MUTATIONS OF THE GLUCOSE-PTS IN SUBLANCIN-RESISTANT *B. SUBTILIS* 168 ΔSPB

BY

MARC RYAN GANCAYCO

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Chemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2018

Urbana, Illinois

Adviser:

Professor Wilfred van der Donk

ABSTRACT

The mode of action of the glycocins is still unknown. The glycocin family is unusual due to the rarity of glycosylated peptides in bacteria and of cysteine S-glycosylations in general. S-glycosylation of cysteine has previously been shown to result in a metabolically more stable conjugation than the more common O-glycosylation of serine. This stabilization is supported by the high stability of sublancin and may explain the need for glycosylation in sublancin for antimicrobial activity. Previous studies have shown that the deletion of and mutations in the glucose phosphotransferase system confer resistance to sublancin. In this work we add to this knowledge by generating new sublancin-resistant mutants in *B. subtilis* 168 ΔSPβ. The mutations found support previous findings but also demonstrate the need to look for larger genome changes that may affect the regulation of the glucose phosphotransferase system.

ACKNOWLEDGEMENTS

I would like thank my principal investigator, Wilfred A. van der Donk. I could not have gotten this far without your kindness and compassion. Thank you for giving me the opportunity to work in your lab.

I dedicate this to my wife, Priya. You give me strength I didn't know I had.

TABLE OF CONTENTS

CHAPTER 1: SUBLANCIN AND THE GENERATION OF SUBLANCIN-RESISTANT MUTANT *B. SUBTILIS* 168 ΔSPB

1.1 INTRODUCTION

The discovery and development of antibiotics is one of the greatest achievements of modern medicine. However, the overuse and abuse of antibiotics has led to a spreading of antibiotic resistance that may become a significant threat to world health. The World Health Organization (WHO) estimates that 700,000 deaths per year worldwide are due to antimicrobial resistance. If nothing is done to counteract this resistance, the number of deaths is estimated to increase to 10 million by 2050.[1] The van der Donk lab seeks to help fill the void for novel antibiotics by studying the biosynthesis of ribosomally-synthesized and post-translationally modified peptide (RiPP) natural products.[2]

The increase in antibiotic resistance less than a century after the first development of antibiotics for human use is an evolutionary inevitability that has been complicated by the overuse, abuse, and poor regulation of antibiotics. New antibiotics will always be needed; however, the number of new antibiotics entering the market has strongly declined in recent decades. This is largely due to the high cost of developing a lead compound, the low success rate of regulatory approval, and the relatively low return on investment. Development of new antibiotics in the pharmaceutical industry is also hampered by the rediscovery of already known antibiotics.[3]

For the first few decades of antibiotic discovery and development, activity was typically detected as zones of growth inhibition on lawns of sensitive bacteria in response to either the antibiotic producer or a lead compound. This ‘low hanging fruit’ approach was so successful that

by the 1970s the rate of antibiotic discovery had slowed dramatically.[1,4,5] With the advent of powerful DNA sequencing technologies, many laboratories have suggested that the future of antibiotic discovery lies in mining the genomes of bacteria for silent biosynthetic gene clusters. By looking for unique gene clusters we may be able to predict the structure and function of novel antibiotics with undiscovered antimicrobial activity.

One antibiotic discovered by the activity-based fractionation approach is the RiPP produced by *B. subtilis* 168, sublancin. First described in 1998 by Paik et al. it was not until 2011 that study of the biosynthetic enzymes revealed sublancin was not a lantibiotic but an S-linked glycopeptide.[6,7] The purpose of this novel glycosylation is still under investigation but it has been found to be essential for bioactivity.

1.2 INTRODUCTION TO GLYCOCINS / SUBLANCIN

Sublancin is a member of a fast-growing collection of RiPP natural compounds. RiPP modifications typically convert a linear peptide into a more rigid, often cyclic, structure. Many of the studied RiPPs are comprised of an N-terminal leader peptide followed by a C-terminal core peptide. The N-terminal peptide is believed to direct enzymatic modification of the core peptide. Subsequent cleavage of the N-terminal peptide leads to a fully modified biologically active core peptide.[2,8]

First identified in *B. subtilis* 168, sublancin has shown extreme thermal stability and good antimicrobial activity against medically relevant Gram-positive bacteria such as methicillin resistant *Staphylococcus aureus* (MRSA).[6] The mechanisms used by glycocins to kill bacteria are currently unknown. Glycocins are characterized by the glycosylation of either a cysteine or serine residue and two nested disulfide bonds. While sublancin was initially thought to be a

lantibiotic, the mass spectrometry (MS) and nuclear magnetic resonance (NMR) data have shown sublancin to be a novel S-linked glycopeptide.[9]

1.3 GENE ORGANIZATION

Glycocin biosynthetic genes are clustered and do not currently have a universal gene designation. Sublancin biosynthetic genes have been designated as *sun* and serve as an excellent model for almost all known glycocins. The biosynthetic gene cluster encodes a precursor peptide (SunA), a glycosylation enzyme (SunS), two thiol-disulfide oxidoreductases (BdbA and BdbB), an immunity protein (SunI), and a bifunctional transporter/protease (SunT). All of these genes are located on a SPβ prophage which has integrated into the chromosome of *B. subtilis* 168. Biosynthesis of sublancin has been shown to first require glycosylation by SunS, followed by peptide folding by BdbA and BdbB. Finally, the leader peptide is removed and transported out of the cell by SunT.[7]

1.4 THE ROLE OF THE PHOSPHOTRANSFERASE SYSTEM IN SUBLANCIN SENSITIVITY

The importance of the phosphotransferase system (PTS) was first discovered independently by the van Dijl and van der Donk groups. The van Dijl Lab found that the deletion of the PTS operon (*ptsGHI*) resulted in resistance to sublancin in a sensitive strain of *B. subtilis* 168. Of note for later discussion, the van Dijl group also found that the deletion of *glcT,*[10] which acts as a regulator of the PTS operon, did not confer sublancin resistance. At the same time, resistant strains generated in the van der Donk Lab were shown to have mutations in the PTS gene cluster.

The glucose phosphotransferase system is comprised of three genes. *ptsG* encodes the major glucose transporter and is comprised of three domains (EIIA, EIIB, & EIIC).[11] PtsH and PtsI, better known as Histidine Protein (HPr) and Enzyme E I (EI) respectively, are general components of the phosphotransferase system, capable of aiding in the transport of any sugar molecule that uses the PTS, not only glucose. In *B. subtilis* 168, the phosphoryl group of phosphoenolpyruvate is transferred successively to conserved histidine residues in EI, HPr, and then EIIA. EIIA is specific for glucose and transfers the phosphoryl group to a cysteine residue on EIIB. Finally, EIIB phosphorylates glucose as it crosses the plasma membrane through the transmembrane domain EIIC forming glucose-6-phosphate.[12] The phosphorylated sugar can then enter glycolysis.

To investigate resistance to sublancin, Dr. Garcia de Gonzalo performed comparative genomic analysis of sublancin-resistant *B. halodurans* C-125. Mutants were generated by treating *B. halodurans* C-125 with sublethal concentrations of sublancin in Luria broth (LB). Of the four sublancin-resistant mutants that she analyzed for SNPs and small indels, sublancin-resistant *B. halodurans* C-125 001 contained three mutations in the region between the transcriptional antiterminator (*glcT*) and the pts operon (*ptsGHI*), while sublancin-resistant *B. halodurans* C-125 003 contained one non-sense mutation in the PTS glucose-specific subunit IIA. Sublancin-resistant *B. halodurans* C-125 002 had one missense mutation in the gene for mannitol-1-phosphate 5-dehydrogenase, which is indirectly involved in the glucose pts. Sublancin-resistant *B. halodurans* C-125 004 had no significant SNPs or indels, yet maintained resistance to sublancin.

Based on these results, new sublancin-resistant mutants were generated in this study using the sublancin sensitive strain prepared by the van Dijl group, *B. subtilis* 168 ΔSPβ. This

strain has been shown to be more sensitive to sublancin than *B. halodurans* C-125 and has more tools available for genetic manipulation and analysis. The mutants were also generated in media where the carbon source was restricted to potentially force mutations outside of the glucose pts.

1.5 GENERATION OF SUBLANCIN-RESISTANT MUTANTS IN MINIMAL MEDIA

Dr. Garcia de Gonzalo's bioassays of reconstituted sublancin and its analogs carrying different sugars against *B. halodurans* C-125 and sublancin-resistant *B. halodurans* C-125 demonstrated that the mutation(s) that allowed for sublancin-resistance were unable to grow in M9 minimal media where glucose is the sole carbon source. These mutants were generated using sublethal concentrations of sublancin in LB media where the bacteria had ample access to alternative carbon sources. In order to investigate potential resistance mechanisms further, I generated sublancin-resistant mutants in M9 minimal media where the sole carbon source was either mannose or glucose. By generating sublancin-resistant mutants under these conditions we hoped it would force the bacteria to make mutations outside of the glucose-phosphotransferase system, since in one case (mannose) the pts is not essential, whereas in the other (glucose) mutations in the pts would not allow the bacteria to grow.

For these experiments, *B. subtilis* 168 ΔSPβ was used as the wild-type. *B. subtilis* 168 is a long-studied bacterium which is auxotrophic for tryptophan, requiring supplemental amounts of trp for proper growth. In *B. subtilis* 168, the SP-beta prophage contains all the genes required for the production of sublancin. Deletion of the SP-beta prophage has been shown to result in increased sensitivity to sublancin compared to *B. halodurans* C-125.[13]

The complete genome sequences of eight sublancin-resistant mutants grown in M9 (glucose) and seven mutants grown in M9 (mannose) of *B. subtilis* 168 ΔSPβ were analyzed for

SNPs and indels. In all but one genome, mutations were found in either the glc-PTS or genes involved in the regulation of the glc-PTS, *glcT* and *glcR*. Mutations in the *glcT* gene were found in one of the mutants (Mut 6) grown in M9 glucose and four of the mutants (Mut 1, 3, 4, & 7) grown in M9 mannose. These findings in part contradict what the van Dijl group found when they deleted the *glcT* gene and found no resistance was conferred. All of the mutations in the *ptsG* gene were found in subunit EIIC.

Table 1.1: Sublancin-resistant mutants of *B. subtilis* 168 ΔSPβ grown in M9 minimal media with glucose as the sole carbon source.

Mutant	Position (s)	Gene	SNP	INDEL	Allele freq.	Misc.
Mut1GlcBacillus	161239	rrna32 (exon, noncoding modifier)	-	C (insertion)	1	Multiple Alignments, Ribosomal RNA 16S
Mut2GlcBacillus	1086539	hemE (silent, low impact)	C -> T	-	0.125	Uroporphyrinogen decarboxylase. Only found in Mut2GlcBacillus
Mut3GlcBacillus	3739868	glcR (missense, moderate impact)	C -> T	-	0.125	Transcriptional regulator of deoR (deoxyribonucleo-side regulator). Only found in Mut3GlcBacillus
Mut4GlcBacillus	40112-40123	yaaQ (frame shift, moderate impact)	-	TGATTCCTATG (deletion)	0.125	Uncharacterized protein. Only found in Mut4GlcBacillus
	3739743 - 3739744	glcR (frame shift, high impact)	-	T (deletion)	0.25	Transcriptional regulator of deoR (deoxyribonucleosid e regulator). Only found in Mut4 & Mut5

Table 1.1 (cont.)

Mut5GlcBacillus	1709209	flhA (frame shift, high impact)	-	ATCGT (insertion)	0.125	Membrane protein involved in the flagellar export apparatus. Only found inMut5GlcBacillus
	3739743 - 3739744	glcR (frame shift, high impact)	-	T (deletion)	0.25	Transcriptional regulator of deoR (deoxyribonucleo-side regulator). Only found in Mut4 & Mut5
Mut6GlcBacillus	1456219 - 1456220	glcT (Frame shift, high impact)	-	A (deletion)	0.125	Antiterminator for *ptsGHI*. Only found in Mut6GlcBacillus
Mut7GlcBacillus	1457586	ptsG (Frame shift, high impact), subunit EIIC	-	TA (insertion)	0.125	PTS system glucose-specific transporter subunit IIABC. Only found in Mut7GlcBacillus
	4157072	yycE (silent, low impact)	C --> T	-	0.125	Hypothetical protein
Mut8GlcBacillus	-	-	-	-	-	Identical to wild-type

Table 1.2: Sublancin-resistant mutants of *B. subtilis* 168 ΔSPβ grown in M9 minimal media with mannose as the sole carbon source.

Mutation	**Position (s)**	**Gene**	**SNP**	**INDEL**	**Allele freq.**	**Misc.**
Mut1ManBacillus	161239	rrna32 (exon, noncoding modifier)		C (insertion)	1	Multiple alignments, Ribosomal RNA 16S
	1456609	glcT (Stop gained, high impact)	C --> A		0.143	Antiterminator for ptsGHI. Only observed in Mut1ManBacillus

Table 1.2 (cont.)

	3746387 - 3746462	flhO (Codon change + deletion, moderate impact)		GCGCTTGT CACATCCA CGTTTGAA AGCTCTGA GACGCCTT GTTTTAATG AATAGGCT ACCTGATT ATTTCCGG CT (deletion)	0.143	Flagellar basal-body rod protein. Only observed in Mut1ManBacillus
Mut2ManBacillus	1458267	ptsG (Stop gained, high impact), subunit EIIC	C --> T		0.143	PTS system glucose-specific transporter subunit IIABC. Only observed in Mut2ManBacillus
	1699071	fliK (frame shift, high impact)		AGCTGAAA GCC (insertion)	0.714	Probable flagellar hook-length control protein. Only observed in Mut 2, 3, 4, 6, & 7
	2771956	oatA (silent, low impact)	C --> T		0.714	Peptidoglycan O-acetlytransferase. Only observed in Mut 2, 3, 4, 6, & 7
Mut3ManBacillus	1456219 - 1456220	glcT (frame shift, high impact)		A (deletion)	0.429	Antiterminator for ptsGHI. Only observed in Mut3, 4, & 7
	1699071	fliK (frame shift, high impact)		AGCTGAAA GCC (insertion)	0.714	Probable flagellar hook-length control protein. Only observed in Mut 2, 3, 4, 6, & 7
	2771956	oatA (silent, low impact)	C --> T		0.714	Peptidoglycan O-acetlytransferase. Only observed in Mut 2, 3, 4, 6, & 7
	3011820 - 3011821	argH (frame shift, high impact)		C (deletion)	0.143	Argininosuccinate lyase
Mut4ManBacillus	1456219 - 1456220	glcT (frame shift, high impact)		A (deletion)	0.429	Antiterminator for ptsGHI. Only observed in Mut3, 4, & 7

Table 1.2 (cont.)

	1699071	fliK (frame shift, high impact)		AGCTGAAA GCC (insertion)	0.714	Probable flagellar hook-length control protein. Only observed in Mut 2, 3, 4, 6, & 7
	2771956	oatA (silent, low impact)	C --> T		0.714	Peptidoglycan O-acetlytransferase. Only observed in Mut 2, 3, 4, 6, & 7
Mut5ManBacillus	1458301	ptsG (missense, moderate impact), subunit EIIC	C --> A		0.143	PTS system glucose-specific transporter subunit IIABC. Only observed in Mut5ManBacillus
	2105952	yocR (missense, moderate impact)	C --> A		0.143	Sodium-dependent transporter
Mut6ManBacillus	1457429 - 1457430	ptsG (frame shift, high impact), subunit EIIC		G (deletion)	0.143	PTS system glucose-specific transporter subunit IIABC. Only observed in Mut6ManBacillus
	1699071	fliK (frame shift, high impact)		AGCTGAAA GCC (insertion)	0.714	Probable flagellar hook-length control protein. Only observed in Mut 2, 3, 4, 6, & 7
	2771956	oatA (silent, low impact)	C --> T		0.714	Peptidoglycan O-acetlytransferase. Only observed in Mut 2, 3, 4, 6, & 7
Mut7ManBacillus	1456219 - 1456220	glcT (frame shift, high impact)		A (deletion)	0.429	Antiterminator for ptsGHI. Only observed in Mut3, 4, & 7
	1699071	fliK (frame shift, high impact)		AGCTGAAA GCC (insertion)	0.714	Probable flagellar hook-length control protein. Only observed in Mut 2, 3, 4, 6, & 7
	2088973	yocD (frame shift+stop gained, high impact)		ATATAAT (insertion)	0.143	Putative carboxypeptidase. Only observed in Mut7ManBacillus

Table 1.2 (cont.)

	2088974 - 2088977	yocD (codon change + codon deletion, moderate impact)		GCG (deletion)	0.143	Putative carboxypeptidase. Only observed in Mut7ManBacillus
	2088982	yocD (silent, low impact)	G --> C		0.143	Putative carboxypeptidase. Only observed in Mut7ManBacillus
	2105457	yocR (nonsense, Stop gained, high impact)	G --> T		0.143	Sodium-dependent transporter. Only observed in Mut7ManBacillus
	2771956	oatA (silent, low impact)	C --> T		0.714	Peptidoglycan O-acetyltransferase. Only observed in Mut 2, 3, 4, 6, & 7

It is important to note that two of the four sublancin-resistant mutants of *B. halodurans* C-125 prepared by Dr. Garcia de Gonzalo also conferred resistance despite the lack of significant SNPs and indels found in the analyzed genomes.

Although other methods of glucose transport exist, the PTS is the major glucose uptake system in enteric bacteria.[14] In the future it may be necessary to search for larger architectural changes such as large intergenic mutations. It is possible that to avoid deleterious mutations the regulation of the PTS is being altered by large transposable elements.

Thus, whereas we had hoped that using minimal media with either mannose or glucose as the sole carbon source would lead to mutations in additional potential targets of sublancin, our results show that the PTS was still the locus of mutant generation. These results suggest that the PTS is likely the molecular target of sublancin.

CHAPTER 2: EXPERIMENTAL PROCEDURES FOR THE GENERATION AND ANALYSIS OF SUBLANCIN-RESISTANT MUTANTS OF B. SUBTILIS 168 ΔSPB

2.1 PRODUCTION OF SUBLANCIN

Isolation of soluble sublancin was described previously by Dr. Trent Oman.[7] Overnight cultures of *B. subtilis* 168 were grown under aerobic conditions at 37 ºC for 12-15 h. The overnight cultures were used to inoculate 500 mL of Medium A (at 1%) in 2 L flasks. Medium A is comprised of 900 mL of a Medium A nutrient broth and 100 mL of 10X Medium A salts. The Medium A nutrient broth consists of 20 g sucrose, 11. 7 g citric acid, 4 g Na_2SO_4, 4.2 g $(NH_4)_2HPO_4$, and 5 g yeast extract all dissolved in 900 mL of deionized water. The pH was adjusted to 6.8-6.9 using NaOH and the medium was autoclaved. The 10X Medium A salts is comprised of 7.62 g KCl, 4.18 g $MgCl_2$-6 H_2O, 0.543 g $MnCl_2$-4 H_2O, 0.49 g $FeCl_3$-6 H_2O, and 0.208 g $ZnCl_2$ dissolved in 1 L of deionized water followed by sterilization using a 0.22 mm filtration device. The cultures were grown in medium A under aerobic conditions at 37 ºC for 28-48 h with agitation at 210 rpm. A pinkish-brown color change and drop in pH (~5.0-6.0) is typically observed after 15 h and is consistent with sublancin production. The cultures are then grown an additional 24-48 h after the color change is observed.

The cultures were acidified to pH 2.5 with concentrated phosphoric acid (85% in water) and centrifuged at 6000x rcf for 10 min at 4 ºC to remove cells and insoluble material. Soluble sublancin was isolated from the supernatant by ammonium sulfate precipitation. In a 2 L flask, 490 g $(NH_4)_2SO_4$ per 1 L was combined with the sublancin-containing supernatant. It is important to dissolve the $(NH_4)_2SO_4$ in several small batches over 20-30 min at 4 ºC. The precipitation mixture was stirred for 24 h at 4 ºC. Crude sublancin was precipitated by

centrifugation at 5000x rcf for 30 min at 4 ºC. After carefully disposing of the supernatant, a thin brownish film of sublancin-containing precipitant was re-suspended in 1:1 ACN:water and analyzed by MALDI-TOF MS to confirm it contained sublancin. A 1 μL aliquot of the precipitant was combined with 1 μL of saturated solution of α-cyano-4-hydroxy- cinnamic acid matrix in 1:1 ACN: water with 0.1% TFA. The total volume was spotted onto a MALDI plate and dried under ambient conditions prior to analysis. The crude sublancin suspension can be lyophilized, and stored under N_2 at -80 ºC until purification by preparative HPLC.

Preparative HPLC was performed on a Shimadzu HPLC with a Phenomenex Jupiter C12 column equilibrated in 2% solvent B (solvent A = 0.1% TFA in water, solvent B = 0.0866% TFA in 80% ACN/20% water). Crude sublancin was re-suspended in 10 mL of 2% solvent B. The solution was spun down at 4500x rcf for 10 min at 4 ºC to pellet insoluble peptides. The crude sublancin was filtered using a 0.45 PDVF filter before loading the sample on to the HPLC column. Sublancin was eluted by maintaining the mobile phase at 2% B for 1 min, followed by an increase to 100% B over 35 min with a flow rate of 8.0 mL/min. Sublancin eluted at 23.0-24.2 min. All fractions were analyzed by MALDI-TOF MS as described above. Purified sublancin was lyophilized and stored under N_2 at -80 ºC. Yields were typically 6-8 mg sublancin per liter of Medium A culture.

2.2 GENERATION OF SUBLANCIN-RESISTANT *B. SUBTILIS* 168 ΔSPβ

Sublancin-resistant *B. subtilis* 168 ΔSPβ was generated by treating cells with sublethal amounts of sublancin in liquid M9 minimal medium with 50 μg/mL tryptophan and either 4 g/L mannose or 4 g/L glucose as the sole carbon source. Enriching for mutants was achieved by plating on agar plates with the same nutrient composition but 4x the MIC of sublancin. Single

colonies were selected and grown in identical liquid minimal medium again with 4x the MIC of sublancin. The minimal inhibitory concentration (MIC) against the wt strain was determined to be 1 µM (man) and 25 µM (glc) under these growth conditions. The experiment was performed as described below. For cells grown in M9 minimal media with 50 µg/mL tryptophan (required for this *B. subtilis* strain) where mannose was the sole carbon source, dilutions of 0.03125-2.0 µM sublancin were used.[15] For cells where glucose was the sole carbon source dilutions of 1.56-100 µM sublancin were required. Stock sublancin solutions were filtered using costar spin-x filters (0.22 µm) for 3 min at 1.7 x 10^3 rpm. Using a sterile 96 well plate the total volume of culture in each well was 100 µL. 10 µL of 10x stock sublancin at defined concentrations and 90 µL of a 1-in-10 dilution (approximately 1 x 10^8 colony-forming units (CFU) mL^{-1}) of a culture of *B. subtilis* 168 ΔSPβ with M9 minimal medium (glc or man). The following negative controls were prepared 1.) 90 µL of a 1-in-10 dilution of a culture of *B. subtilis* 168 ΔSPβ and 10 µL of M9 minimal medium (glc or man) and 2.) 100 µL of M9 minimal medium (glc or man). The 96-well plate was incubated at 37 ºC, with shaking at 250 rpm for 28 h. The optical density at 600 nm (O.D.600 nm) was recorded at half hour intervals from 0 to 28 h using a BioTek Synergy 4H plate reader. The readings of triplicate experiments were averaged.

10 µL of sublancin-resistant *B. subtilis* 168 ΔSPβ that was generated at the MIC (1 µM-man, 25 µM-glc) was used to inoculate M9 minimal media (5 mL with 50 µg/mL tryptophan, 4 g/L mannose or glucose) at 4x MIC to retain any mutations. After incubating at 37 ºC, 250 rpm for 4-5 h the cultures were streaked for colony isolation on M9 minimal media agar plates (50 µg/mL tryptophan, 4 g/L mannose or glucose, and 4x MIC sublancin) and placed in the incubator overnight at 37 ºC. Several colonies were picked and grown overnight in liquid M9 minimal media (50 µg/mL tryptophan, 4 g/L mannose or glucose, and 4x MIC sublancin). These samples

were now considered to only contain sublancin-resistant *B. subtilis* 168 ΔSPβ. Glycerol stocks were prepared and stored at − 80 ºC.

2.3 READ PREPARATION OF MUTANT SAMPLES

The following procedure was developed by Gloria Rendon at the Roy J. Carver Biotechnology Center. The shotgun genomic libraries were prepared with the Hyper Library construction kit from Kapa Biosystems. The libraries were quantitated by qPCR and sequenced on one lane for 151 cycles from each end of the fragments on a HiSeq 4000 using a HiSeq 4000 sequencing kit version 1. Fastq files were generated and demultiplexed with the bcl2fastq v2.17.1.14 Conversion Software (Illumina). Each paired read was 150 nt in length. The size of the genome is 4.14 Mb and the read coverage is ~3000x.

The file names of the samples were assigned by me and they include strings that indicate the kind of treatment that they received prior to DNA extraction and sequencing. Samples that received the glucose treatment had filenames with this pattern Mut#GlcBacillusMRG where [#] is a digit. Samples that received the mannose treatment had filenames with this pattern Mut#ManBacillusMRG.

Paired reads of each sample were aligned against the reference genome of Bacillus subtillis 168[16] with bwa[17] version 1.7.12, with default parameters. Then deduplication was performed with MarkDuplicates, part of Picard-tools[18] version 2.4.1. Finally sorting by coordinates was done with novosort, part of novocraf[19] version 3.02. The resulting file is sometimes called “analysis-ready reads” and its format is BAM.[20]

The same variant discovery pipeline for DNA samples was run on the raw reads of the wild type strain, and 30 raw variants compared to the reference genome were discovered in this

sample. These raw variants were then annotated with functional information in the same way used for the 15 mutant samples; that is, functional annotations with SnpEff and with GATK's VariantAnnotator. Figure 2.1 shows the result of one of the reports generated by SnpEff on this sample.

Number variantss by type

Type	Total
SNP	7
MNP	0
INS	15
DEL	8
MIXED	0
INTERVAL	0
Total	30

Figure 2.1. SnpEff classification of variants found in WT strain.

Figure was prepared by Gloria Rendon, HPCBio, University of Illinois

REFERENCES

(1) O'Neill, J. Rev. on Antimicrobial Resistance 2014. "Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations". https://amr-review.org/Publications.html

(2) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Goransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Muller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Sussmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A. Nat Prod Rep 2013, 30, 108. "Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature".

(3) Cooper, M. A. & Shlaes, D. Nature 2011, 472, 32. "Fix the antibiotics pipeline".

(4) Fair, R. J.; Tor, Y. Perspect Medicin Chem 2014, 6, 25. "Antibiotics and bacterial resistance in the 21st century".

(5) Lange, R. P.; Locher, H. H.; Wyss, P. C.; Then, R. L. Curr Pharm Des 2007, 13, 3140. "The targets of currently used antibacterial agents: lessons for drug discovery.".

(6) Paik, S. H.; Chakicherla, A.; Hansen, J. N. J Biol Chem 1998, 273, 23134. "Identification and characterization of the structural and transporter genes for, and the chemical and biological properties of, sublancin 168, a novel lantibiotic produced by Bacillus subtilis 168".

(7) Oman, T. J.; Boettcher, J. M.; Wang, H.; Okalibe, X. N.; van der Donk, W. A. Nat Chem Biol 2011, 7, 78. "Sublancin is not a lantibiotic but an S-linked glycopeptide".

(8) Oman, T. J.; van der Donk, W. A. Nat Chem Biol 2010, 6, 9. "Follow the leader: the use of leader peptides to guide natural product biosynthesis".

(9) Garcia De Gonzalo, C. V.; Zhu, L.; Oman, T. J.; van der Donk, W. A. ACS Chem Biol 2014, 9, 796. "NMR structure of the S-linked glycopeptide sublancin 168".

(10) Stülke, J.; Martin-Verstraete, I.; Zagorec, M.; Rose, M.; Klier, A.; Rapoport, G. Mol Microbiol 1997, 25, 65. "Induction of the Bacillus subtilis ptsGHI operon by glucose is controlled by a novel antiterminator, GlcT".

(11) Gonzy-Treboul, G.; de Waard, J. H.; Zagorec, M.; Postma, P. W. Mol Microbiol 1991, 5, 1241. "The glucose permease of the phosphotransferase system of Bacillus subtilis: evidence for IIGlc and IIIGlc domains".

(12) Deutscher, J.; Francke, C.; Postma, P. W. Microbiol Mol Biol Rev 2006, 70, 939. "How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria".

(13) Garcia De Gonzalo, C. V.; Denham, E. L.; Mars, R. A. T.; Stülke, J.; van der Donk, W. A.; van Dijl, J. M. Antimicrob. Agents Chemother. 2015, 59, 10.1128/AAC.01519. "The Phosphoenolpyruvate:Sugar Phosphotransferase system is involved in sensitivity to the glucosylated bacteriocin sublancin".

(14) Jahreis, K.; Pimentel-Schmitt, E. F.; Bruckner, R.; Titgemeyer, F. FEMS Microbiol Rev 2008, 32, 891. “Ins and outs of glucose transport systems in eubacteria”.

(15) Zeigler, D. R.; Prágai, Z.; Rodriguez, S.; Chevreux, B.; Muffler, A.; Albert, T; Bai, R.; Wyss, M.; Perkins, J. B. J Bacteriol 2008, 190, 6983. “The Origins of 168, W23, and Other Bacillus subtilis Legacy Strains”.

(16) Reference genome: Bacillus subtilis subsp. subtilis str. 168 https://www.ncbi.nlm.nih.gov/genome/?term=Bacillus%20subtilis%20168

(17) Li H. and Durbin R. (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. Bioinformatics, 25:1754-60. [PMID: 19451168].

(18) Picard. http://broadinstitute.github.io/picard/

(19) Novocraft www.novocraft.com

(20) Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). "The Sequence Alignment/Map format and SAMtools". Bioinformatics 25 (16): 2078–2079. doi:10.1093/bioinformatics/btp352