The GbsR family of transcriptional regulators: functional characterization of the OpuAR repressor

Stefanie Ronzheimer^{1¶}, Bianca Warmbold^{1¶}, Christian Arnhold¹, and Erhard Bremer^{1,2}*

¹Laboratory for Microbiology, Department of Biology, Philipps-University Marburg,

Marburg, Germany

²LOEWE-Center for Synthetic Microbiology, Philipps-University Marburg, Marburg, Germany

[¶]These authors have contributed equally to this work

For correspondence please contact:

Dr. Erhard Bremer, Philipps-University Marburg, Dept. of Biology, Laboratory for Microbiology, Karlvon-Frisch-Str. 8, D-35032 Marburg, Germany. Phone: (+49)-6421-2821529. Fax: (+49)-6421-2828979. E-Mail: bremer@staff.uni-marburg.de

	50	60	70	80	90	100	110
	1	 	I	I	I	<u> </u>	<u>l</u>
GbsR Bacillus subtilis subtilis 168	RDE-MTLDEMREELQ	MSKPSMST	GVKKLQDLNV	VKKTFHRGIR	KHTEVAEKD	FKFFTNFF	PPKWEREV
Alteromonas naphthalenivorans SN2	ADEPISAQEIADALN	ISRGNTSM	IGLKELQSWRL	VKQHLVPGER	KEFFIAGGD	WILANRVE	EERRKREI
Cellvibrio japonicus Ueda107	SKEPLCADDITEALG	ISRSNVSN	IGL KELTSWEL	VKLQHRPGER	KEYYSAPGD	MOTAKTLI	EQRRKREM
Herbaspirillum seropedicae SmR1	CGRALNADEIADYLS	FSRSNVSN	IGLKELQSWRL	VKLLHKPNDR	REYFEPPGD	WDIFKVLL	EERRRRE I
Collimonas fungivorans Ter331	LGQPLNADQIADYLT	FSRSNVSN	IGLKELQSWRL	VKLLHQPGDR	REYFEPPKD	WDIFKTLL	EERRRRE I
Methylibium petroleiphilum PM1	SPTPLNADDIAEQLE	FSRSNVSM	IGLKELQAWRL	VRLRHLPGDR	REYFEAPTD	WEIFRILA	EERRRRE I
Delftia acidovorans SPH-1	SPEPLNADQIAETLE	FSRSNVSN	IGLKELQAWRL	VQLRHQPGDR	REYFEAPQD	WE I F R R L AI	EERRRRE I
Variovorax paradoxus B4	SERALNADEIAELLE	FSRSNVSM	IGLKELQAWRL	VHLRHHPGDR	REYFEAPAD	AME I FRVLA	EERRRRE I
Ramlibacter tataouinensis TTB310	SPRPLNADEIAETLE	FSRSNVSN	IGLKELQSWRL	VRLKHLPGDR	REYFEAPSD	AME I FRVLA	EERRRRE I
Jannaschia sp. CCS1	SSTPLNAEQITEGLG	VSRSNTSM	IGL KELQAWNL	VRLRHVPNDR	RDYFTTPEDL	WEITRILL	AERKKRE I
Chromobacterium violaceum ATCC 12472	SEKPLNADEIGEAIG	CSRSNVSN	IGL KELQ SWRL	VKLQHFPGDR	REYFSTPED	AWLIFKTLA	EERKKREV
Pseudogulbenkiania sp. NH8B	SPKPLNADDIVEAIG	CSRSNTSM	IAL KELQAWRL	VKLQHLPGDR	REYFSTPDD	/WAIFKTLA	EERQKREV
Arenimonas donghaensis DSM 18148, HO3-R19	SARPL TADE I TDRL G	VSRSNVSN	IGLKELSSWRL	VRLSHQPGDR	RDFYTAPED	AWA I FKTLA	EERQRREV
Thiomonas arsenitoxydans 3As	SERPLNADEIAAALD	FSRSNVSN	IGL KEL GAWNL	VRMQHLPGDR	RDYFSAPED	AWA I FRTLA	EERRKRE I
Thiomonas intermedia K12	SERPLNADEIAAALD	FSRSNVSN	IGLKELGAWNL	VRMQHLPGDR	RDYFSAPED	AWA I FRTLA	EERRKRE I
Polaromonas naphthalenivorans CJ2	SPRAL NADDI GEALA	FSRSNVSM	IGL KELQ SWNL	VRLQHLPNDR	REYFOAPED	AWA I FRTLA	EERRKRE I
Polaromonas sp. JS666	SPRAL NADE I GEALA	FSRSNVSM	IGL KELQ SWNL	VRLQHLPNDR	REYFSAPDD	AWA I FRTLA	EERRKRE I
Cupriavidus metallidurans CH34	ASRPLNADEIAESLG	FSRSNVS	GLKELESWSL	VRL SHQPGDR	REYFSAPDD	AWA I FRTLA	EERRRRE I
Ralstonia solanacearum GMI1000	SREPINADEIAEALG	FSRSNVS	GLKELESWKL	VRLTHKPGDR	REYFSAPDD	WTIFRTLA	EERRKRE I
Ralstonia solanacearum Po82	AREPINADEIAEALG	FSRSNVS	GLKELESWKL	VRLTHKPGDR	REYFSAPDD	WTIFRTLA	EERRKRE I
Ochrobactrum anthropi ATCC 49188	SPEPL CADQ I VDAL G	VSRSNVSM	IG I RELQGWNL	VLLKHIPGDR	RDFFTTPDD	/WQILRTLA	EERKKRE I
Brucella microti CCM 4915	SPEPL CADE I VEVL G	VSRSNVSM	IG I RELQ GWNL	VLLKHIPGDR	RDFFTTPDD	/WQILRTLA	EERKKRE I
Brucella suis ATCC 23445	SPEPL CADE I VEAL G	VSRSNVSM	IG I RELQGWNL	VLLKHIPGDR	RDFFTTPDD	AWQ IL RTLA	EEPKKRE I
Brucella canis ATCC 23365	SPEPL CADE I VEAL G	VSRSNVSM	IG I RELQ GWNL	VLLKHIPGDR	RDFFTTPDD	/WQILRTLA	EERKKRE I
Brucella pinnipedialis B2/94	SPEPL CADE I VEAL G	VSRSNVSM	IG I RELQ GWNL	VLLKHIPGDR	RDFFTTPDD	MQILRTLA	EERKKRE I
Brucella abortus NCTC 10505	SPEPL CADE I VEAL G	VSRSNVSM	IG I RELQ GWNL	VLLKHIPGDR	RDFFTTPDD	ANQ IL RTLA	EERKKREI
Brucella melitensis ATCC 23457	SPEPL CADE I VEAL G	VSRSNVSM	IG I RELQ GWNL	VLLKHIPGDR	RDFFTTPDD	/WQILRTLA	EERKKRE I
Brucella ovis ATCC 25840	SPEPL CADE I VEAL G	VSRSNVSM	IG I RELQ GWNL	VLLKHIPGDR	RDFFTTPDD	AWQ IL RTLA	EERKKRE I
Agrobacterium tumefaciens F2	SPQPL CADDIVEALG	ISRSNVSM	ISLKELQAWNL	AILKHFPGDR	RDFFTTPED	AWQILRTLA	EERKKREI
Agrobacterium sp. H13-3	SPKPLCADDIVESLG	ISRSNVSM	ISLKELQAWNL	AILKHFPGDR	RDFFTTPED	WQILRTLA	EERKKREI
Acidiphilium cryptum JF-5	SDRPLPADE AETLG	FSRSNVSM	IGLKELQSWRL	VRLQHLPGDR	REHFSTPDD	WQIVRTLA	EERRRRE I
Acidiphilium multivorum AIU301	SDRPLPADE AETLG	FSRSNVSM	IGL KELQ SWRL	VRLQHLPGDR	REHFSTPDD	WQIVRTLA	EERRRRE I
Methylobacterium radiotolerans JCM 2831	AERPLNADEIVERLG	VSRSNVSM	IGL KELQAWNL	VRLQHRPGDR	RDYFTTPED	WQIVRTLV	EERKKREV
Methylobacterium nodulans ORS 2060	AERALNADEIVERLG	LSRSNVSM	IGL KELO AWNL	VRLQHVPGDR	RDYFSTPED	WQIVRTLV	EERKKREV
Methylobacterium sp. 4-46	AERALNADEIVERLG	LSRSNVSM	IGLKELQAWNL	VRLLHRPGDR	RDYFSTPED	WQIVRTLV	EERKKREV

Supplementary Figure S1 Amino acid sequence alignment of CydE-type proteins. The amino acid sequences of 34 CydE-type proteins (Xia et al., 2018) were compared to the amino acid sequence of the *B. subtilis* GbsR protein (Nau-Wagner et al., 2012) by an alignment using the MAFFT server (Katoh et al., 2017). Depicted are only the N-terminal domains of these proteins, containing the amino acids corresponding to the winged helix-turnhelix DNA-binding motif (lite blue), the inter-domain linker (purple) and of the likely inducer-binding site (reddish) of GbsR. Highly conserved amino acids are shaded in grey



Supplementary Figure S2 In *silico* models of the GbsU substrate-binding proteins. Depicted are the substrate-binding proteins encoded within the putative glycine betaine synthesis gene clusters of (A) *Halobacillus halophilus* DSM 2266 and (B) *Virgibacillus sp.* SK37 based on the crystal structure of (C) ProX of *Borrelia burgdorferi* [PDB entry 3TMG; (SSGCID, Gardberg, A., Fox, D., Staker, B., Stewart, L.; *to be published.*)]. All of these substrate-binding proteins show a ligand-binding site similar to the OpuAC protein of *B. subtilis.* Its crystal structure in complex with the ligand glycine betaine [PDB entry 2B4L (Horn et al., 2006)] is depicted in panel **D**. Aromatic amino acids likely to be involved in ligand-binding of the *H. halophilus* DSM 2266 and *Virgibacillus sp.* SK37 GbsU proteins are highlighted in red. Glycine betaine, the substrate of the OpuAC protein of *B. subtilis* (Horn et al., 2006) is shown in blue.

Bacillus subtilis subtilis 168 Bacillus sp. JS Bacillus subtilis spizizenii TU-B-10 DSM 15029 Bacillus athrophaeus 1942 Bacillus amyloliquefaciens SQR9 Bacillus methylotrophicus JJ-D34 Bacillus amyloliquefaciens DSM Bacillus licheniformis ATCC 14580 Bacillus pumilus SAFR-032 Bacillus pumilus B6033 Bacillus selenitireducens MLS10 Halobacillus halophilus DSM 2266 Tetragenococcus halophilus NBRC 12172 Staphylococcus saprophyticus saprophyticus ATCC 15305 Staphylococcus xylosus HKUOPL8 Staphylococcus aureus aureus ATCC 25923 Staphylococcus epidermidis ATCC 12228 Staphylococcus warneri SGI Staphylococcus pasteuri SP1 Staphylococcus haemolyticus JCSC1435 Staphylococcus naemolyticus 5CSC1455 Staphylococcus carnosus carnosus TM300 Staphylococcus lugdunensis HKU09-01 Staphylococcus pseudintermedius HKU10-03 Virgibacillus sp. SK37 Bacillus megaterium DSM 319 Bacillus lehensis G1 Bacillus clausii KSM-K16 Bacillus atrophaaeus 1942 Bacillus subtilis subtilis 168 Bacillus sp. JS Bacillus subtilis spizizenii TU-B-10 DSM 15029 Bacillus amyloliquefaciens SQR9 Bacillus methylotrophicus JJ-D34 Bacillus amyloliquefaciens DSM 7 Bacillus atrophaeus 1942 Bacillus subtilis subtilis 168 Bacillus sp. JS Bacillus subtilis spizizenii TU-B-10 DSM 15029 Bacillus amyloliquefaciens SQR9 Bacillus methylotrophicus JJ-D34 Bacillus amyloliquefaciens DSM 7 Bacillus licheniformis ATCC 14580 Bacillus pumilus SAFR-032 Bacillus pumilus B6033 Bacillus coagulans 2-6 Virgibacillus sp. SK37 Bacillus lehensis G1 Desulfotomaculum reducens MI-1 Geobacillus sp. Y412MC10 Paenibacillus lactis 154 Paenibacillus polymyxa M1 Paenibacillus terrae HPL-003 Paenibacillus alvei DSM 29 Paenibacillus sabinae T27 Paenibacilius sabinae 12/ Thermobacilius composti KWC4 DSM 18247 Paenibacilius sp. JDR-2 Paenibacilius larvae larvae 08-100 DSM 25719 Bacilius infantis NRRL B-14911 Bacillus selenitireducens MLS10 Bacillus cellulosilyticus DSM 2522 Bacillus lehensis G1 Bacillus clausii KSM-K16 Virgibacillus p. SK37 Halobacillus halophilus DSM 2266 Alkaliphilus oremlandii OhILAs

50	60	70	80	90	100	110
RDE - MTLDEMREE	LOMSKPSMST	GVKKLODLN	VVKKTEHRGI	RKHTEVAEKDE	FKEETNEE	PPKWEREV
RDE - MTLDEMREE	LQMSKPSMST	GVKKLQDLN	VVKKTFHRGI	RKHTFVAEKDF	FKFFTNFF	P P KWE R E V
RDE-MTLDEMREE	LQMSKPSMST	GVKKLQDLN	VVKKTFHRGI	RKHTFVAEKD <mark>F</mark>	FKFFTNFF	P P KWE RE V
RDE - MTLDEMREE	LOMSKPSMST	GVKKLODLN	VVKKTFHRGI	RKHTFVAEKDF	FKFFTNFF	PPKWEREV
RDE MILDEMREE	LUMSKPSMSI	GARKFODEN	VVKKTEHRGI	RKHIFVAEKDE	ETEETNEE	PRIVUREV
RDE - MTLDEMREE	LOMSKPSMST	GVKKLODLN	VVKKTFHRGT	RKHTEVAEKDE	ETEETNEE	PPKWDREV
SDE - MTLDEMREE	LQMSKPSMST	SVKKLODLN	VVKKTFHRGR	RKHSFVAEKDF	FKFFTSFF	PKKWEREV
NEE - MTLDEMREE	LQMSKPSMST	GVKKLQDMN	IVKKTFHRGR	RKHSFVAEKD <mark>F</mark>	FKFFMNFF	PQ KWE RE V
NEE - MTLDEMREE	LOMSKPSMST	GVKKLODMN	IVKKTFHRGR	RKHSFVAEKDF	FKFFMNFF	POKWEREV
KDQ-MILDEMREE	LGMSKPSMST	SVRRLOEID	WVKKTFTRGS	RKHTYAAFKNE	ERSEMSEY	
EDQ - MNLDEMREK	LGMSKPSMST	NVRHLQEIG	WVKKKFQRGS	RKHTYTAEKNF	FHSFMSYF	COMWEREV
EDS-MTLDEMREQ	LQMSKPSMSA	GVKRLQEF <mark>D</mark>	IVKQQFTRGS	RKQHFIAEKDF	FNFFSNFF	TRKWRRE I
EDS-MTLDEMREQ	LQMSKPSMSA	GVKRLQEFD	IVKQQFTRGS	RKQHFIAEKDF	FNFFRNFF	RKWRREI
EGS-MILDEMRHQ	LOMSKPSMSA	GVKKLQEFD	IVKQQFIRGS	RKQHFIAEKDE	FIFFRNFF ETEEDNEE	TKKFQREI
EGS-MTLDEMRQQ		GVKKLQETD	IVKQQFTRGS	RKOHFIAEKDE	ETEERNEE	TKKWERET
EGS-MTLDEMRQQ	LQMSKPSMSA	GVKKLQEYD	IVKQQFTRGS	RKQHFVAEKDF	FTFFRNFF	TKKWEREI
EGS-MTLDEMREH	LQMSKPSMSA	GVKKLQEF <mark>D</mark>	IVKQQFTRGS	RKQHFIAEKNF	FHFFRNFF	f Q KWE RE V
EDN - MTL DQMREE	LOMSKPSMST	GVKKLQEYD	LVKQQFTRGS	RKQHFVAEKDF	FTFFRNFF	NIKFNREI
KES MTLDEMREH	LOMSKPSMSA	GVKKLOEFD	LVKQQFIRGS	RKQHFTAEKDF	FIFFRNFF	TO KNANDEL
KHEPITLDEMKDA	LGMSKPSMST	SVRKLOEIN	IVQKVWQKGS	RKDSFIAEKNF	FNYFTQFY	GMKWEREV
MDEPVTLDQMSEE	LGMSKPTMST	SIRSLONID	WVHKVWKKGV	RKDLYEAEKDF	SKSFFSFF	CKKWQREI
SDDPMNLNELRDE	LGMSKGSMSI	GVRKLLDEN	IIHRVYRKGE	RKDLYEAEQDF	FQFFTSFF	I RRWERER
NDGPMTLDDLRDE	LGMSKGSMSI	GVRKLLEEK	IIHRVYRKGE	RKDLYEAEKDF	FQFFISFF	
NRKPMILNELSEE	TGMSKTRMSG	VVREMLDVN	IAEKVFEKGV	RKDLYEVEODY	YOTFITLE	SANWSRVV SATWSKVV
NRKPMTLMELSEA	TGMSKTRMSC	VVREMLDAN	IAEKVFEKGV	RKDLYEVEQDY	YQTFITLF	SATWSKVV
NRKPMTLTELSEA	TGMSKTRMSC	VVREMLDAN	IAEKVFEKGV	RKDLYEVEQD <mark>Y</mark>	YQTFITLF	r a twsk v v
NRKPMTLTELSEA	TGMSKTRMSC	AVREMLDVN	LAEKVFEKGV	RKDLYDVEQDY	YQTFITLF	FATWSKVV
NRKPMILIELSEA	TGMSKIRMSC		LAEKVFEKGV	RKULYDVEQUY	YOTELTLE	TATWISKVV TATWISKVV
NRKPMTLSELSEA	TGMSKTRMSC	VVREMIDAN	LAEKVFERGV	RKDLYDVEQDY	YQTFISLF	TANWTKVV
NRKPMTL SELSEA	TGMSKTRMSC	VVREMIDAN	IAEKVFEKGV	RKDLYDVEQD <mark>Y</mark>	YQTFISLF	A A NWT K V V
NRKPMTL SELSEA	TGMSKTRMSC	VVREMIDAN	IAEKVFEKGV	RKDLYDVEQDY	YQTFISLF	AANWTKVV
NRKPMILSELSEA	TGMSKIRMSG	VVREMIDAN	I AEKVFEKGV L AEKVEEKGV	RKULYDVEQDY	YOTELSLE	A A NWIK V V Ta Martka V
NRKPMTLNELSEA	TGMSKTRMSC	VVREMTDAN	LAEKVFEKGV	RKDLYDVERDY	YQTFISLF	TANWTKAV
NRKPMTLNELSEA	TGMSKTRMSG	VVREMTDAN	LAEKVFEKGV	RKDLYDVERD <mark>Y</mark>	YQTFISLF	TANWTKTV
NRKPMTLNELSEE	TGMSKTRMSC	VVREMIDLN	IAEKVFEKGV	RKDLYDVEQDY	YQTFISLF	FANWSKVA
NRKPMILEELSEA	TGMSKIRMSG		IAEKVYEKGI	RKULYDVEQUQ	YOTELSLE	A ANWIRVV 6 6 MM/SRVV
NHRPMTLDELSQA	TGMSKTRMSC	VVREMVDMN	IAEKVFEKGV	RKDLYNVGLDF	YDTEVTIE	ISNLOKII
NRTPMTLDQLADE	TGMSKTRMSG	VMRQMMSLN	IAEKEFVKGS	RKEYYKVENDY	IQTFISLF	TTNWKEVV
NRQAMTLDELAAK	TGMSKTRMSC	VMRQMISLN	IAEKEFVKGS	RKEYYNVESDY	VQTFISLF	ITSWKEVI
QDKPMILDEMGEA	LGMSKISMSI MKMSKISMSI	GVRILLDLN	MVNKVWMKGI MVNKVWEKCS	RKDHYETEQDW	YQNFIDYF	SIKWRKGI
ADKPMNLDEMGQE	MKMSKTSMST	GVRTLLDNK	WVNKVWEKGS	RKDLYEVEYD	YOTEVDEE	DIKWRKSV
ADKPMTLDEMGRE	MEMSKTSMST	GVRTLLDLK	MVNKVWSKGS	RKDLYEVEYDW	YQTETDEE	AIKWRKAV
ADKPMTLDEMGRE	MEMSKTSMST	GVRTLLDL <mark>K</mark>	MVNKVWSKGS	RKDLYEVEYD	/HQTFTDFF	AIKWRKAV
SDHPMTLDDMKREI	MEMSKTSMST	GVRTLVDMK	MVHKVWEKGT	RKDLYEAEQD	YQTFTDYF	GLKWRKAV
HNGPVTLDELSKA	MGMSKTSMSI		HIDKVWGKGI	RKDOYEVVPD	YANESDEE	SIRWRRAV
QNDP I TL DQMSQQI	MGMSKTSMST	GMRTLMDLK	HIDKVWGKGS	RKDFYEVVPDW	YQNFTDFF	SIKWRKAV
SDGPVTLDDMREE	MGMSKTSMST	GMRTLMDL <mark>K</mark> I	MVNKVWGKGC	RKDLYETEFDW	/HQNEVDYE	SIKWRKSM
ENKPMTLDELKEE	LEMSKTSMST	GVRTLLELN	MVE KVWRKGE	RKDLYQVKGDW	YQNFIDRF	CTOWRKGT
ENKPMILDELKEE		GURSI VEAD	WVERVWRKGE	RKDLYQVKGDW	NKSESTVE	LIQWRKGT
AEEPMTLDEMSKA	MGMSKTSMST	GIRSLLDAN	WVERVWERGI	RKDLYKAEDDW	YKSFSNVF	INRWRNAT
EENALTLDEMSDH	LGMSKTSMST	SIRQLTEAN	MARKVWQKGV	RKDLYTSEDDW	YSSFIAIF	TKOWRTAI
ENDTLTLDEMSRR	LGMSKTSMST	GIRQLVDAN	MARKVWKKGV	RKDLYIGEEDW	MASFIAIF	SKKWRQGI
ADDPLTLDEMSEA	LGKSKTSMST	STRSL FDL N		RKDLYEANAQL	FKSFSTIY	ISKWLDST
ENNPLTLDDMSQA	LGMSKTSMST	GIRSLIDGE	MVEKTWKKG	RKDLYTAFESI	YKTESNIE	VHOWLSVV
				T :1	volv indu	or
DNA-	binding do	main	Flexible	linker	· · ·	

DNA-binding domain

binding-site

Supplementary Figure S3 | Amino acid sequence alignment of GbsR-type proteins associated with osmotic stress. The amino acid sequences of 27 GbsR-type proteins associated with glycine betaine synthesis (red), 18 OpuAR-type proteins (blue), 7 GbsR-type proteins associated with OpuB-type transporters (green) and 13 GbsR-type proteins associated with OpuC-type ABC transporters (yellow) were aligned using the MAFFT server (Katoh et al., 2017). Depicted are the N-terminal domains of these proteins. The amino acids corresponding to the winged helix-turn-helix DNA-binding motif (green), the inter-domain linker (reddish), and of the likely inducer-binding site (blue) are highlighted. Highly conserved amino acids are shaded in grey.



Supplementary Figure S4 | Structural arrangements of the OpuAC substrate-binding proteins. Shown are the *in silico* derived models of the OpuAC proteins of (A) *B. infantis* based on the crystal structure of (B) the *B. subtilis* OpuAC protein in complex with its ligand glycine betaine [PDB entry 2B4L (Horn et al., 2006)] and, (C) the modeled substrate-binding domain of the fused OpuABC protein of *P. larvae larvae* 08-100 DSM 25719 based on the crystallographic data on the (D) *L. lactis* OpuAC complexed with glycine betaine [PDB entry 3L6H (Wolters et al., 2010)]. Aromatic amino acids involved in the binding of glycine betaine are highlighted in red, whereas the ligand glycine betaine is depicted as blue sticks.



Supplementary Figure S5 Clade analysis of GbsR-type proteins among *Bacteria* and *Archaea*. 179 GbsR-type proteins were identified by bioinformatics. Information on fully sequenced microbial genomes were obtained from the IMG/M Web-server and homologs of the GbsR protein of *B. subtilis* JH642 (Nau-Wagner et al., 2012) were searched for via the JGI Web-server (Chen et al., 2017) using the BLAST-P algorithm (Altschul et al., 1990). The retrieved amino acid sequences were aligned using the MAFFT Web-server (Katoh et al., 2017) with bootstrapping setting (100 bootstraps) proved by the Web-server to analyze their phylogenetic relationship. The protein-homology tree was visualized using the iTOL web-tool (Letunic and Bork, 2016). GbsR homologs were grouped according to their gene neighborhood as depicted by colored boxes. If the *gbsR*-type gene was encoded in proximity of genes associated with glycine betaine synthesis or uptake systems for osmostress protectants it was classified as a member of the sub-group of GbsR-type proteins associated with osmotic stress (highlighted in red). Otherwise the GbsR homologs were assigned as non-associated with osmotic stress (highlighted in blue).

	50	60	70	80	90	100	110
CheP Racillus subtilis subtilis 168	RDE MTLDEMREE		GYKKLODI N	VVKKTEHRGI.	RKHTEVAEKDE	EKEETNE	F D D KWE R E V
Stanhylococcus sanronhyticus sanronhyticus ATCC 15305	EQKSMTLDEMRYE	LOMSKPSMSA	GVKRLOEFD	RVKOOFIRGS	RKOHFTAEKDF	FTFFGNF	FSOKRNREI
Staphylococcus suprophyticus s	EQKSMTLDEMRYE	LOMSKPSMSA	GVKKLQEFD	GVKQQFTRGS	RKQHFVAEKDF	FTFFGNF	FSOKWNREI
Paenibacillus mucilaginosus 3016	QNKPMTLDEMGRA	MEMSKTSMST	GVRTL VDL KI	MVHKVWEKGS	RKDLYEVEMDW	FQTFSDF	FSLKWRKAL
Terribacillus aidingensis MP602	ADQPMTLDDMRDA	LEMSKTSMST	GVRALAEMKI	MVETAYKKGM -	RKDLYKSEEDW	YKSFTSI	FGNKWEKST
Halobacillus halophilus DSM 2266	AEEPMTLDDMREA	LAMSKTSMST	GVRAL SDMK	MVEPSFKKGV	RKDLYKSEEDW	Y K S F T S L	FSSRWKHHT
Virgibacillus sp. SK37	SDNPMTLDDMRDA	LEMSKTSMST	GVRTL SDMK	MVESTFKKGI	RKDLYRSEEDW	YKSFTSL	FGNRWRQYT
Bacillus cellulosilyticus N-4 DSM 2522	SSDPMTLDDMRES	LAMSKTSMST	GVRTL VDMN	MVEPVFRRGV	RKDLYKAEEDW	/KSFTAL	FSRQWRKAS
Natranaerobius thermophilus JW/NM-WN-LF	SUEPKILDELRUM	MGMSKISMSI	GVRRLENNN		RAHLYEGEIDE	ANTELNE	FIPMWKRET
Natranaerobius thermophilus JW/NM-WN-LF	EDROMTI FEMKTSI	MNMSKSNMSY	AVRSI TESO	MITKAMQROD.		ABNELNE	EGAKLOREL
Bacillus clausii KSM-K16	NRKPMTLEAL SDE	TGMSKTRMSO	VVREMI SI N	VAHKVEEKGV.	RKDLYDVEKDY	YOTELSI	EVANWOAVI
Syntrophomonas wolfei Goettingen DSM 2245B	SEQPLGLEDIASE	LEVSKATVSI	HIRFLEGMK	NVRKVWVKGS	RRDYYEAERNT	GKIMTEH	LOSSFINER
Fimbriimonas gindengisoli Gsoil 348	TGVPLEVNEIMDR	LQISRGNASM	INL RELMDWG	VVRRFRQPGD-	RKDTYVSETDP	FLTLVKI	VKERKRRE I
Fimbriimonas gindengisoli Gsoil 348	TGVPLEVNEIMDR	LQISRGNASM	INL RELMDWG	VVRRFRQPGD-	RKDTYVSETDP	FLTLVKI	VKERKRRE I
Bradyrhizobium sp. BTAil	AEKPMTAEDIADT	LGMARSNVSN	ISLKELLAWD	L I RRVP I L GD -	RRDHFEAETD IN	NEVAQRI	AAGRKEREI
Bradyrhizobium sp. ORS278	AEKPMTAEDIAET	LGMARSNVSN	ISLKELLAWD	LIRRVPIMGD	RRDHFEAETDI	MEVAQRI	A A G R K E R E I
Bradyrhizobium oligotrophicum S58	AEKPMTAEDIADT	LGMARSNVSN	ISLKELLAWD		RRDHFEAETDI	MEVAORI	AAGRKEREI
Bradyrhizobium japonicum SEMIA 5079	AEAPMIAEDIADI	LGMARSNVSN	SL KELLAWN		RRDHYEAEIDI	WE VAARI	AARRKEREL
Methylobacterium radiotolerans JCM 2832 Mathylobacterium oblorom othenisum CM4	SERPLIAEDIAGE	LGIARSNVSN	SI KEL A CAM	LIKKVPVMGE	DDDEEVAETDI	MEMVIRI	AAGRKEREI
Methylobacterium chloromethanicum CM4 Methylobacterium extorguens AMI	SDKPLAAEDIAGT		SIKELAGAN		RRDFFVAFTDI	MEMVTRI	AAGRKEREL
Methylobacterium populi B 1001	SDKPLAAEDIAST	LGIARSNVSN	SLKELAAWN		RRDFFVAETDL	NEMVIRI	AAGRKEREI
Sphingomonas sanxanigenens DSM 19645	SERPLNADEIGEK	LGIARSNISN	SLKELVGWR	LVRRVPVMGD	RRDHFEAEVDL	NOMLTRI	AEGRKOREI
Octadecabacter temperatus SB1	SIKPVNAEEISDQ	LGIARSNVSN	SLKELVGWK	LIRRVPIQGD	RREHFVAEVD <mark>W</mark>	MEMAMRI	AQGRKERE I
Sideroxydans lithotrophicus ES-1	SHVPLNADDITYL	LSFSRSNVSM	IGLKELQSWR	LLRSEYRAGD -	RREYFKAPDD VI	ME I FRTL	ADERRRREV
Comamonas testosteroni TK102	SPEPL NADQ I ADT	LEFSRSNVSM	IGLKELQAWR	L VQL RHQPGD -	RREY FQAPED V	ME I F K R L	AEERRRRE I
Enhydrobacter aerosaccus SK60	AGKPMNAEEIQET	LGVARSNVSN	ISIKELQSLN	LVQTVHTLGD	RRDYFTTGTD	MELAKVI	VEERYRREL
Niastella koreensis GR20-10 DSM 17620	SPDPL SQDE I MEQ	LSISRGNVNM	INTROL I DWG		RKEFFSAEKDN	WKVATQI MKVATCI	VKERKKREL
Chitinophage pinensis UQM 2034 DSM 2588	ADOALSADDIMEE			LVDKVLIPGE	DKEEELAEKDM	MANNALET	ARERARKEL
Hallscomenobacter nyarossis O DSM 1100	SEADI SADDIMEE	LAISRGSSNM		LVYKRIKAGE	RREVEVAEKDM	MKVVROL	THREE
Spirosoma linguale DSM 74	SPEAL STEDVMEQ	LOISRGNASM	INL RDL MDWG	LVYKOLKPGE	RREFEVAEKDIN	MEVAROV	AKERBREL
Spirosoma radiotolerans DG54	SPDALAIEDIMEQ	LQISRGNASM	IN L R D L M DWG	LIYKQLKPGE	RREFFIAEKDI	NKVARQV	AKERRRREI
Emticicia oligotrophica GPTSA100-15 DSM 17448	SPKSMNADEIMAE	LQISRGNVNM	IN L R D L M DWG	LIYKOLLPGE	RKEYFIAEKDI	NKVAKQI	AKERRRRE I
Leadbetterella byssophila 4M15 DSM 17132	SEKALNADEIMEE	LQISRGNVNM	INL RDL I SWG	LIFKQLIPGE	RKEYFVAEKD IN	MKVT RQ I	SRERRREL
Owenweeksia hongkongensis DSM 17368	SKDAMTTEEIMDT	LKISRGNANM	IN I RAL I EWG	LVKREMVAGE -	RKEYFSAGKD <mark>W</mark>	ME I ARQ I	AKERAKREL
Hymenobacter sp. APR13	SPGALSTEDIMEQ	LQISRGNVNL	NVRALMDWG	I VRKEL RPGE	RREFFSGEKD	HRVATLI	LQERRKREL
Hymenobacter swuensis DY53	SPGALSTEDIMEQ	LQISRGNVNL	NVRALMDWG	I VRKEL RPGE	RREFFSGEKDI	ARVAILI	LQERRKREL
Flavobacterium johnsoniae UW101 ATCC 1/061	SNEPVSMEETMEET			IVINEFNAGE	- RREFFIAERULI		SRERSAREI
Flavobacterium columnare AICC 49512	SHEDVSMEDIMEE			IVYKEEKAGE	DDEEETAEKDI	DELAVRI	SPERSKREI
Dokdonia sp. 4H-3-7-5	SPEAL SMEDIMSE	LHISBGNASM	INLEST LDWG	LIFKEYKAGE	RREYEVAEGN	DELARKI	AKERSKRET
Dokdonia sp. PRO95	SPEAL SMED I MGE	LHISRGNASM	INL RSL I DWG	IIFKEYKAGE	RREYFVAESNI	DELARKI	AKERSKREI
Cellulophaga algicola IC166 DSM 14237	STKPLSTEEIMEE	LQISRGNTSM	INVRQL I DWG	IVTKELVPGE	RKEYFSTEKDV	QELARVI	AKERSRRE I
Lacinutrix sp. 5H-3-7-4	STKPLSMEEIMDE	LKISRGNTSM	INL RQL I DWG	IVSKTIIAGE -	RKEFFTTEKDV	QELARIV	AKERSRRE I
Cellulophaga lytica LIM-21 DSM 7489	SSEPL SMEE I MEE	LKISRGNTSM	INL RQL I DWG	I VSKEFKAGE -	. RKEYFTTEKD <mark>V</mark> (QELARII	AKERSRREL
Nonlabens dokdonensis DSW-6	HPDGL STDEIMEE'	VQLSRGNVNT	NVRELINWR	LVRKETVLGE	RKEFFIALHDVI	HSTAQNI	MEERKRREL
Optitutaceae sp. TAV5	SDRPLHIDEIMKE	LSISRGSANI	SLRELVNWG	VARSVIMKGE-	RKEHFESERDW	WRMFCAI	ARERKREI
Methanothermococcus okinawensis IHI	SDEDITISDIMEE	LGISKGNVSM	INL NALEALO		DKNYYTAVDGE	551 1 GM	KDIAKDKY
Methanocaldococcus sp. F5400-22 Mothanocaldococcus splagning M7 DSM 12004	SDKPLTISDYMEE	LKISKGNVSM	ISLKKLEELO		RKNYYEAVDGE	ISM	KDIARKKY
Methanocaldococcus iannaschii DSM 2661	SDKPLTISDIMEE	LKISKGNVSM	SLKKLEELG	FVRKVWIKGE	RKNYYEAVDGE	551	- KDIAKRKH
Methanocaldococcus fervens AG86	SNKPLTISDIMEE	LKISKGNVSM	ISLKKLEELG	FVKKAWIKGE	RKNYYIIVDGF	T S F	- KDVVKKKY
Methanotorris igneus Kol5 DSM 5666	SDKPLCIADIMDE	LGISKGNVSM	IALNKLEELG	FVRKVWIKGK	RRNYYEAMDG <mark>F</mark>	5 S Y	- KDIVKKKH
Methanococcus aeolicus Nankai-3	SHNPLCIDDIMEE	LGISRGNVSM	INLNKLEKLG	FIKKVWVKGE	. RKQYYEPLCG <mark>F</mark>	SSI	- L D I VGHKH
Methanocaldococcus infernus ME	YERPMSLDEIVEE	LKISKGNASM	ISLKKLEELG	FVKRTWIEGE	RKNFYQISESF	SSI	- KDIARRKH
Methanococcus voltae A3	SEKPLCMDETTDF	LKISKGTAST	TIRKLEELK	AIKKVWVEGD	RKNYYKISGSI	PLLDSMF	KRDM
Pyrococcus abyssi GES	ANDPESESDIALI	TGYSISHISS	AMEVLEGVG	LVORIKKPGD	- RRAY FVAI KNE	SEWRSSA	FYEKILRDI
Pyrococcus sp. NA2 Pyrococcus sp. ST04	AKEDISISEISSI	TGYSLSHVSS	AMKVLEGVG	LVORIKKPGD	RKAYFLATKNE	GEWRSSA	FYENILRDI
Pyrococcus furiosus DSM 3638	SEDPL SLGDISEL	TGYSI SHVSS	AMKVLESVG	VRRIKKPGD	RKVYEVATKTE	SEWRSSA	FYDNIMDNI
Pyrococcus yavanosii CH1	SDEPLSLAEIAEI	TGYSLSHVSS	AMKVLEGVG	LVORVKKPGD	RKAYEVATKS	GEWRSAA	FYSRLLRDV
Thermococcus onnurineus NA1	AKEPMSLSEIAER	TGYSLSHVST	ALKSMESLG	LVVRIKKPGD	KKAYYKATKLLI	K DWRQ A A	YYSRILEDI
Palaeococcus pacificus DY20341	EDEPLSLGEIAER	TGYSLSHVST	ALKLLESVG	LVKRVKKPGD	KKAYYTAIKN <mark>i</mark>	REWRRE A	YKRLEEDI
Thermococcus sibiricus MM 739	EDEPLSLGKIAER	TGYSLSHVST	ALKLLENIG	L V K R I K K P G D -	. KRAYYTA I KN <mark>i</mark> I	REWRKE A	YYKKIEEDV.
Thermococcus litoralis DSM 5473	EDEPLSLGEIAER	TGYSLSHVST	ALKLLENIG	LVKRIKKPGD	KRAYYTAIKNI	REWRKEA	YYKKIEEDV
Thermococcus paralvinellae ESI	ADEPLSLGEIAEK	TGYSISHVSS	ALKLLESVG	LVTRIKKPGD	KKAYFTAIKNL	KEWRKAA	YYNKLLEDV
Inermococcus barophilus MP DSM 11830	ADEPLSLGETAER	IGTSISHVSS	AVPOLEOLO	LVIKIKKPUD- MVPPSMOKCD	DOVECTVETO	ACTALSY	I CI PUKDEC
Halorhabdus utahensis 4X-2 DSM 12040	AEEPL SI DTI AAP	SEYAKSTVST	AMSDLORYH	MVTRRSI PGEO	GKKAFYFAFTDE	MOLERAE	LNNEVRREI
Haloferax mediterranei R-4	VGEPMSLDELVAF	SGYAKSTVSN	AMSALEPYH	LVRRQSSPGF	GRKVYFEAETDE	NT VMOE F	LEQOGRREI
Halogeometricum boringuense DSM 11551	AGGPKSLDELAEE	TGYAKSTVSN	AMGTLERYH	LVHRRSAPGE	GRKVYFEAETDF	NTVLTEL	FEREGRREI
Haloarcula hispanica CGMCC 1.2049	ASDPLSIPELVEE	TGYAKSTVSN	IVTRTL SRIG	LIHRRSSAGG	GRRVRFEAEREW	NFILQDV	FQQYIQREV
Haloarcula hispanica N601	ASDPL SIPELVEE	TGYAKSTVSN	IVTRTL SRIG	LIHRRSSAGGO	GRRVRFE A E RE <mark>M</mark>	NFILQDV	FQQY I QRE V
Haloarcula marismortui ATCC 43049	AAEPLSIPELVDE	TGYAKSTVSN	VTRTLSRVG	LIHRKSSAGGO	GRRVRFEAERD <mark>W</mark>	NFILQDV	FQQYIQREV
Salinarchaeum sp. Harcht-Bskl	AEEPLSLDVLAER	SGYAKSTVSD	VTNSLEEIY	FARRVSGAGGG	GRKSYFEAERD	NY AMRQ A	MQDAGRREV
Hallangium ochraceum SMP-2 DSM 14365	APEPESAAETAQR	LSLSSGAVSM	ILLGELTHWG		REDHYESESNI		RERELRKI
Corauomargarita akajimensis DSM 45221 Mathanothrir sochnachij GP 6	ACRELATOULVER	TGYSKETVEL	NMNOL KNL G	VKRVVIDCD	KRHLYADITON	ARLASUP ETTKINN	I DALTKEY
Rhodothermus marinus R-10 DSM 4252	OPEPL SI DE LATI	LNRSKGPISS	TIRFLASIG		RRDYYVAHPDI	FLNNEKE	NMATVRKNR
nononormus murmus n-10 DBM 4232	a. e. cococinite		THE PHOTO I				

Supplementary Figure S6 | Amino acid sequence alignment of GbsR-type proteins nonassociated with osmotic stress. The amino acid sequences of 7 GbsR homologs associated with various transport systems (pink) and 73 GbsR-type proteins, encoded in a gene neighborhood which does not allow the assignment of a particular function (pale purple) were aligned using the MAFFT server (Katoh et al., 2017). Depicted are only the N-terminal domains of these proteins. The amino acids corresponding to the winged helix-turn-helix DNA-binding motif (lite blue), the inter-domain linker (purple) and of the likely inducerbinding site (reddish) of GbsR are highlighted. Highly conserved amino acids are shaded in grey.



Supplementary Figure S7 *B. infantis* cannot use glycine betaine as carbon source. Cultures of *B. infantis* NRRL B-14911were grown in basal medium supplemented with either glucose (27.7 mM final concentration) or glycine betaine (33.24 mM final concentration) as the sole carbon source. Growth yields were measured after 46 hours of incubation at 37 °C. Cultures without the addition of a carbon source served as control.

Strain	Relevant genotype	Source or reference
JH642	trpC2 pheA1	J. Hoch; BGSC [*] 1A96
TMB118	$\Delta(opuA::tet) \Im \Delta(opuC::spc) \Im \Delta(opuD::neo) \Im \Delta(opuB::erm) \Im$	(Teichmann et al., 2017)
GNB37	Δ (treA::erm)2	(Nau-Wagner et al., 2012)
GNB40	$\Delta(gbsR::neo) 1 \Delta(treA::erm) 2$	(Nau-Wagner et al., 2012)
CAB1	$\Delta(opuA::tet) \Im \Delta(opuC::spc) \Im \Delta(opuD::neo) \Im \Delta(opuB::erm) \Im amyE::pX$	This study
CAB2	$\Delta(opuA::tet)$ $\Delta(opuC::spc)$ $\Delta(opuD::neo)$ 2 $\Delta(opuB::erm)$ 3 $amyE::opuAR-opuA_{B.i}$	This study
STHB04	Δ (treA::erm)2 Δ (gbsR::neo)1 Δ (yvaV::tet)2	This study
STHB05	Δ (treA::erm)2 Δ (gbsR::neo)1 Δ (yvaV::tet)2 Δ (opcR::zeo)2	This study
STHB06	Δ (treA::erm)2 Δ (gbsR::neo)1 Δ (opcR::zeo)2	This study
STHB07	$\Delta(treA::erm)2 \Delta(yvaV::tet)2 \Delta(opcR::zeo)2$	This study
STHB08	$\Delta(treA::erm)2 \Delta(opcR::zeo)2$	This study
STHB10	$\Delta(treA::erm)2 \Delta(gbsR::neo)1 \Delta(yvaV::tet)2 \\\Delta(opcR::zeo)2 [amyE::\Phi(opuAA_{B,i}'-treA); \\opuAR^+]1$	This study
STHB11	Δ (treA::erm)2 Δ (gbsR::neo)1 Δ (yvaV::tet)2 Δ (opcR::zeo)2 [amyE:: Φ (opuAA'-treA)]1	This study
STHB65	Δ (treA::erm)2 Δ (gbsR::neo)1 Δ (yvaV::tet)2 [amyE:: Φ (opuAA' _{B.i} -treA)]1	This study
STHB66	Δ (treA::erm)2 Δ (gbsR::neo)1 Δ (opcR::zeo)2 [amyE:: Φ (opuAA' _{B.i} -treA)]1	This study
STHB67	$\Delta(treA::erm)2 \Delta(yvaV::tet)2 \Delta(opcR::zeo)2 [amyE::\Phi(opuAA'_{B.i}-treA)]1$	This study

Supplementary Table S1 | *Bacillus subtilis* strains used in this study

*BGSC: Bacillus Genetic Stock Center (Columbus, OH, USA).

Supplementary Table S2 | Primers used in this study

During an array of	$\mathbf{D}_{\mathbf{n}}$
Primer name	
Oper Knout PI	
Oper Knout Zeo P2	
	CGTATGCATATTTT
OpcR Knout Zeo P3	CCATATCAAGATAACTTCGTATAATGTATGTTGAAGGCATTCCAAA
	CGTATGCATATTTT
OpcR Knout P4	GTAAAGCAATACTCGTCTGCTTTTGTTTTA
OpcR Knout Zeo P5	AAAATATGCATACGTTTGGAATGCCTTCAACATACATTATACGAAG
.r	TTATCTTGATATGG
OncR Knout Zeo P6	TATTTTAGAGAGCTGCATTCTTTTGTTTTCTAATGTATGCTATACGA
open miour Zeo I o	
VeroV Verout D1	
YVAV KNOULPI Vyvay V a gyt Tat D2	
Yvav Knout Tet P2	
	TTTCTATAAAATG
YvaV Kneu2 Tet P3	AAATTGTTATCCGCTCACAATTCCACACAACATATTTGAGAGCGA
	AGACATTTTTAAATATGTG
YvaV Knout P4	CAGTGAAATAAACCGGTAAATCTAGGTCTC
YvaV Knout Tet P5	CATTTTATAGAAAGAATCGCGGAAAACATGGAAAAGAGTAATGC
	TAACATAGCATTACGG
YvaV Kneu2 Tet P6	CACATATTTAAAAATGTCTTCGCTCTCAAATATGTTGTGTGGAATT
1,4, 111042 10010	GTGAGCGGATAACAATTT
Opu A P frud	
OpuAR Iwa	
OpuAR lev OpuAR Poy1 for	
OpuAK_Box1_lol	
OnuAR Box1 rev	CGTACAAAAGCGATCGATGAAATTCTGGAACCAATCCCCTTTGAC
opunit_box1_iev	CTGG
OpuAR Box3 for	GGTACCAGAATTTCATCGATTTCTTTTGTACGCAATGGAGAAAAG
opunt_bons_for	G
OpuAR Box3 rev	CCTTTTCTCCATTGCGTACAAAAGAAATCGATGAAATTCTGGTAC
	C
OpuAR Box1+3 for	CCAGGTCAAAGGGGATTGGTTCCAGAATTTCATCGATTTCTTTG
1	TACG
OpuAR Box1+3 rev	CGTACAAAAGAAATCGATGAAATTCTGGAACCAATCCCCTTTGA
·	CCTGG
OpuAR N98Fn for	GTACAAAAGCGATCGATGAAAAACTGGTACCAATCCCCTTTGAC
OpuAR N98Fn rev	TCAAAGGGGATTGGTACCAGTTTTTCATCGATCGCTTTTGTAC
OpuAR W95F for	GATTTATACCAGGTCAAAGGGGATTTCTACCAGAATTTCATCGAT
	CG
OpuAR W95F rev	CGATCGATGAAATTCTGGTAGAAATCCCCTTTGACCTGGTATAAA
	ТС
CA3-opuARA fwd	AAAGGATCCGCATCTACTCTTCCTTTTTTTC
CA3-opuARA rev	AAAGGATCCCATTAATCTAAGTCTTCGATTG
OpuAR treA Frag1/4 for	AAACCCGGGGCACCTTTTTGGAGAGCATCTAC
OpuAR treA Frag2 for	AAACCCGGGGAACTCCGTATAAATGCATGTTTC
OpuAR treA Frag1/2 rev	AAAGGATCCCTGAAGCAGCTTCGATTGTGAC

Plasmid	Description	Resistance	Reference	
pASK-IBA3plus	Expression plasmid for <i>E. coli</i> with a AHT-inducible <i>tet</i> -promoter and a C-terminal <i>Strep</i> -tag II	bla	IBA (Göttingen, Germany)	
рХ	Integration vector for <i>B. subtilis amyE::cat::amyE</i>	bla, cat	(Kim et al., 1996)	
pJMB1	amyE::treA	bla, cat	(Hoffmann et al., 2013)	
pDG1515	Tetracycline resistance cassette	tet	(Guerout-Fleury et al., 1995)	
p7Z6	Zeocin resistance cassette	<i>Ze0</i>	(Yan et al., 2008)	
pCA-opuARA	<i>opuA</i> operon and <i>opuAR</i> gene of <i>B</i> . <i>infantis</i> with native promoters cloned into pX- <i>amyE</i> -site	bla, cat	This study	
pSTH33	$amyE::\Phi(opuAA_{B.i}'-treA); opuAR^+$	bla, cat	This study	
pSTH34	$amyE::\Phi(opuAA_{B,i}'-treA)$	bla, cat	This study	
pMP_AR1	<i>B. infantis opuAR</i> gene cloned into pASK-IBA3plus	bla	This study	
pSTH49	Site directed mutagenesis of <i>B. infantis</i> opuAR in pMP_AR1: opuAR [W ⁹³ F]	bla	This study	
pSTH50	Site directed mutagenesis of <i>B. infantis</i> opuAR in pMP_AR1: opuAR [N ⁹⁶ F]	bla	This study	
pOpuAR_Y94F	Site directed mutagenesis of <i>B. infantis</i> opuAR in pMP_AR1: opuAR [Y ⁹⁴ F]	bla	This study	
pOpuAR_R100F	Site directed mutagenesis of <i>opuAR</i> in pMP_AR1: <i>B. infantis opuAR</i> [R ¹⁰⁰ F]	bla	This study	
pOpuAR_Y94F_R100F	Site directed mutagenesis of <i>B. infantis</i> opuAR in pMP_AR1: opuAR [Y ⁹⁴ F/R ¹⁰⁰ F]	bla	This study	

Supplementary Table S3 | Plasmids used in this study

REFERENCES

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. 215, 403-410.
- Chen, I.A., Markowitz, V.M., Chu, K., Palaniappan, K., Szeto, E., Pillay, M., et al. (2017). IMG/M: integrated genome and metagenome comparative data analysis system. *Nucleic Acids Res.* 45, D507-D516. doi: 10.1093/nar/gkw929.
- Guerout-Fleury, A.M., Shazand, K., Frandsen, N., and Stragier, P. (1995). Antibiotic-resistance cassettes for *Bacillus subtilis. Gene* 167, 335-336.
- Hoffmann, T., Wensing, A., Brosius, M., Steil, L., Völker, U., and Bremer, E. (2013). Osmotic control of opuA expression in *Bacillus subtilis* and its modulation in response to intracellular glycine betaine and proline pools. J. Bacteriol. 195, 510-522. doi: 10.1128/JB.01505-12.
- Horn, C., Sohn-Bösser, L., Breed, J., Welte, W., Schmitt, L., and Bremer, E. (2006). Molecular determinants for substrate specificity of the ligand-binding protein OpuAC from *Bacillus subtilis* for the compatible solutes glycine betaine and proline betaine. *J. Mol. Biol.* 357, 592-606. doi: 10.1016/j.jmb.2005.12.085.
- Katoh, K., Rozewicki, J., and Yamada, K.D. (2017). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform*, 1-7. doi: 10.1093/bib/bbx108.
- Kim, L., Mogk, A., and Schumann, W. (1996). A xylose-inducible *Bacillus subtilis* integration vector and its application. *Gene* 181, 71-76.
- Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242-W245. doi: 10.1093/nar/gkw290.
- Nau-Wagner, G., Opper, D., Rolbetzki, A., Boch, J., Kempf, B., Hoffmann, T., et al. (2012). Genetic control of osmoadaptive glycine betaine synthesis in *Bacillus subtilis* through the cholinesensing and glycine betaine-responsive GbsR repressor. *J. Bacteriol.* 194, 2703-2714. doi: JB.06642-11 10.1128/JB.06642-11.
- Teichmann, L., Chen, C., Hoffmann, T., Smits, S.H.J., Schmitt, L., and Bremer, E. (2017). From substrate specificity to promiscuity: hybrid ABC transporters for osmoprotectants. *Mol. Microbiol.* 104, 761-780. doi: 10.1111/mmi.13660.
- Wolters, J.C., Berntsson, R.P., Gul, N., Karasawa, A., Thunnissen, A.M., Slotboom, D.J., et al. (2010).
 Ligand binding and crystal structures of the substrate-binding domain of the ABC transporter
 OpuA. *PloS one* 5, e10361. doi: 10.1371/journal.pone.0010361.

- Xia, X., Wu, S., Li, L., Xu, B., and Wang, G. (2018). The cytochrome bd complex Is essential for chromate and sulfide resistance and is regulated by a GbsR-type regulator, CydE, in Alishewanella sp. WH16-1. Front. Microbiol. 9, 1849. doi: 10.3389/fmicb.2018.01849.
- Yan, X., Yu, H.J., Hong, Q., and Li, S.P. (2008). Cre/lox system and PCR-based genome engineering in Bacillus subtilis. Appl. Environ. Microbiol. 74, 5556-5562. doi: AEM.01156-08 10.1128/AEM.01156-08.