAFF303099*, MII; *Caulobacter crescentus CB15*, CC. Residues conserved in 75% of the sequences are highlighted (BLOSUM\_62). Regions 2 and 4 conserved among  $\sigma^{70}$  family are shown above the alignment.

## **Other Supporting Information Files**

Table S1 Table S2 Table S3

VAS PNAS