Supplementary materials

The role of the *ganSPQAB* operon in degradation of galactan by *Bacillus subtilis*

Authors: Hildegard Watzlawick, Kambiz Morabbi Heravi[#], Josef Altenbuchner

Affiliation:

Institut für Industrielle Genetik, Universität Stuttgart, Allmandring 31, 70569 Stuttgart, Germany

Running title: The galactan utilization system of *B. subtilis*

Keywords: galacto-oligosaccharide, galactanase, β-galactosidase, sugar-binding lipoprotein, ABC transporter, transcription regulation

[#]**Corresponding author:** Kambiz Morabbi Heravi

Email: kambiz.morabbi@iig.uni-stuttgart.de

Tel.: +49 711 68566982

Fax: +49 711 68566973

Materials and Methods

DNA manipulation and plasmid construction

Standard molecular techniques, such as *E. coli* transformation, were carried out as described by Sambrook and Russell (1). Natural transformation of *B. subtilis* was performed as described before (2). The desired DNA fragments were amplified in PCRs utilizing Phusion® HF DNA polymerase (Catalog #M530S; New England BioLabs®, Frankfurt am Main, Germany) on a PTC-200 Peltier Thermal Cycler (MJ Research Inc., Watertown, Massachusetts, USA). Unless otherwise specified, the *B. subtilis* 168 chromosomal DNA was added to all PCRs as the template. Oligonucleotides used in this study are listed in Table S2. Commercial kits including DNeasy® Blood & Tissue Kit (Cat. #69506; Qiagen, Hilden, Germany) for chromosomal DNA extraction, "innuPREP Plasmid mini Kit" (Analytic Jena AG, Jena, Germany) for plasmid isolation and "NucleoSpin® Gel and PCR Clean-up Kit" (Machery-Nagel, Düren, Germany) for PCR fragment purification were applied through-out this study. Restriction enzymes were provided by New England BioLabs® (Frankfurt am Main, Germany). T4 DNA Ligase was purchased from Thermo Fisher Scientific Inc. (Karlsruhe, Germany). The constructed plasmids were sequenced by GATC Biotech (Konstanz, Germany).

All plasmids used in this study are listed in Table S1. To characterize the GanA and GanB enzymes, they were produced as N-terminal His_6-tagged fusion proteins in *E.coli*. Plasmid pJOE5751.1 harboring an L-rhamnose-inducible promoter (*rha*P_{BAD}) was used as the expression vector for this purpose. The *ganB* gene without the signal sequence was amplified in a PCR with oligonucleotides s10013 and s10014, while oligonucleotides s9872 and s9873 were added to PCR for amplification of the *ganA* gene. The amplified *ganB* and *ganA* DNA fragments were then inserted into pJOE5751.1 via BamHI and BsrGI restriction sites to create plasmids pHWG1119

and pHWG1111, respectively. The *cycB* gene lacking its signal peptide encoding sequence was amplified in a PCR with oligonucleotides s10505 and s10506. The PCR product was then inserted into pJOE5751.1 through BamHI and BsrGI restriction sites to construct pHWG1151. To produce and purify the GanR-Strep tagged (C-terminus) fusion protein, its encoding gene was amplified in a PCR using oligonucleotides s10275 and s10277. The amplified GanR DNA-fragment was inserted between the BamHI and NdeI restriction sites of the rhamnose-inducible vector pJOE6089.4 in order to construct pJOE8798.6.

The *ganR*-P_{gan} DNA fragment containing the *ganR* gene and the promoter of the *gan* operon located downstream of *ganR* was amplified in a PCR with oligonucleotides s10175 and s10176. This fragment was inserted between the AgeI and NdeI site upstream of the promoterless superfold-GFP gene of a derivative of pUC18/pUB110 shuttle vector (unpublished) and created plasmid pJOE8774.1. The plasmid pJOE8790.6 is derived from pJOE8774.1 by cutting pJOE8774.1 with AgeI and NruI and circularization of the vector to remove *ganR* from the vector.

Protein purification

The cell pellet of *E. coli* over-expressing the desired protein was broken by passing them twice through a French press cell. After centrifugation of the cell lysate at 12,000 × g for 15 min, the fraction of the crude extract proteins in the supernatant was used for purification. Purification of the His_6-tagged proteins was performed by immobilized metal affinity chromatography (IMAC) using TALON® Metal Affinity resins (Clontech Laboratories, Inc, Mountain View, USA) according to the supplier's instruction. Fractions of the purified proteins were combined and the imidazole was removed with illustra NAP-10 columns (GE Healthcare, Munich, Germany). Usually, the yield of the purified protein was between 0.6 to 3 mg from a 50 ml

induced *E. coli* culture. Purification of the Strep-tagged fusion proteins was carried out using *Strep*-Tactin® matrices according to manufacturer's instructions (IBA GmbH, Göttingen, Germany). Protein concentration was determined by the method developed by Bradford (3) with bovine serum albumin as a standard.

DNA sequencing

For DNase I footprinting and primer extension experiments, the desired plasmids were sequenced with relevant oligonucleotides using Thermo Sequenase™ Cycle Sequencing Kit (Affymetrix, High Wycombe, UK). The sequencing master mix was prepared by the addition of 2 µl of the desired plasmid (30 fmol/µl) to 2 µl reaction buffer, 1 µl oligonucleotide (100 pmol/µl), 1 µl DMSO, 2 µl DNA polymerase and 9.5 µl ddH_2O. 4 µl of the sequencing master mix was then added to 4 µl aliquots of ddATP, ddCTP, ddGTP and ddTTP. The sequencing reaction was accomplished in a PCR using a PTC-200 Peltier Thermal Cycler (MJ Research Inc., Watertown, Massachusetts, USA). The PCR program was: 95 °C for 2 min, 30 cycles of (95 °C for 30 sec, (T_m-5) °C for 30 sec, 72 °C for 1 min) and the final extension for 1 min at 72 °C. Finally, 4 µl of stop solution was added to each reaction and 5 µl was loaded onto a polyacrylamide gel. The sequencing reaction was analysed by ALFexpress DNA sequencer (GE Healthcare, Munich, Germany).

Electrophoretic mobility shift assay

A 5'-end FITC-labeled DNA fragment was generated in a PCR using FITC-labeled oligonucleotides. FITC-labeled P_{gan} DNA fragments were amplified with oligonucleotides s10311 and s10312 from genomic DNA of *B. subtilis* as the template. Electrophoretic mobility shift assay (EMSA) was carried out in a total volume of 10 µl containing 2 µl FITC-labeled P_{gan} DNA fragment (88 fmol/µl), 2 µl of 5× shift buffer (50 mM Tris-HCl, pH 7.5; 250 mM KCl; 10

mM dithiothreitol [DTT]; 25% [vol/vol] glycerol; 250 µg/ml BSA; 25 µg/ml Herring sperm DNA) and 1 µl purified GanR-Strep tag (31 pmol/µl). 5 µl of the inducer candidates (10 mM) were added at a final concentration of 5 mM. The reaction mixture was incubated at 8 °C for at least 15 min, and all of the reaction mixture was loaded onto a 6% (wt/vol) native polyacrylamide gel. The gel was run at 20 mA for 40 min to separate the free DNA and DNA-protein complexes. The migration of the bands of free DNA and the DNA-protein complexes was visualized by a PhosphorImager (Storm 860 PhosphorImager; Molecular Dynamics).

DNase I footprinting

Determination of the GanR DNA binding site was carried out using DNase I DNA footprinting. Cy5-labelled DNA fragments were generated in PCRs using Cy5-labelled 5'-end oligonucleotides s10512 and s10725. To label the non-coding strand, the P_{gan} DNA was amplified from pJOE8790.6 using oligonucleotides s10512 and s10523 (PCR254), while for the coding strand, the P_{gan} DNA was amplified from pJOE8774.1 using oligonucleotides s10725 and s10726 (PCR261). 6 µl of the Cy5-labelled DNA (100 fmol/µl; PCR254 or PCR261) was then mixed with 18 µl of the purified GanR-Strep tag (1.2 mg/ml) and 12 µl of the 5× shift buffer (20 mM HEPES pH 7.4, 80 mM KCl, 1 mM TCEP, 8 mM $MgCl_2$, 25% glycerol, 0.25 mg/ml BSA, 0.25 mg/ml Herring sperm DNA) in a final reaction volume of 60 µl. The strep tag elution buffer (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 2.5 mM desthiobiotin) was added to the control reaction instead of purified GanR. All reactions were incubated for 15 min on ice followed by incubation for 5 min at 25 °C. The DNase I digestion was carried out by the addition of 0.4 µl of DNase I (2,000 U/ml; NEB, Frankfurt, Germany), 7 µl of the 10× DNase I buffer and 2.6 µl ddH_2O. Each reaction was stopped after 1 min by the addition of 70 µl stop buffer (50 mM EDTA pH8.0, 15 µg/ml calf thymus DNA) and 140 µl phenol:chloroform:isoamylalcohol (25:24:1). After 10 min

of centrifugation, 100 µl of the upper phase was mixed with 200 µl of ethanol and kept at -70 °C for 30 min. The precipitated DNA was washed once with ethanol and dissolved in 8 µl ddH_2O and 4 µl stop solution of the Thermo Sequenase™ Cycle Sequencing Kit (Affymetrix, High Wycombe, UK). 5 µl of each reaction was loaded on to a polyacrylamide gel which was then analyzed by ALFexpress DNA sequencer (GE Healthcare, Munich, Germany). To identify the GanR binding site, DNA sequencing reactions were carried out using pJOE8790.6 as a template and oligonucleotide s10512 for sequencing of the P_{gan} non-coding strand, while coding strand was sequenced using pJOE8774.1 and oligonucleotide s10725.

Primer extension

To identify the transcription start site (TSS) of the *cycB-ganPQAB* operon, *B. subtilis* KM0 was transformed with pJOE8790.6 expressing the *sGFP* by $P_{gan.}$ Next, a single colony of KM0 pJOE8790.6 was cultivated in 5 ml LB supplemented with kanamycin and incubated overnight at 37 °C. Afterwards, the bacterial culture was harvested and the total RNA was extracted from the cell pellet with Qiagen RNeasy mini kit (Hilden, Germany) according to the manufacture's instruction. Approximately, 30 µg of the total RNA was precipitated using sodium acetate (3 M, pH 6.3) and ethanol. The RNA pellet was then dissolved in 5 µl RNase-free ddH_2O. 0.5 µl RNasin® Ribonuclease Inhibitor (40 U/µl; Promega, Mannheim, Germany) was added to the RNA. The mixture was then incubated for 3 min at 65 °C followed by the addition of 0.5 µl oligonucleotide s10512 (10 pmol/µl) and 2 µl of the 5× Avian myeloblastosis virus reverse transcriptase (AMV-RT) reaction buffer for hybridization. The reaction mixture was incubated for 20 min at 55 °C followed by incubation for 5 min at RT. Reverse transcription reaction was carried out by the addition of 1 µl dNTP (10 mM), 1 µl AMV-RT (20 U/µl) and incubated for 1 h at 42 °C. The reverse transcription reaction was stopped by the stop solution provided by Thermo

Sequenase™ Cycle Sequencing Kit (Affymetrix, High Wycombe, UK). 5 µl of the generated Cy5-labeled cDNA was loaded onto a polyacrylamide gel and analysed by ALFexpress DNA sequencer (GE Healthcare, Munich, Germany). To find the TSS, a DNA sequencing reaction was carried out as described before using pJOE8790.6 as template and oligonucleotide s10512 as primer.

Thermal shift assay

Thermal shift assay (also known as differential scanning fluorimetry) was carried out to find the possible ligands of GanR and CycB. The experiments were conducted with a Mastercycler® ep realplex (Eppendorf, Hamburg, Germany) utilizing the melting capability. Reactions were carried out in a total volume of 50 µl by mixing purified GanR-Strep tag (12 µg), purified His_6-CycB (20 µg) with(out) the effectors, galactobiose, galactotriose and galactotetraose at a final concentration of 1 mM, with SYPRO® Orange (5 µl of the 50× stock solution in DMSO; Sigma, Munich, Germany). The fluorescence intensity profile (or melting curve) of the SYPRO® Orange-protein complex was measured at 520 nm with intervals of 0.2°C from 20°C to 90°C. The thermal stability (melting temperature) was recorded from the Δfluorescence/ΔT.

TABLE S1 Strains and plasmids used in this study

Strain or plasmid	Genotype or description	Source or reference
E. coli		
JM109	*recA1*, *endA1*, *gyrA96*, *thi*-1, *hsdR17*(r_K^-, m_k^+), *mcrA*, *supE44*, *gyrA96*, *relA1,* λ^-, Δ(*lac-proAB*), F' (*traD36*, *proAB*$^+$, *lacI*^{q}, (Δ*lacZ*)M15)	(4)
B. subtilis		
KM0	a tryptophan prototroph derivative of strain 168	(5)
BKE32550	*trpC*2 Δ*yurJ*::*loxP-ermC-loxP*	BGSC[a]
BKE34120	*trpC*2 Δ*ganB*::*loxP-ermC-loxP*	BGSC
BKE34130	*trpC*2 Δ*ganA*::*loxP-ermC-loxP*	BGSC
BKE34160	*trpC*2 Δ*cycB*::*loxP-ermC-loxP*	BGSC
BKE34170	*trpC*2 Δ*ganR*::*loxP-ermC-loxP*	BGSC
BKE38810	*trpC*2 Δ*msmX*::*loxP-ermC-loxP*	BGSC
KM462	Δ*cycB*::*loxP-ermC-loxP*	pKAM041→BKE34160
KM468	Δ*cycB*::*loxP*	pJOE6732.1→KM462
KM580	*trpC*2 Δ*ganB*::*loxP*	pJOE6732.1→BKE34120
KM581	*trpC*2 Δ*ganA*::*loxP*	pJOE6732.1→BKE34130
KM588	Δ*ganB*::*loxP*	pKAM041→KM580
KM589	Δ*ganA*::*loxP*	pKAM041→KM581
Plasmids		
pKAM041	ori_{pUC18}, *bla*, *ter*-'*trpD-trpC-trpF*'-*ter*	(5)
pJOE5751.1	ori_{pBR322}, *rop*, *bla*, *rha*P_{BAD}-His_6-*eGFP*-ter_{rrnB}	(6)
pJOE6089.4	ori_{pBR322}, *rop*, *bla*, *rha*P_{BAD}-*eGFP*-Strep-ter_{rrnB}	(6)
pJOE6732.1	$ori_{pAMβ1}$, *repDE*, P_{xyl}-cre_{P1}, *spc*, ori_{pUC18}, *bla*	Altenbuchner (unpublished)
pJOE8774.1	ori_{pUC18}, ori^+_{pUB110}, *kan*, rep_{pUB110}, *kan*, *ter-ganR*-P_{gan}-His_{6}-*sGFP-ter*, *bla*	*ganR*-P_{gan}→pUB110/pUC18 derivative (unpublished)
pJOE8790.6	ori_{pUC18}, ori^+_{pUB110}, *kan*, rep_{pUB110}, *kan*, *ter-'ganR*-P_{gan}- His_{6}-*sGFP-ter*, *bla*	pJOE8774.1 Δ*ganR*
pJOE8798.6	ori_{pBR322}, *rop*, *bla*, *rha*P_{BAD}-*ganR*-Strep-ter_{rrnB}	*ganR*→pJOE6089.4
pHWG1111	ori_{pBR322}, *rop*, *bla*, *rha*P_{BAD}-His_6-*ganA*-ter_{rrnB}	*ganA*→pJOE5751.1
pHWG1119	ori_{pBR322}, *rop*, *bla*, *rha*P_{BAD}-His_6-*ganB*-ter_{rrnB}	*ganB*→pJOE5751.1
pHWG1151	ori_{pBR322}, *rop*, *bla*, *rha*P_{BAD}-His_6-*cycB*-ter_{rrnB}	*cycB*→pJOE5751.1

[a] *Bacillus* Genetic Stock Center

TABLE S2 Oligonucleotides used in this study

Primer	Sequence (5′→3′)	Application
s 9872	GCG GAT CCT CAA AGC TTG AAA AAA CGC ACG	*ganA* amplification
s 9873	GCC TGT ACA CTA ATG TGT GTT TAC GAC AAT TC	*ganA* amplification
s10013	GCG GAT CCA TTG AGA AGG AGA AGC ACG TGT C	*ganB* amplification
s10014	GCC TGT ACA TCA GCG TTT CTT TGG AAT ATG TC	*ganB* amplification
s10175	AAA AAA CAT ATG TTG TCA CCC TTT CGG TAG GA	*ganR*-P_{gan} amplification
s10176	AAA AAA CCG GTT CCG AAT GAT ACG AA	*ganR*-P_{gan} amplification
s10275	AAA AAA GGA TCC ATT CAT ACT CTT CCT GAC GA	*ganR* amplification
s10277	AAA AAA CAT ATG GCG ACA ATT AAA GAT ATC G	*ganR* amplification
s10505	GCG GAT CCA AAT CGA GTT CAA AAG GGT CAG	*cycB* amplification
s10506	GCC TGT ACA TTA TTT TCC GCT GTG CTT TGC C	*cycB* amplification
s10523	TCC TCT GAC GAC GTT TCA	P_{gan} amplification
s10726	TGA TGA TGG TTA TTG TCT GC	P_{gan} amplification
5'-end labelled		
s10311	FITC-GCG CAG AGC ATG AAT ACG	EMSA
s10312	FITC-GCT TGA ACA AGT GCA GGA C	EMSA
s10512	Cy5-CGT TTA CGT CGC CGT CCA	DNA footprinting
s10725	Cy5-AGC GGA CCC GAT TGC AGT	DNA footprinting

TABLE S3 The GanB activity assays using different type of substrates and methods.

Substrate	**Activity**	**Method**
Arabino-galactan	negative	TLC, HPLC
*p*NP-βGal	negative	Colorimetric
AZCL-galactan	positive	Colorimetric
Galactan (Lupin)	positive	TLC, HPLC, DNS-Test

TABLE S4 The GanA activity assays using different type of substrates and methods.

Substrat	**Activity**	**Method**
*p*NP-βGal	80 U/mg	Colorimetric
Lactose	positive	HPLC
Arabino-galactan	negative	TLC
AZCL-galactan	negative	Colorimetric
Galactan (Lupin)	negative	HPLC
$(Gal)_4$, $(Gal)_3$, $(Gal)_2$	positive	HPLC

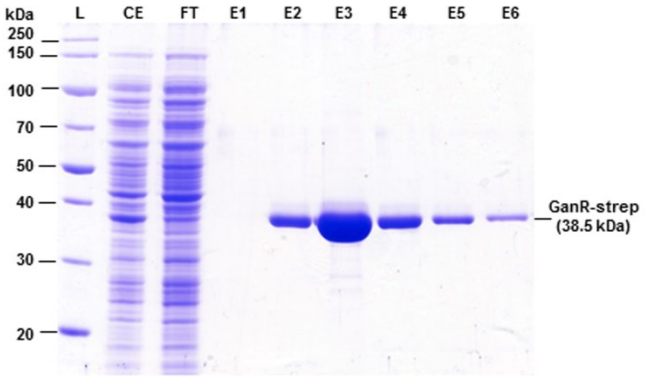

FIG S1 Purification of GanR-Strep tag. The crude extract of the induced strain JM109 pJOE8798.6 was used for purification of the Strep-tagged GanR as explained in methods and materials.

References

1. **Sambrook J, Russell DW.** 2001. Molecular cloning : a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
2. **Harwood CR, Cutting SM.** 1990. Molecular biological methods for *Bacillus*. Wiley, Chichester ; New York.
3. **Bradford MM.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem **72:**248-254.
4. **Yanisch-Perron C, Vieira J, Messing J.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103-119.
5. **Rahmer R, Morabbi Heravi K, Altenbuchner J.** 2015. Construction of a super-competent *Bacillus subtilis* 168 using the P_{mtlA} -*comKS* inducible cassette. Front Microbiol **6:**1431.
6. **Hoffmann J, Bóna-Lovász J, Beuttler H, Altenbuchner J.** 2012. *In vivo* and *in vitro* studies on the carotenoid cleavage oxygenases from *Sphingopyxis alaskensis* RB2256 and *Plesiocystis pacifica* SIR-1 revealed their substrate specificities and non-retinal-forming cleavage activities. FEBS J **279:**3911-3924.