



Viviana Gomes Correia

Graduated in Molecular and Cell Biology

**Genetic and Functional Analysis
of the *Bacillus subtilis* BSP1
gam Cluster**

Dissertation to obtain the
Master degree in Biotechnology

Supervisor: Isabel M. G. de Sá Nogueira,
Associated Professor, FCT-UNL

Jury:

President: Prof. Doutor Pedro Miguel Ribeiro Viana Baptista
Examiner: Prof. Doutor Adriano José Alves Oliveira Henriques
Member: Prof.^a Doutora Isabel Maria Godinho de Sá Nogueira



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

December 2012

Genetic and Functional Analysis of the *Bacillus subtilis* BSP1 *gam* Cluster

Copyright© reserved to Viviana Gomes Correia, FCT/UNL

The Faculty of Science and Technology and the New University of Lisbon have the perpetual right, and without geographical limits, to archive and publish this dissertation through press copies in paper or digital form, or by other known form or any other that will be invented, and to divulgate it through scientific repositories and to admit its copy and distribution with educational or research objectives, non-commercial, as long as it is given credit to the author and editor.

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

ACKNOWLEDGEMENTS

*“Sê todo em cada coisa. Põe quanto és
No mínimo que fazes.”*

Ricardo Reis in *Odes*

First and foremost I offer my sincerest gratitude to my Professor and Supervisor, Isabel de Sá Nogueira, not only for her support and guidance throughout my thesis, but mainly for her contagious passion for science.

Thank you Isabel, Mário and Lia for the wise advices, precious teachings and of course, for the cheerful atmosphere at the 327 lab, even at late hours. One simply could not wish for a better environment to work in. I've learned so much from working with you guys! Also thanks Sónia, Diana, Liliana and Maria João.

I would also like to extend my thanks to every worker, researcher and professor at DCV, specially to Raquel, for the constant good mood and to Dr. Alan Phillips and Jorge Santos, for excellent previous scientific orientation.

Thanks Inês and Joana for the friendship; Pernes and Gralheira for the stupid jokes at lunch time; and all the Biotechnology Master's colleagues of 2010/11 class. Thanks Sofia, Vasco, Tânia, Nadine, Filipa, Cátia, Rui and Marino, BCM's future brightest scientists, for all the good moments.

I would also like to acknowledge all the teachers and friends that I've made along the way, especially to Professors Manuela, Mónica, Pedro and Ana Bela. Their commitment and enthusiasm motivated me to proceed with my studies and always give my best.

Most importantly, I wish to thank to the ones that made this possible, my family! My Mom, Anabela, for being my cornerstone, for made me believe in my value and for teaching me how to never give up. My Dad, João, for the financial support and the valuable lessons about hard working and dedication. Thanks Jonny, my not so little brother, for making me repeat 8th grade all over again while doing the thesis. Thanks Aliete and Gaspar for everything. Special thanks to my Grandparents, for being so proud of me. My work is dedicated to you.

Finally, I would like to thank Sérgio, for so many years of friendship, love and laughter. Thank you so much for always being there above all and no matter what.

A special thanks to Professor Adriano Henriques, Cláudia Serra (Microbial Development Lab - ITQB) and Dr. Ghislain Schyns (DSM) for kindly provide BSP1 strain and for sharing unpublished data.

I could not forget *Bacillus subtilis*, for its, not always easy, collaboration.

Thank you all.

ABSTRACT

Mannans (linear mannan, glucomannan, galactomannan and galactoglucomannan) are the major constituents of the hemicellulose fraction in softwoods and show great importance as a renewable resource for fuel or feedstock applications. As complex polysaccharides, mannans can only be degraded through a synergistic action of different mannan-degrading enzymes, mannanases.

Microbial mannanases are mainly extracellular enzymes that can act in wide range of pH and temperature, contributing to pulp and paper, pharmaceutical, food and feed, oil and textile successful industrial applications. Knowing and controlling these microbial mannan-degrading enzymes are essential to take advantage of their great biotechnological potential.

The genome of the laboratory 168 strain of *Bacillus subtilis* carries genes *gmuA-G* dedicated to the degradation and utilization of glucomannan, including an extracellular β -mannanase. Recently, the genome sequence of an undomesticated strain of *B. subtilis*, BSP1, was determined. In BSP1, the *gmuA-G* operon is maintained, interestingly, however, a second cluster of genes was found (*gam* cluster), which comprise a second putative extracellular β -mannanase, and most likely specify a system for the degradation and utilization of a different mannan polymer, galactoglucomannan.

The genetic organization and function of the *gam* cluster, and whether its presence in BSP1 strain results in new hemicellulolytic capabilities, compared to those of the laboratory strain, was address in this work.

In silico and *in vivo* mRNA analyses performed in this study revealed that the *gam* cluster, comprising nine genes, is organized and expressed in at least six different transcriptional units. Furthermore, cloning, expression, and production of Bbsp2923 in *Escherichia coli* was achieved and preliminary characterization shows that the enzyme is indeed a β -mannanase. Finally, the high hemicellulolytic capacity of the undomesticated *B. subtilis* BSP1, demonstrated in this work by qualitative analyses, suggests potential to be used in the food and feed industries.

KEYWORDS

Bacillus subtilis; undomesticated strain; Galactoglucomannan; Galactomannan; Glucomannan; Mannanase; *gam* cluster; Hemicellulose;

RESUMO

Os mananos (manano linear, glucomanano, galactomanano e galactoglucomanano) são as principais hemiceluloses presentes em gimnospérmicas (*softwoods*), constituindo uma importante fonte de material renovável. Como polissacáridos complexos, os mananos são apenas completamente degradados pela acção conjunta de várias enzimas hidrolíticas, denominadas mananases.

As mananases de origem microbiana são essencialmente enzimas extracelulares que podem actuar numa ampla gama de temperatura e pH, o que lhes confere excelentes qualidades para serem usadas nas indústrias da pasta e do papel, farmacêutica, alimentar e de rações, têxtil ou do petróleo. Conhecer e controlar estas enzimas é fundamental para tirar partido do seu potencial biotecnológico.

O genoma da estirpe de laboratório *Bacillus subtilis* 168 contém genes, *gmuA-G*, dedicados à degradação e utilização do glucomanano, os quais incluem uma β -mananase extracelular. Recentemente, foi determinada a sequência do genoma de uma estirpe “não domesticada” de *B. subtilis*, BSP1. Em BSP1, a operação *gmuA-G* é mantida mas, curiosamente, está presente um segundo aglomerado de genes (*gam cluster*) que inclui uma segunda possível β -mananase extracelular. O *gam cluster* deverá constituir um sistema para a degradação e utilização de um outro tipo de manano, o galactoglucomanano.

Neste projecto, foi estudada a organização genética e funcional do *gam cluster*, bem como se a sua presença no genoma de BSP1 lhe confere capacidades hidrolíticas adicionais, quando comparada com a estirpe laboratorial *B. subtilis* 168.

As análises realizadas *in silico* e *in vitro* ao mRNA, revelaram que o *gam cluster* inclui nove genes e está organizado em pelo menos seis unidades transcricionais. Além disso, Bbsp2923 foi clonada, expressa e produzida em *Escherichia coli*. A sua caracterização preliminar mostra que a enzima é de facto uma β -mananase. Finalmente, a elevada capacidade hemicelulolítica da estirpe “não domesticada” *B. subtilis* BSP1, demonstrada neste trabalho, sugere o seu potencial para aplicações nas indústrias alimentar e de rações-animais.

PALAVRAS-CHAVE

Bacillus subtilis; estirpe “não-domesticada”; Galactoglucomanano; Galactomanano; Glucomanano; Mananase; *gam cluster*; Hemicelulose;

CONTENTS

Acknowledgements.....	V
Abstract.....	VII
Resumo.....	IX
Contents.....	XI
Figures Index.....	XIII
Tables Index.....	XV
Abbreviations, Symbols and Notations.....	XVII
1. General Introduction.....	1
1.1. <i>Bacillus subtilis</i> – An overview.....	3
1.2. Carbohydrate uptake.....	4
1.2.1. <i>PTS</i> system.....	5
1.2.2. <i>ABC</i> transporters.....	5
1.3. Hemicellulose degradation.....	6
1.3.1. <i>Mannan structure</i>	7
1.3.2. <i>Mannan-degrading enzymes</i>	9
1.3.3. <i>Biotechnological application of Mannanases</i>	11
1.4. Heteromannans degradation in <i>B. subtilis</i>	14
1.5. A novel mannan degrading system – <i>gam</i> cluster.....	16
1.6. Scope of the thesis.....	18
2. Materials and Methods.....	19
2.1. Substrates.....	21
2.2. Bacterial strains and growth conditions.....	21
2.3. DNA manipulation and sequencing.....	22
2.4. Plasmids construction.....	23
2.5. Total RNA extraction and Reverse transcription-PCR analysis.....	24

2.6.	Qualitative and comparative analysis of substrate hydrolysis.....	25
2.7.	Hydrolytic activity assays	25
2.9.	Protein analysis and Quantification	27
2.10.	β -Mannanase activity assay	27
3.	Results and Discussion	29
3.1.	Genetic Analysis of the <i>gam</i> Cluster	31
3.1.1.	<i>Transcriptional organization of the gam cluster</i>	<i>31</i>
3.1.2.	<i>Sequence Analysis of the gam cluster</i>	<i>35</i>
3.2.	Functional Characterization of the <i>gam</i> Cluster	43
3.2.1.	<i>Qualitative analysis of substrate hydrolysis</i>	<i>43</i>
3.2.2.	<i>Hydrolytic activity assays</i>	<i>45</i>
3.3.	Cloning and Preliminary Characterization of the Bbsp2923	46
3.3.1.	<i>Sequence analysis of Bbsp2923</i>	<i>47</i>
3.3.2.	<i>Expression of Bbsp2923 in E. coli.....</i>	<i>51</i>
3.3.3.	<i>β-Mannanase activity assay.....</i>	<i>52</i>
4.	Concluding Remarks and Future Perspectives	55
5.	References	61
6.	Appendices	69
Appendix 6.1	71
Appendix 6.2	72
Appendix 6.3	73
Appendix 6.4	75
Appendix 6.5	76
Appendix 6.6	77

FIGURES INDEX

Figure 1.1 – Typical mannan and heteromannans structure.....	7
Figure 1.2 – Mannan-based saccharides and the mannanases responsible for their degradation.....	9
Figure 1.3 – GmuA-G model for glucomannan utilization in <i>B. subtilis</i>	15
Figure 1.4 – <i>B. subtilis</i> 168 chromosome region (A) replaced by <i>gam</i> cluster in <i>B. subtilis</i> BSP1 strain (B).	17
Figure 3.1 – RT-PCR results obtained with RNA extracted from cells grown in different conditions and using distinct pairs of primers	33
Figure 3.2 – Transcriptional organization of the region <i>Bbsp2929-2924</i> of <i>gam</i> cluster	34
Figure 3.3 – Annotated sequence of the putative promoter regions of <i>gam</i> cluster.....	40
Figure 3.4 – Stem-loop structure of a putative transcription terminator found downstream of <i>Bbsp2924</i>	42
Figure 3.5 – Complete <i>gam</i> cluster scheme with identified promoter regions and hairpins.....	42
Figure 3.6 – Hydrolysis test on Locust Bean Gum (LBG) and Birchwood Xylan	44
Figure 3.7 – Illustration of modular organization of <i>Bbsp2923</i> , adapted from BLASTp analysis.	47
Figure 3.8 – N-terminal sequence alignment of <i>Bbsp2923</i> and other GH-26 members correlated with secondary structure prediction information	50
Figure 3.9 – Overproduction of recombinant <i>Bbsp2923</i> -His ₆ in <i>E. coli</i> BL21 pLysS (DE3) and <i>E. coli</i> Rosetta™ (DE3) pLysS strains harboring pVC2.....	52
Figure 3.10 – Chemical structure of used substrates	54

TABLES INDEX

Table 1.1 – List of the major industrial enzymes produced by <i>Bacillus</i> spp.	4
Table 1.2 – Galactose and mannose content of popular galactomannans	8
Table 1.3 – Available information on commercial mannanases.....	11
Table 2.1 – List of bacterial strains used in this work.	22
Table 2.2 – List of oligonucleotides used in this project.....	23
Table 2.3 – List of used or constructed plasmids.....	24
Table 2.4 – List of oligonucleotides used for RT-PCR assays.....	25
Table 3.1 – Predicted signals for translation initiation in <i>gam</i> cluster sequence.....	36
Table 3.2 – Comparison of <i>gam</i> cluster products to other proteins in available databases using BLASTp tool.	37
Table 3.3 – Identified <i>cre</i> sequences and location.	38
Table 3.4 – Typical <i>B. subtilis</i> sigma-factor (σ) binding sequences identified in some <i>gam</i> genes promoters.....	39
Table 3.5 – Extracellular hydrolytic activity of <i>B. subtilis</i> BSP1 and 168 strains towards distinct substrates.....	45
Table 3.6 – Activity rates of recombinant Bbsp2923, against several tested substrates	53

ABBREVIATIONS, SYMBOLS AND NOTATIONS

aa – Amino acid

ABC – ATP-binding cassette

ATP – Adenosine triphosphate

BLAST – Basic Local Alignment Search Tool

bp – Base pairs

BSA – Bovine Serum Albumin

CAZy – Carbohydrate-Active enZymes Database

CBM – Carbohydrate Binding Domain

CCR – Carbon catabolite repression

Cm – Chloramphenicol

CRE – Catabolic responsive elements

DNA – Deoxyribonucleic acid

dNTPs – Deoxynucleotide Triphosphates

EDTA – Ethylenediamine tetra-acetic acid

gam – Galactoglucomannan utilization cluster

GH – Glycosyl Hydrolase

gmu – Glucomannan utilization operon

IPTG – Isopropyl- β -D-galactopyranoside

kb – Kilobase pairs

kDa – KiloDalton

Km – Kanamycin

LA – Luria-Bertani Agar, solid LB medium

LB – Luria-Bertani medium

LBG – Locust Bean Gum, source of galactomannan

MCS – Multiple Cloning Site

mRNA – Messenger ribonucleic acid

NBD – Nucleotide-binding domain

NCBI – National Center for Biotechnology Information

OD – Optical Density

ORF – Open Reading Frame

PC buffer – Citrate-Phosphate buffer

PCR – Polymerase chain reaction

PEP – Phosphoenolpyruvate

PMSF – Phenylmethylsulfonylfluoride

PTS – Phosphotransferase system

Rbs – Ribosome binding site

RNA – Ribonucleic acid

rpm – Revolutions per minute

RT-PCR – Reverse transcription polymerase chain reaction

SBP – Solute-binding protein

SDS-PAGE – Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SF – Sugar Free medium

SFA – Sugar Free agar

TE – Tris-EDTA

TMD – Transmembranar domain

tRNA – Transfer ribonucleic acid

TSS – Translational start site

UV – Ultraviolet light

Chapter 1

General Introduction

1. General Introduction

1.1. *Bacillus subtilis* – An overview

Bacillus subtilis is a gram-positive, rod-shaped, endospore-forming bacterium, ubiquitous found in soil, water sources or in association with plants and animal gastrointestinal tract (Priest, 1993; Casula & Cutting, 2002). *B. subtilis*, whose genome has been completely sequenced (Kunst *et al.*, 1997), was the first non-pathogenic microorganism to be transformed (Anagnostopoulos & Spizizen, 1961) and since then has been used as the gram-positive paradigm for fundamental and applied research.

Bacillus strains, including *B. subtilis*, have developed a series of responses in order to cope with nutrient supply fluctuation, commonly felt in their natural environment. These responses can be observed in the laboratory towards the end of exponential growth phase or in stationary phase, and include onset of motility, production of hydrolytic enzymes, competence to take up exogenous DNA and sporulation (Harwood, 1992).

Sporulation is the ability to form endospores – highly differentiated dormant structures, resistant to desiccation, starvation or radiation, and used as a survival as well as new niche colonization techniques. Works of Casula & Cutting (2002) and Tam and colleagues (2006) proved that spores indeed germinate in the gastrointestinal animal tract, disproving the idea that vegetative cells only develop in the soil or in association with decaying vegetative material. Spores are widely used as probiotics in human diet improving gastrointestinal microflora health by preventing pathogenic bacteria onset (Casula & Cutting, 2002).

The metabolic diversity of *Bacillus* spp. led to the use of these organisms in a wide range of industrial microbial processes, also known as “white biotechnology”, which includes the production of carbohydrate-, lipid- and protein-degrading enzymes, antibiotics, fine biochemicals (vitamins) and insecticides (Harwood, 1992).

Bacillus strains are extremely attractive for microbial biotechnology, due to their ability to secrete large quantities of proteins (grams per litre) into the extracellular medium, their high growth rates leading to short fermentation cycles and their GRAS (generally regarded as safe) status approved by FDA (Food and Drug Administration) (Harwood, 1992; Schallmeyer *et al.*, 2004). Table 1.1 summarizes the major industrial enzymes produced by *Bacillus* spp. *Bacillus subtilis* and close relatives such as *Bacillus amyloliquefaciens* and *Bacillus licheniformis* are of

particular relevance for extracellular protein secretion (Ferrari *et al.*, 1993). It's estimated that 50-60% of the total enzymes market is produced or derived from *Bacillus* spp. (Schallmeyer *et al.*, 2004; Westers *et al.*, 2004).

Table 1.1 – List of the major industrial enzymes produced by *Bacillus* spp. Adapted from Harwood (1992) and Schallmeyer *et al.* (2004).

Industry	Enzymes	Producer strain
Starch	α -Amylases	<i>B. amyloliquefaciens</i> <i>B. liqueniformis</i> <i>B. subtilis</i> <i>B. circulans</i> <i>B. stearothermophilus</i>
Baking, Beverage and Textile	Amylases	<i>B. amyloliquefaciens</i> <i>B. liqueniformis</i> <i>B. stearothermophilus</i> <i>B. halodurans</i>
Pharmaceutical	Alkaline phosphatases	<i>B. liqueniformis</i>
Detergent	Alkaline proteases	<i>B. amyloliquefaciens</i> <i>B. clausii</i> <i>B. Halodurans</i>
Beverage	β -Glucanases	<i>B. subtilis</i>
Starch	Glucose isomerases	<i>B. acidipullulyticus</i> <i>B. deramificans</i>
Pharmaceutical	β -Lactamases	<i>B. liqueniformis</i>
Starch	Pullulanases	<i>B. liqueniformis</i> <i>B. stearothermophilus</i>

1.2. Carbohydrate uptake

Either as soil or gut associated bacteria, *B. subtilis* must be able to sense, degrade and uptake different types of plant-derived cellulose and hemicellulose polymers available as carbon sources. Cellulose and hemicellulose polysaccharides are hydrolyzed by a wide variety of extracellular polysaccharide degrading enzymes, cellulases and hemicellulases. After enzymatic cleavage of these polysaccharides into smaller units, membrane transport systems are vital for the uptake of those essential nutrients, enabling selective passage across the lipid bilayer membrane that separates the cell from the environment.

Membrane transport systems found in all living organisms, including *B. subtilis*, can be classified into four major types: channel proteins, characterized by an energy-independent transport process; secondary active transporters, which use chemiosmotic energy to drive transport; primary active transporters such as ATP-binding cassette (ABC) transporters, that use chemical, electrical or solar energy as transportation force; and group translocating systems, including the phosphotransferase system (PTS) which transport and concomitantly phosphorylate the substrate. In *B. subtilis*, well characterized carbohydrates uptake systems belong mainly to PTS and ABC transporters families (Deutscher *et al.*, 2002).

After entering the cell, all carbohydrates are converted into intermediates of one of the central carbohydrate-degrading pathways – glycolysis, pentose phosphate or Entner-Doudoroff pathways (Deutscher *et al.*, 2002).

1.2.1. PTS system

The phosphoenolpyruvate (PEP): carbohydrate phosphotransferase system or simply PTS, is a complex carbohydrate transportation system that exists in both gram-positive and gram-negative bacteria. The PTS system is involved not only in the transport and phosphorylation of different carbon sources and in movement towards the carbohydrates (chemotaxis) but also in the regulation of other metabolic pathways (Postma *et al.*, 1993).

This system catalyses the simultaneous transport and phosphorylation of the substrate through a phosphorylation cascade involving protein enzyme I (EI), HPr (histidine-containing protein) and PTS sugar-specific transporter proteins – IIA, IIB and IIC. PEP is used as energy source and phosphoryl group provider (Deutscher *et al.*, 2002).

1.2.2. ABC transporters

Existing from microorganisms to man, ABC transporters are membrane proteins that couple the translocation of several substrates across cellular membrane to the hydrolysis of ATP (Higgins, 1992).

ABC transporters are of the largest protein families with a great diversity of physiological functions. Particularly in bacteria, these transporters catalyse the uptake of essential nutrients

(*e.g.* saccharides) or the excretion of toxic compounds, contributing to drug and antibiotic resistance (Davidson & Chen, 2004).

Depending on the direction of the transport, ABC transporters can be classified as exporters or importers. The latest type is only present in prokaryotes and requires a solute-binding protein (SBP, also known as periplasmic binding protein) for substrate capture and delivery to the external face of the transporter. In gram-positive bacteria, like *B. subtilis*, the SBP is often a lipoprotein bound to the cytoplasmic membrane external side or fused to the ABC transporter, conferring high affinity, specificity and directionality to a certain substrate or solute family (Higgins, 2001).

Generally, the basic unit of an ABC transporter comprises four core domains: two transmembranar domains (TMDs) that provide a passageway for substrates and two cytoplasmic nucleotide-binding domains (NBDs) that bind and hydrolyse ATP. Usually, in ABC importers the TMDs and NBDs are encoded as separate polypeptides, although it is known that these domains can be fused to each other as multidomain polypeptides, in any combination (Higgins, 2001).

1.3. Hemicellulose degradation

Hemicelluloses are polysaccharides of the plant cell wall with an important structural role in association with cellulose, lignin and pectin (McNeil *et al.*, 1984).

Hemicellulose fraction corresponds to one third of total components available in plants and is the second most abundant heteropolymer in nature with great importance as a renewable resource for fuel or feedstock applications. This makes hemicellulases, the major industrially important enzymes right after proteases and cellulases (Dhawan & Kaur, 2007; Moreira & Filho, 2008; Chauhan *et al.*, 2012).

According to the main sugar unit, hemicelluloses can be grouped into mannans, xylans, galactans or arabinans. Xylans are the major hemicellulose in hardwoods (angiosperms) and grasses, while mannans comprise the more abundant hemicellulose in softwoods (gymnosperms) and specialized plant structures (*e.g.* seeds and fruits) (Moreira & Filho, 2008; Chauhan *et al.*, 2012).

1.3.1. Mannan structure

Mannan is an important hemicellulose and can be classified in two main categories – homo- or linear mannans and heteromannans (glucomannans, galactomannans and galactoglucomannans) (Fig. 1.1). Each of these polysaccharides possesses a backbone of β -1,4-linked D-mannose residues (linear mannan) which might be combined with D-glucose (glucomannan) or/and D-galactose (galactomannan/galactoglucomannan). The ratio between different sugar residues alters mannan properties and is characteristic of each plant species.

Mannans can be found in gums and are used as thickening agents in the food (ice-creams and sauces) and pharmaceutical (hair gels, shampoos or tooth-pastes) industries (Chauhan *et al.*, 2012).

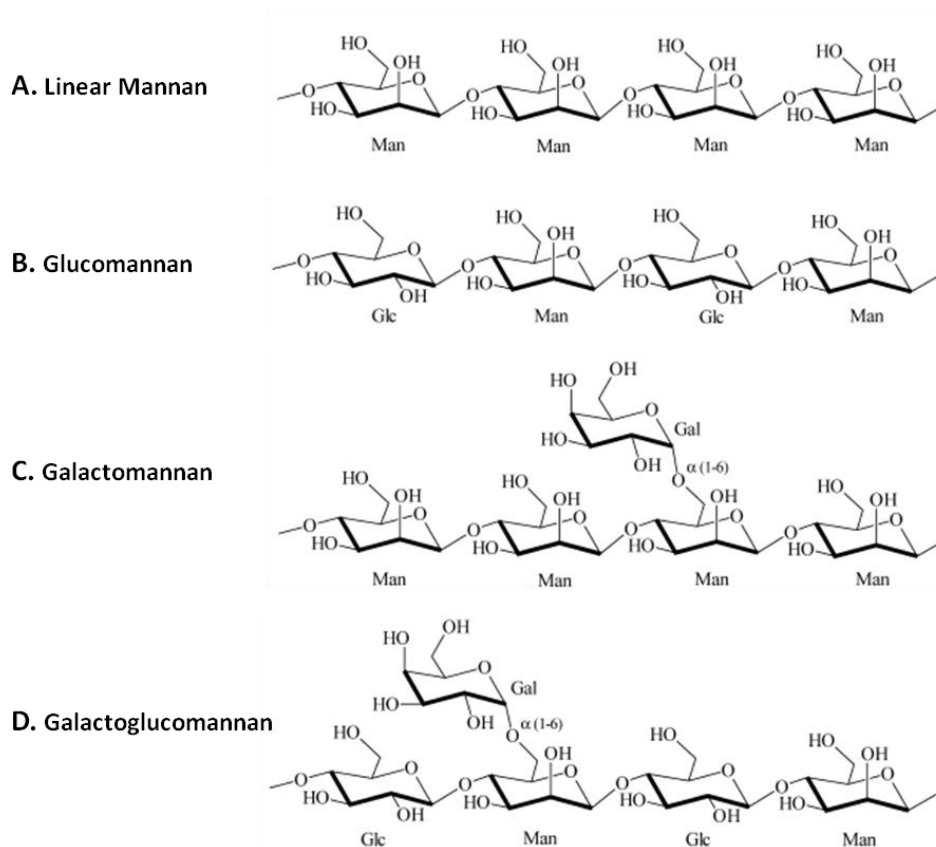


Figure 1.1 – Typical mannan and heteromannans structure. (A) Linear mannan with a backbone of β -1,4 linked mannose (Man) residues; (B) Glucomannan structure formed by a main chain of β -1,4 linked mannose and glucose (Glc) residues; (C) Galactomannan general structure, a main chain of β -1,4 linked mannose residues with α -1,6 linked galactose (Gal) residues attached to some mannose residues; (D) Galactoglucomannan structure constituted by a main chain of Man and Glc β -1,4 linked residues with α -1,6 linked galactose (Gal) side residues. Adapted from Dhawan & Kaur, 2007.

Linear mannans are homopolysaccharides with a backbone of β -1,4 linked D-mannose (β -D-mannopyranosyl) residues (Fig. 1.1-A). These mannans play a major structural role in woods and seed of several plants like ivory nuts (*Phytelephas* spp.), green coffee beans (*Coffea* spp.) or coconut kernel (copra). Mannans with linear chains are known to be highly insoluble in water (Moreira & Filho, 2008).

Glucomannans are formed by a chain of randomly arranged β -1,4 linked D-mannose and β -1,4 linked D-glucose (β -D-glucopyranosyl) residues (Fig. 1.1-B). Usually the observed ratio is of one glucose residue for each three mannans (1:3). Glucomannans backbone mannose or glucose units can be acetylated in the O-2 or O-3 position, providing branching points. These heteromannans can be obtained from several plants like ramie (*Boehmeria nivea*), lupins (*Lupinus* spp.) or Konjac (*Amorphophallus konjac*) (Moreira & Filho, 2008).

Table 1.2 – Galactose and mannose content of popular galactomannans. Ratio and polysaccharide source are referred. Adapted from Nishinari *et al.* (2007) and Moreira & Filho (2008).

Common name	Source	Ratio (Gal:Man) ^a
Fenugreek gum	<i>Trigonella foenum-graecum</i>	1:1
Guar gum	<i>Cyamopsis tetragonoloba</i>	1:2
Tara gum	<i>Caesalpinia spinosa</i>	1:3
Locust bean or carob gum	<i>Ceratonia siliqua</i>	1:4
Cassia gum	<i>Senna obtusifolia</i>	1:5

^a Ratio between galactose (Gal) and mannan (Man) residues.

Galactomannans are formed by a backbone of β -1,4 linked D-mannose residues with side chains of single D-galactose (β -D-galactopyranosyl) groups attached through α -1,6 bonds along the main chain (Fig. 1.1-C). Similarly to what happens in other heteromannans, the ratio between different sugar residues characterizes mannan source, altering mannan properties. In table 1.2, galactose/mannose ratio of different popular galactomannans can be seen. The increasing amount of D-galactose side branches (highly hydrophilic) increases galactomannan polymer solubility in water (Moreira & Filho, 2008).

Galactoglucomannans are the more complex heteromannans and constitute the dominant hemicellulose in softwoods, comprising 25% of the wood dry weight. These mannans contain a backbone of β -1,4-linked D-mannose and D-glucose residues with side

branches of α -1,6-linked D-galactose along the chain (Fig. 1.1-D). Some mannose units can be partially substituted by *O*-acetyl groups in the *O*-2 or *O*-3 position, on the average of one acetylated residue for each 3-4 residues. Like galactomannan, galactoglucomannan presents high solubility in water due to D-galactose side chains that prevents macromolecules from aggregate. Norway spruce (*Picea abies*) is an abundant softwood species that can contain up to 10-20% of acetylated galactoglucomannan (Moreira & Filho, 2008).

Due to the great variety of hemicelluloses available in nature, including mannans, microorganisms have developed over the time many enzymatic systems to overcome this problem and be able to take advantage of this carbon source abundance.

1.3.2. Mannan-degrading enzymes

Complex polysaccharides such as mannans depend on a synergistic action of different mannan-degrading enzymes for complete degradation.

Mannan-degrading enzymes or mannanases involved in linear mannan degradation include β -mannanases, β -mannosidases and β -glucosidases. Heteromannans processing requires some additional enzymes like α -galactosidases and acetyl mannan esterases, for the side-chains substituents removal. Figure 1.2 illustrates the enzymatic action involved in the degradation of two representative mannan-based polymers, galactoglucomannan (Fig. 1.2 - A) and mannobiose (Fig. 1.2 - B).

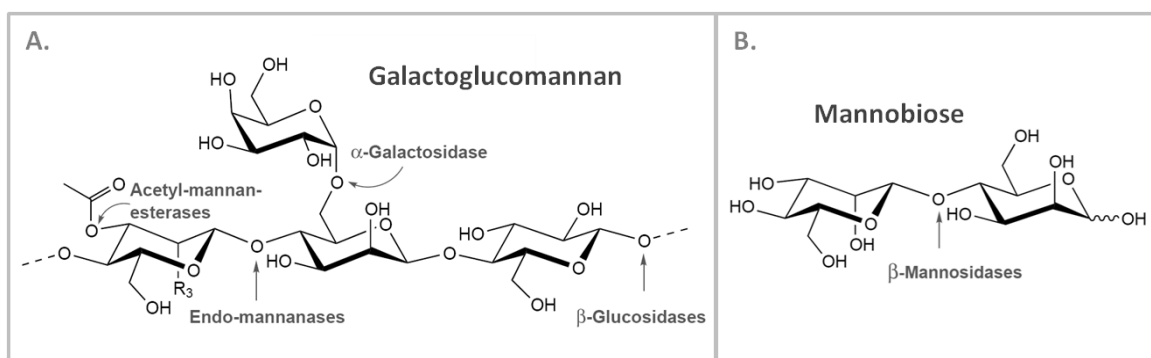


Figure 1.2 – Mannan-based saccharides and the mannanases responsible for their degradation. Adapted from Shallom & Shoham (2003). (A) Galactoglucomannan is a polysaccharide formed by a backbone of D-glucose and D-mannose β -1,4 linked units, containing D-galactose residues attached to both of them. Some of the mannose units can be partial substituted by *O*-acetyl groups (*O*-2 or *O*-3 positions). Galactoglucomannan degradation requires the action of an endo-mannanase, a β -glucosidase, an α -galactosidase and an acetyl mannan esterase; (B) Mannobiose is a disaccharide of mannose, linked by β -1,4 mannosidic bonds. Mannobiose is hydrolysed by a β -mannosidase.

β -Mannanases or endo-1,4- β -mannanases (EC 3.2.1.78, mannan endo-1,4- β -mannosidases) are hydrolases that catalyze the random cleavage of β -1,4 mannosidic linkages within the backbone of different homo- and heteromannans. β -Mannanase activity releases short β -1,4-manno-oligosaccharides, mainly mannobiose and mannotriose. Galactomannan and galactoglucomannan backbone substitution pattern possesses a great influence over β -mannanases activity (Dhawan & Kaur, 2007; Moreira & Filho, 2008). These are the most studied mannanases, mainly due to its abundance and biotechnological importance.

β -Mannosidases (EC 3.2.1.25, mannan exo-1,4- β -mannosidases) are exo-type enzymes that cleave β -1,4 mannosidic linkages releasing mannose units from the nonreducing ends of mannans and manno-oligosaccharides (Dhawan & Kaur, 2007; Moreira & Filho, 2008).

β -Glucosidases (EC 3.2.1.21, 1,4- β -D-glucoside glucohydrolases) constitute a group of exo-acting enzymes that hydrolyze β -1,4 bonds, releasing D-glucose residues from the nonreducing end of glucomannan or galactoglucomannan oligomers (Dhawan & Kaur, 2007; Moreira & Filho, 2008).

α -Galactosidases (EC 3.2.1.22, 1,6- α -D-galactoside galactohydrolases) act as debranching enzymes, cleaving α -1,6-side linked D-galactose units of galactomannans and galactoglucomannans backbones (Dhawan & Kaur, 2007; Moreira & Filho, 2008).

Acetyl mannan esterases or N-acetyl mannosamine esterases (EC 3.1.1.6) also act as debranching enzymes by releasing acetyl groups from galactoglucomannan polymers (Dhawan & Kaur, 2007; Moreira & Filho, 2008).

Mannanases as well as other hemicellulases are listed and classified in CAZy, The Carbohydrate-Active EnZymes database (Cantarel, 2009). CAZy members are classified according to sequence, structure and organization similarities of their catalytic domains and additional non-catalytic modules like carbohydrate binding modules (CBMs). In relation to the catalytic module, mannanases are considered to be glycoside hydrolases (GH) (EC 3.2.1.-), a widespread group of enzymes that hydrolyze the glycosidic bonds of oligo- and polysaccharides (Moreira & Filho, 2008).

1.3.3. Biotechnological application of Mannanases

Over the past years, mannanases have been isolated from many microbial sources like fungi, actinomycetes and bacteria.

Microbial mannanases are mainly extracellular enzymes that can act on a wide range of pH and temperature, which have contributed to many successful industrial applications (Dhawan & Kaur, 2007).

The induction of mannanases production by natural strains normally requires the use of expensive mannan-rich substrates, fact that compromises industrial use. So cloning and expression of mannanases in heterologous hosts for protein overproduction have become a very popular procedure. The production of genetic engineered mannanases with specific alterations and properties that suits commercial applications constitutes other market opportunity (Chauhan *et al.*, 2012). Table 1.3 provides an updated list of commercial mannanases and their suppliers.

Table 1.3 – Available information on commercial mannanases. Adapted from Dhawan & Kaur, 2007.

Product name	Supplier	Microorganism	Applications
Hemicell ^a	ChemGen, USA	<i>T. longibrachiatum</i> and <i>B. lentus</i>	Animal feed supplement
Gamanase ^a	Novo Nordisk, Denmark	<i>Aspergillus niger</i>	Coconut oil extraction
Mannaway ^a	Novozyme, Denmark	<i>Bacillus</i> spp.	Detergent
Purabrite, Mannastar ^a	Genencor, USA	Fungal mannanase	Detergent
Pyrolase 160 or 200 ^b	Diversa (NASDAQ-DVSA), USA	Organism from hydrothermal vents	Oil recovery and well drilling

^a Endo-mannanase.

^b Endo- and exo-mannanase.

In 2010, the global market of industrial enzymes worth about 3.6 billion dollars. The market for 2016 is estimated to reach about 6 billion dollars, with an annual growth rate prediction of 9.1% (Dewan, 2012). Hydrolytic enzymes represent the majority of used industrial enzymes, including proteases and glycoside hydrolases, in which mannanases are catalogued (Moreira & Filho, 2008).

Following paragraphs summarizes current industrial applications of mannanases.

✓ *Biobleaching in pulp and paper industry*

The lignin extraction from wood is an essential step for pulp bleaching. Usually this was achieved through an alkaline pre-treatment that hydrolyze hemicelluloses covalently bound to lignin, thus facilitating lignin removal. However, this treatment releases extremely pollutant chlorinated compounds. Enzymatic treatment of pulps, using mannanases along with other hemicellulases as xylanases, constitutes an environmental friendly alternative for bleaching process. Mannanases specific for galactomannan (major hemicellulose in pulps) that remain active at high temperatures and pH are excellent for this application (Dhawan & Kaur, 2007; Chauhan *et al.*, 2012).

✓ *Hydrolytic agent in detergent industry*

Mannans have high tendency to adsorb to fibres and can be found in gums or thickening agents in many food or beauty products like ice-creams, sauces and hair gels. So, alkaline mannanases stable in detergents have found many applications in laundry segments as stain removal booster for fabrics, health care products and sanitization or hard surface cleaners (Chauhan *et al.*, 2012). Mannaway, Purabrite and Mannastar are some commercially available products (Table 1.3).

✓ *Textile and cellulosic fiber processing*

Mannanases in combination with other hemicellulases can be used in textile industry, by helping on fibres preparation and cleaning before transformation into yarn. Mannanases are also important in enzymatic scouring and desizing processes for cellulosic material preparation before dyeing (Dhawan & Kaur, 2007).

✓ *Hydrolysis of coffee extract*

Coffee extracts are rich in mannan polymers. The use of mannanases to hydrolyze mannans in the coffee results in a significant viscosity reduction, enabling coffee bean extracts to be concentrated by evaporation, a low-cost industrial procedure (Dhawan & Kaur, 2007; Chauhan *et al.*, 2012).

✓ *Food additives*

Mammalian digestive enzymes aren't able to degrade mannan based substrates. But when mannans reach the large intestine, probiotic bacteria of the genera *Bifidobacteria* or *Lactobacillus*, which naturally produce mannanases, can readily metabolize mannans. The food enrichment with mannan degradation products, mannanoligosaccharides, was proved to have a probiotic effect, enhancing growth and proliferation of human beneficial intestinal microflora (Dhawan & Kaur, 2007).

Mannanases can also be used in food industry, helping fruit and vegetables maceration as well as fruit juices clarification (Chauhan *et al.*, 2012).

✓ *Feed improvement*

Feed ingredients are rich in mannans, xylans and arabinoxylans. Due to the lack of appropriate degrading enzymes in animals gut, these substrates are misused and nutritional value wasted. Mannanases can be incorporated in animal diets helping the digestion of these feeds, improving nutrient absorption and decreasing digested food viscosity. For this purpose, mannanases must be active over a wide pH range and resistant to proteases (*e.g.* trypsin, pepsin) in order to resist gastric conditions (Chauhan *et al.*, 2012). Hemicell supplied by ChemGen is an example of industrial use (Table 1.3).

✓ *Pharmaceutical applications*

Mannanases can be used for economical production of mannose from low-cost mannan rich substrates such as guar gum, copra meal or palm kernel cake. Mannose can be useful in urinary tract infections and intestinal disorders treatment. Besides, mannan can help reduce cholesterol and body fat without protein mass loss. Mannan is also used for conferring fast dissolving and structure properties to tablets (Chauhan *et al.*, 2012).

✓ *Oil drilling and extraction*

Mannanases can be used in oil and gas industries for galactomannan hydrolysis, enhancing outflow in drilling operations. Natural galactomannan rich polymers, like guar gum, are used in flooding, capping and pressurizing of the well, thus helping extract the oil and gas. The polymer solution is thinned by adding mannanases, which facilitates the outflow. Termostable

mannanases are required, due to the extreme temperatures in the wells (>80°C) (Moreira & Filho, 2008).

Mannanases can also be used in enzymatic oil extraction of coconut meat rich in mannans. The enzymatic process eliminates aflatoxin contamination, oxidative rancidity of the products and reduces refinement costs (Dhawan & Kaur, 2007).

✓ *Bioethanol production*

An enzymatic cocktail, including mannanases, xylanases and cellulases, can be used to effectively hydrolyze lignocellulosic biomass to fermentable sugars for bioethanol production. This enzymatic treatment eliminates the usual expensive heat pre-treatment of biomass (Várnai *et al.*, 2011).

✓ *Slime control agents*

Alkaline mannanases combined with proteases can prevent slime formation in water purification or cooling systems and sewers or waste water treatment equipments. By degrading mannans rich substrates mannanases can control biofilms formation and bacterial adhesion (Chauhan *et al.*, 2012).

1.4. Heteromannans degradation in *B. subtilis*

The genes coding for polysaccharide-hydrolysing enzymes are generally organized in an operon or regulon, together with genes encoding proteins responsible for the uptake of the extracellular hydrolyzed products and the first intracellular steps in their metabolism (Deutscher *et al.*, 2002).

In 2008, Sadaie and co-workers characterized a novel operon of *B. subtilis* involved in the utilization of a heteromannan. Besides that study, no report has been presented describing mannan utilization in *B. subtilis*. Little is known about the mannan degradation pathway in *B. subtilis*. Although several mannanases with significant biotechnological potential have been isolated and tested, no further genetic characterization has been performed (Chauhan *et al.*, 2012).

The operon *gmuBACDREFG* or simply *gmuA-G* (formerly *ydhMNOPQRST*) was found in the strain 168 and consists in eight genes encoding all the necessary enzymes for the uptake and degradation of glucomannan (Fig. 1.3) (Sadaie *et al.*, 2008).

The *gmuA-G* operon codes for an external β -mannanase (GmuG), responsible for endo-cleavage of the β -1,4 linked glucomannan backbone; a transportation system (GmuA, GmuB and GmuC) homologous to the enzymes IIA, IIB and IIC of lactose-class phosphotransferase systems; a phospho- β -glucosidase (GmuD), a fructokinase (GmuE) and a mannose-6-phosphate isomerase (GmuF) for further oligo-mannans processing after cell internalization; and a DNA-binding transcriptional repressor (GmuR) containing a small molecule-binding UbiC transcription regulator-associated (UTRA) domain of the GntR family (named after the gluconate operon repressor) (Sadaie *et al.*, 2008).

Experimental works suggests that the expression of the *gmuA-G* operon is induced by degraded glucomannan products, such as cellobiose and mannobiose, and repressed not only by the GmuR, but also by glucose in a CcpA-protein dependent manner (Sadaie *et al.*, 2008).

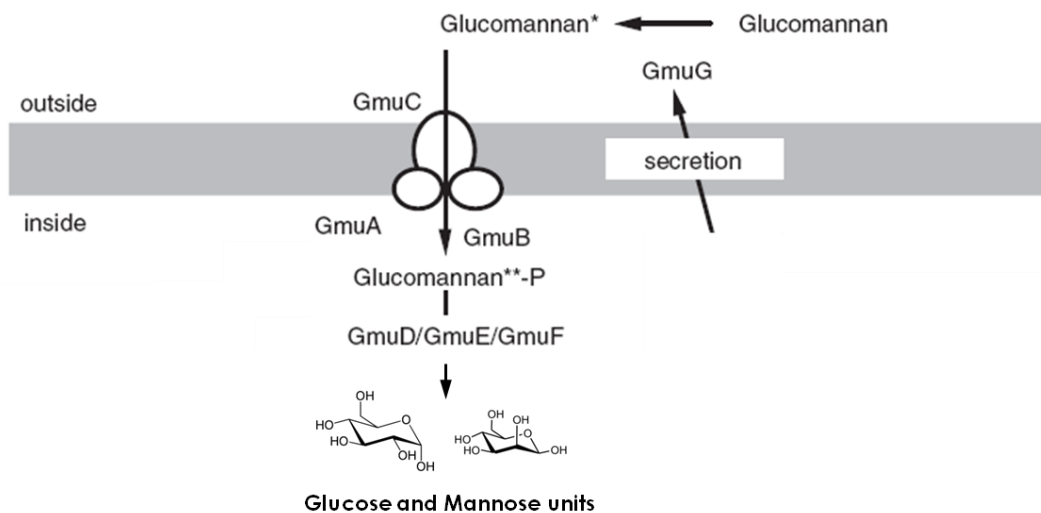


Figure 1.3 – GmuA-G model for glucomannan utilization in *B. subtilis*. Adapted from Sadaie *et al.* (2008). First glucomannan polymers are partially digested by the GmuG extracellular β -mannanase. The resulting oligo-glucomannans are then uptaken by a specific PTS system (GmuBAC) and further processed by GmuD (phospho- β -glucosidase), GmuE (fructokinase) and GmuF (mannose-6-phosphate isomerase) enzymes.

1.5. A novel mannan degrading system – *gam* cluster

Recently, Barbosa and colleagues (2005) reported the discovery of an undomesticated gut-associated strain of *B. subtilis*, the BSP1 (#200). Like *B. subtilis* 168, BSP1 is a gram-positive, endospores-former bacterium, naturally competent and easily transformable by plasmid or genomic DNA.

BSP1 carries about 200 genes not found in *B. subtilis* 168 and some of these genes appear to be related with colony morphology, biofilm formation, mucosal adhesion or hemicellulose utilization. Most of these genes can be considered signatures for probiotic bacteria and evidence of *B. subtilis* cycling between soil and animal gastrointestinal tract (Schyns *et al.*, 2009).

This strain also presents a potent anti-microbial activity and the ability to sporulate during growth, due to the increased activity of Spo0A, a key regulatory protein in *B. subtilis*. Both traits could be advantageous in the gut environment (Schyns *et al.*, 2009).

In *silico* analysis of the BSP1 genome revealed that *gmuA-G* operon is maintained and all the encoded proteins present high similarity to the 168 strain operon. Moreover, a second cluster of genes was discovered in BSP1, which most likely specifies a putative system for the degradation and metabolization of another heteromannan, possibly galactoglucomannan due to putative function of the encoded enzymes (Fig. 1.4) (Schyns G., Serra C.R., Henriques A.O., *et al.* (unpublished data)).

The *gam* cluster (Fig. 1.4) replaces a region of eleven genes found in the chromosome of *B. subtilis* 168. Most of these substituted genes (*yobE-yobJ*) have an unknown function and none is essential to *B. subtilis*. The *gam* cluster, designation that stands for *galactoglucomannan utilization*, encodes a putative extracellular β -Mannanase (Bbsp2923), a CcpA-like regulatory protein (Bbsp2924), a β -Glucosidase (Bbsp2925), an ABC transportation system (Bbsp2926-29), an intracellular α -Galactosidase (Bbsp2930) and a N-acetyl-mannosamine kinase (Bbsp2931). This system is consistent with the degradation of a galactoglucomannan polysaccharide formed by a backbone of D-glucose and D-mannose β -1,4 linked units, with D-galactose side residues attached through an α -1,6 bond (Moreira & Filho, 2008).

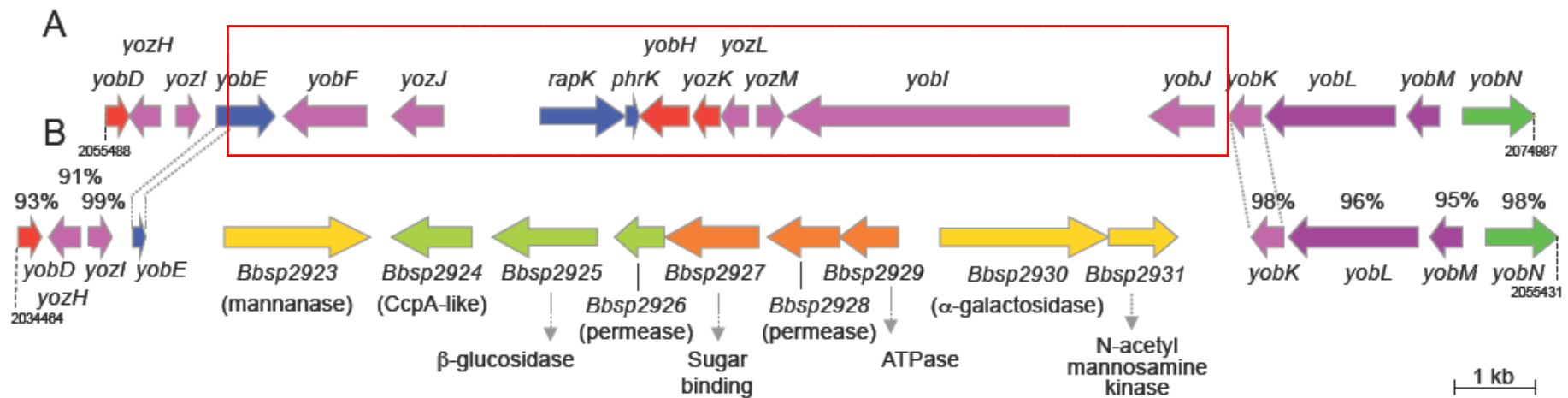


Figure 1.4 – *B. subtilis* 168 chromosome region (A) replaced by *gam* cluster in *B. subtilis* BSP1 strain (B). Genome location, putative function of *gam* cluster protein and similarity (%) with maintained genes are depicted. Adopted from Schyns G., Serra C.R., Henriques A.O., *et al.* (unpublished data).

Schyns G., Serra C.R., Henriques A.O. and co-workers (unpublished data) propose that Bbsp2923 is an endohydrolase responsible for randomly cleaving the β -1,4 bonds within galactoglucomannan backbone, releasing oligo-galactoglucomannans. These oligo-saccharides may then be imported by an ABC transporter, unlike the *gmuA-G* system which functions with a PTS importer, as previously referred. The putative ABC transporter consists of a binding protein (Bbsp2927), a channel of two transmembrane proteins (Bbsp2926 and Bbsp2928) and an ATPase (Bbsp2929).

After entry the cell, the oligo-mannans can be further processed by an α -Galactosidase (Bbsp2930), that acts on D-galactose α -1,6 linkages, a β -Glucosidase (Bbsp2925), that cleaves the β -1,4 bonds, removing D-glucose residues from oligo-saccharides non-reducing ends, and a N-acetyl-mannosamine kinase (Bbsp2931), possible esterase needed for the hydrolysis of D-mannose acetyl substitutions. As in *gmuA-G* operon, this system could be under the control of a CcpA-like protein (Bbsp2924).

1.6. Scope of the thesis

Elucidation of microbial physiology and genetic organization is crucial to modulate and control metabolic pathways, leading interesting features discovery towards new biotechnological applications.

The *gam* cluster comprises a putative system for degradation of galactoglucomannan, a major component of softwood hemicellulose.

The current study intends to elucidate the function and genetic organization of the *gam* cluster as well as to identify the signals that modulate their expression.

Chapter 2

Materials and Methods

2. Materials and Methods

2.1. Substrates

Konjac Glucomannan (60% mannose, 40% glucose; low viscosity), Carob Galactomannan (78% mannose, 22% galactose; low viscosity), Mannan (1,4- β -D-Mannan; 97% mannose, 3% galactose) and Lupin Galactan (88% galactose, 5% arabinose, 5% galacturonic acid, 1% rhamnose, 1% xylose) were purchased from Megazyme International Ireland, Ltd.

Locust Bean Gum (Galactomannan) from *Ceratonia Siliqua* seeds was acquired from Sigma-Aldrich Co. while Birchwood Xylan (89.3% xylose, 8.3% anhydrouronic acid, 1.4% glucose, 1% arabinose) from Fluka.

All substrates were prepared according to manufacturer's instructions.

2.2. Bacterial strains and growth conditions

The bacterial strains used in this project are listed in Table 2.1.

The strain *E. coli* DH5 α [™] (Gibco-BRL) was used for routine molecular cloning work, while *E. coli* BL21(DE3) pLysS (Studier *et al.*, 1990) and *E. coli* Rosetta[™](DE3) pLysS (Novagen Inc.) were used as hosts for overproduction of recombinant Bbsp2923 enzyme. *E. coli* strains were grown in selective Luria-Bertani (LB) liquid broth (Miller, 1972) and on LA medium - LB solidified with 1.6% (w/v) Agar (Difco). Kanamycin (30 $\mu\text{g}\cdot\text{mL}^{-1}$), Chloramphenicol (25 $\mu\text{g}\cdot\text{mL}^{-1}$) and IPTG (1mM) were added as appropriate.

B. subtilis strains were grown on SF medium (gelatin peptone 7.5 $\text{g}\cdot\text{L}^{-1}$; tryptone 7.5 $\text{g}\cdot\text{L}^{-1}$; sodium chloride 5 $\text{g}\cdot\text{L}^{-1}$; pH7.4). Medium was inoculated with overnight grown culture from a freshly streaked colony, to an initial OD_{600nm} of 0.05. When OD_{600nm} reached 0.15-0.18 (early exponential growth phase - t_0), *B. subtilis* BSP1 and 168 cultures were washed and resuspended in the 20mL of SF medium + Konjac glucomannan or carob galactomannan 0.5% medium. Samples of cell culture were collected 2h (exponential growth phase - t_2) and 4h (late exponential growth phase - t_4) after resuspension. Samples were centrifuged 6min at 13000rpm. Pellets and supernatants stored separately at -80°C. Samples of cultures grown in SF medium without mannans addition were also collected.

Cultures were grown on an Aquatron® Waterbath Rotary Shaker at 37°C and 180rpm. Cellular growth was followed by OD_{600nm} periodically measurements in an Ultraspec™ 2100 pro UV/Visible Spectrophotometer (GE Healthcare Life Sciences).

Table 2.1 – List of bacterial strains used in this work.

Strain	Genotype	Source or Reference
<i>Bacillus</i> spp.		
<i>B. amyloliquefaciens</i> FZB42	Wild-type	Chen <i>et al.</i> , 2007
<i>B. subtilis</i> BSP1 (#200)	Wild-type	Barbosa <i>et al.</i> , 2005
<i>B. subtilis</i> 168	<i>trp2c</i>	Spizizen, 1958
<i>B. subtilis</i> 168T⁺	Prototroph	F. E. Young
<i>B. licheniformis</i> DSM13	Wild-type	Veith <i>et al.</i> , 2004
<i>Escherichia coli</i>		
<i>E. coli</i> DH5α™	F ⁻ Ø80 <i>lacZ</i> ΔM15 Δ(<i>lacZYA</i> ⁻ <i>argF</i>) U169 <i>recA1 endA1 hsdR17</i> (<i>r_K</i> ⁻ , <i>m_K</i> ⁺) <i>phoA supE44 λ thi-1 gyrA96 relA1</i>	Gibco-BRL
<i>E. coli</i> BL21 (DE3) pLysS	F ⁻ <i>ompT hsdS_B</i> (<i>r_B</i> ⁻ <i>m_B</i> ⁻) <i>gal dcm</i> (DE3) pLysS (Cam ^R)	Studier <i>et al.</i> , 1990
<i>E. coli</i> Rosetta™ (DE3) pLysS	F ⁻ <i>ompT hsdS_B</i> (<i>r_B</i> ⁻ <i>m_B</i> ⁻) <i>gal dcm</i> (DE3) pLysSRARE (Cam ^R)	Novagen Inc.

2.3. DNA manipulation and sequencing

Regular DNA manipulations were carried out as described by Sambrook *et al.*, 1989. PCR amplifications were performed with Phusion® High-Fidelity DNA Polymerase (Finnzymes) or NZYDNACchange Polymerase (NZYTech, Lda.). Designed oligonucleotides (Table 2.2) were purchased from Metabion International AG or StabVida Lda. DNA from agarose gels and PCR products were purified using GFX® PCR DNA and Gel Band Purification Kit (GE Healthcare).

All restriction enzymes were purchased from MBI Fermentas or New England Biolabs and used according to manufacturer's instructions. The digested pET30a(+) (Novagen) vector DNA was dephosphorilated with an alkaline phosphatase (FastAP™, Fermentas). DNA ligations were performed using T4 DNA Ligase (MBI Fermentas).

Plasmid DNA was purified with the QIAGEN® Plasmid Midi Kit (Qiagen) or QIAprep® Spin Miniprep Kit (Qiagen). DNA samples were quantified using NanoDrop 2000c™ (Thermo Fisher Scientific Inc.). DNA sequencing reactions were performed by StabVida, Lda.

All amplification products were visualized under ultraviolet (UV) light (ChemiDoc™ XRS Gel Documentation system) in a 0.8-1% (w/v) electrophoresis grade agarose gel, buffered with

1xTAE, and stained with GreenSafe Premium (NZYTech, Lda.). NZYDNA Ladder III (NZYTech, Lda.) was the used molecular marker.

Table 2.2 – List of oligonucleotides used in this project.

Oligonucleotide ^a	Sequence (5'→3') ^b	Complementary Sequence
ARA608 – Fwd	CGCTAGCTTTAGGGATTCTCC	<i>Bbsp2927</i>
ARA609 – Rev	AACTAAGTCTGGCTGGTCACC	<i>Bbsp2927</i>
ARA610 – Fwd	AACAGCCAAGGACTCTTCC	<i>Bbsp2923</i>
ARA611 – Rev	CATCAAGACTGAGCGTATCCC	<i>Bbsp2923</i>
ARA612 – Fwd	GAGGAACT <u>CATATG</u> CTGAAAAAAGG	Used to construct pVC1. <i>NdeI</i> restriction site.
ARA613 – Rev	TATATATATGTGC <u>CTCG</u> AGGCTGGCC	Used to construct pVC1 and pVC2. <i>XhoI</i> restriction site.
ARA664 – Fwd	GCGCGCCCAG <u>CATATG</u> ATTATTAATCTTG	Used to construct pVC2. <i>NdeI</i> restriction site.

^a (Fwd) forward primer and (Rev) reverse primer.

^b Restriction sites in the primer sequence are underlined.

2.4. Plasmids construction

The coding sequences of *Bbsp2923*, with and without the signal peptide (see coding sequence in Appendix 6.3) were amplified by PCR using BSP1 chromosomal DNA as template. A nested PCR amplification step was used to increase obtained amplicon yield. The list of used primers is available in Table 2.2. These designed primers introduced unique restriction sites *NdeI* and *XhoI*, at the 5' and 3' end, respectively.

Table 2.3 summarizes the plasmids used or constructed during the course of this work. To construct pVC1, a 2089bp DNA fragment obtained by PCR using primers ARA612 and ARA613, was digested with *NdeI* and *XhoI* restriction enzymes and cloned between the same sites in pET30a(+). To construct pVC2, a 2035bp DNA fragment without the coding sequence of the signal peptide obtained by PCR using primers ARA664 and ARA613, was digested using the same restriction enzymes and also cloned in pET30a(+). Both plasmids, pVC1 and pVC2, encode a recombinant putative mannanase bearing a C-terminal His₆-tag, under the control of a T7 phage promoter, inducible by IPTG. The inserts' sequences in pVC1 and pVC2 were confirmed by DNA sequencing. The map of both plasmids is illustrated in appendices 6.4 and 6.5, respectively.

Table 2.3 – List of used or constructed plasmids.

Plasmid	Relevant construction ^a	Source
pET30a(+)	Expression vector under T7 promoter control. Allows N- or C-terminal His ₆ -tag insertion. <i>Kan</i> .	Novagen
pVC1	pET30a(+) containing the complete Bbsp2923 coding sequence in the MCS. <i>Kan</i> .	This work
pVC2	pET30a(+) containing Bbsp2923 coding sequence without the signal peptide, in the MCS. <i>Kan</i> .	This work

^a (*kan*) Kanamycin resistant. (MCS) Multiple cloning site.

2.5. Total RNA extraction and Reverse transcription-PCR analysis

B. subtilis BSP1 was grown as described above. Cell samples of 1mL were harvest 2h (t_2) and 4h (t_4) after growing in the absence or presence of different types of mannans - konjac glucomannan or carob galactomannan. Total RNA was extracted using Absolutely RNA™ Miniprep kit (Stratagene – Agilent Technologies, Inc) according to manufacturer’s protocol for RNA isolation from Gram-positive bacteria. RNA samples were quantified using NanoDrop 2000c™ (Thermo Fisher Scientific Inc.).

For specific mRNA detection, a two-step reverse transcription polymerase chain reaction (RT-PCR) was used. In this procedure, first it is used a reverse transcriptase with a specific reverse primer, for generation of the cDNA first stand. The second step consists in a regular PCR using both specific oligonucleotides for the synthesis and amplification of the double stranded cDNA. In the first step, M-MuLV Reverse Transcriptase (MBI Fermentas), an RNA- and DNA-dependent DNA polymerase from Moloney Murine Leukemia Virus, was used according to the manufacturer’s instructions, with the exceptions: incubation at 37°C during 2h. PCR amplifications were performed with Phusion® High-Fidelity DNA Polymerase (Finnzymes). Used primers are listed in table 2.4.

All amplification products were visualized as described above (section 2.3).

Table 2.4 – List of oligonucleotides used for RT-PCR assays.

Oligonucleotide ^a	Sequence (5'→3')	Complementary Sequence
ARA593 – Rev	GTTGCATAAGCATCATTGTCG	<i>Bbsp2924</i>
ARA594 – Rev	CCACGATACATATGTGAATAGC	<i>Bbsp2926</i>
ARA596 – Fwd	GGATCTGATACGTCAAGTGG	<i>Bbsp2927</i>
ARA598 – Fwd	AAAGCTTGGTGTAAAGCAGCG	<i>Bbsp2924</i>
ARA607 – Rev	TTCCACCAGTAGTAGCGATACC	<i>Bbsp2928</i>
ARA648 – Fwd	GCCTACCGCACAAGCATTTC	<i>Bbsp2925</i>
ARA649 – Rev	CTTCATCTCACCTTGAGAGC	<i>Bbsp2925</i>
ARA659 – Fwd	GTCTCTGTGCCCACTGTTGC	<i>Bbsp2929</i>

^a (Fwd) forward primer and (Rev) reverse primer.

2.6. Qualitative and comparative analysis of substrate hydrolysis

For qualitative and comparative analysis of the hemicellulose hydrolytic capabilities of the different *Bacillus* strains (Table 2.1), solid medium plates were prepared using 34 g.mL⁻¹ of SFA (Sugar Free Agar) (LabM) and 10 g.mL⁻¹ of the tested substrate, LBG (Galactomannan) or Birchwood Xylan. Strains were grown in liquid SF medium as described in section 2.2. After growing during 7 hours, cells were collected by centrifugation (6min, 6000rpm) and resuspended in new medium to a final OD_{600nm} ≈ 0.8-0.9. Three blank paper discs, inoculated with 20µL of different strains growth, were placed on each plate. Duplicates of plates were prepared and incubated at room temperature during 4 days. Hydrolysis was recorded as the clear halo surrounding the paper discs. Pictures were taken after 48h and 96h, using ChemiDoc™ XRS Gel Documentation system and Quantity One® software (Bio-Rad).

2.7. Hydrolytic activity assays

After hydrolysis of the tested polysaccharides, the reducing sugar content was determined by the Nelson-Somogyi method (Somogyi, 1952), using D-glucose as standard (Appendix 6.2).

B. subtilis strains were grown in liquid medium as described in section 2.2. Supernatants collected in t_4 (4h after cell growth in the absence or presence of different types of mannans) were filtered (0.45µm filter) and used as enzyme source. The reducing sugar content was assayed in a reaction mixture containing 0.5% (w/v) of each tested polysaccharide (Konjac Glucomannan, Carob Galactomannan or Birchwood Xylan) in 200mM phosphate/100mM citrate buffer (PC buffer, pH6,6) and 25µL of supernatant, with incubation periods of 0–25

minutes at 37°C. Samples were prepared in triplicates. The hydrolytic activity was determined measuring the initial rates obtained from the linear portion of the progress curve and was expressed as the amount of glucose reducing-sugar equivalents (μg) produced per minute of reaction, per mL of supernatant and per $\text{OD}_{600\text{nm}}$ of cell culture.

2.8. Small scale overproduction of recombinant Bbsp2923

Induction tests and small scale growth of *E. coli* BL21 (DE3) pLysS and *E. coli* Rosetta™ (DE3) pLysS cells harboring pVC1 and pVC2 were performed to assess the production and solubility of the recombinant proteins. Cells were grown at 37°C and 180rpm in 10mL of selective LB medium. Half of the culture was induced by IPTG 100mM, when culture $\text{OD}_{600\text{nm}}$ reached 0.6. Both the induced and non-induced cultures were grown for further 3 hours, after which the cells were harvested by centrifugation at 13000rpm for 5 minutes. Cells were resuspended in 100 μL of French Press Buffer (Sodium phosphate buffer 20mM pH7.4, NaCl 500mM, Imidazole 100mM and Glycerol 10%(v/v)). For cells disruption, incubation with lysozyme (1mg.mL⁻¹) for 10 minutes at 37°C, and three cycles of freezing in liquid nitrogen and thawing for 5 minutes at 37°C were performed, followed by incubation with 1 μL of Benzonase Nuclease (Novagen®) and PMSF (10mg.mL⁻¹), a serine protease inhibitor, at 37°C for 10 minutes. After centrifugation at 13000rpm and 4°C for 15 minutes, the soluble fraction (supernatant) was recovered and the insoluble fraction (pellet) was resuspended in 100 μL of French Press Buffer.

When protein solubility was not observed, a small-scale auto-induction for the overproduction of Bbsp2923 recombinant proteins was performed based on the method described by Studier, 2005. *E. coli* BL21 (DE3) pLysS and *E. coli* Rosetta™ (DE3) pLysS cells harboring pVC1 were incubated for 16 hours, at 37°C and 150rpm, in 10mL of LB selective medium, with the addition of 1mM MgSO_4 , 1xNPS and 1x5052 (Studier, 2005). Soluble and insoluble fractions of the protein extracts were obtained as described above.

2.9. Protein analysis and Quantification

The analysis of the presence and molecular mass of the overproduced enzymes, in different tested fractions, were conducted by electrophoresis on 12.5% polyacrylamide gels containing sodium dodecyl sulphate (SDS-PAGE), stained with Coomassie Blue. Samples were prepared in 10x loading buffer solution (Tris-HCl 0.5M, pH6.8; SDS 20%; bromophenol blue 0.2%; β -mercaptonethanol 1M; glycerol 20%). NZYTech LMW Protein Marker (NZYTech, Lda.) was the protein marker used.

Supernatants concentration was determined using Bradford reagent (Bio-Rad Laboratories Inc.) with bovine serum albumin (BSA) as standard.

2.10. β -Mannanase activity assay

To assess the putative β -mannanase activity in different substrates, the reducing sugar content after hydrolysis of the polysaccharides was determined by the Nelson-Somogyi method (Somogyi, 1952), using D-glucose as standard (Appendix 6.2). Filtered supernatants (0.45 μ m filter) of *E. coli* Rosetta™ (DE3) pLysS cells harboring pVC2 were used as enzyme source. *E. coli* Rosetta™ (DE3) pLysS strain carrying pET30a(+) without any insert, was used as negative control. The enzyme activity was assayed in a reaction mixture containing 0.5%(w/v) of each tested polysaccharide (1,4- β -D-mannan, carob galactomannan, konjac glucomannan or lupin galactan) in 200mM phosphate/100mM citrate buffer (PC buffer, pH6.6) and 20 μ g.mL⁻¹ of total protein, with incubation periods of 0-20 minutes at 37°C. The enzyme activity was determined measuring the initial rates obtained from the linear portion of the progress curve and was expressed as the amount of glucose reducing-sugar equivalents (μ mol) produced per minute of reaction and per mg of protein.

Chapter 3

Results and Discussion

3. Results and Discussion

3.1. Genetic Analysis of the *gam* Cluster

According to gene orientation, the *gam* cluster appears to be organized in at least three operons – *Bbsp2923*, *Bbsp2924-Bbsp2929* and *Bbsp2930-Bbsp2931* (Fig. 1.4). To address this question an evaluation of the genetic organization at the transcriptional level combined with *in silico* analysis of the *gam* cluster DNA sequence was performed.

3.1.1. Transcriptional organization of the *gam* cluster

Based on the organization of the open reading frames of the *gam* cluster (Fig. 1.4), gene *Bbsp2923* appears to be monocistronic and *Bbsp2930-Bsp2931* may constitute a bicistronic transcriptional unit. The cluster of genes *Bbsp2924-Bbsp2929* suggests they are organized as a single transcriptional unit comprising six cistrons. To elucidate the transcriptional organization of *gam* cluster *Bbsp2924-Bbsp2929* genes, a two-step reverse transcription polymerase chain reaction (RT-PCR) approach was used.

In this method, firstly it is used a reverse transcriptase using a specific reverse primer, for generation of the cDNA first stand. Secondly, it is performed a regular PCR using both specific oligonucleotides for the synthesis and amplification of the double stranded cDNA. This RNA- and DNA- polymerase dependency makes imperative to ensure that RNA is DNA free, in order to avoid false positives. Specific primers were designed to target all possible mRNAs molecules resulting from transcription of this cluster. Prediction of RNA secondary structure formation that could prevent cDNA amplification was taken into account for primers design. RT-PCR positive (chromosomal DNA as template) and negative controls (RNA and water as different templates) were performed in parallel to warrant results fidelity. RT-PCR is a very sensitive technique, where genomic DNA contaminations can be specially problematic due to the absence of introns in bacterial genomes (Sambrook & Russell, 2001). So, highly pure RNA is required. PCR amplifications were firstly performed with NZYDNAChange Polymerase (NZYTech, Lda.), however this enzyme wasn't able to detect small DNA contaminations, contributing in this manner for false positives (results not shown). Thus, Phusion® High-Fidelity DNA Polymerase (Finnzymes) a more accurate polymerase enzyme was used. In a Gram-positive bacteria, like *B. subtilis* BSP1, RNA extraction can be a challenge not only because of the genomic DNA contaminations but also due to the composition of the cell wall that is

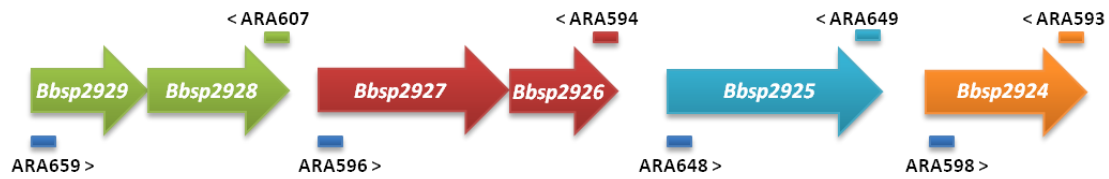
difficult to disrupt (Sambrook & Russell, 2001). Total RNA was obtained using Absolutely RNA™ Miniprep kit, although other approach for total RNA extraction was attempted, using TRIzol® reagent (INVITROGEN™, Life Technologies Ltd) combined with RNeasy Mini kit clean up (QIAGEN Inc.). Even with protocol alterations and longer periods of incubation with DNase, RNA kept on being DNA contaminated or degraded. Only Absolutely RNA™ Miniprep kit allowed the isolation of high-purity RNA. This kit procedure is based on the use of two different spin cups, a prefilter spin cup and a RNA binding cup, which, speeds up RNA extraction and diminishes exposure period to contaminants.

Total RNA was purified and tested from *B. subtilis* BSP1 grown in several conditions: cell samples harvested 2h (exponential growth phase - t_2) and 4h (late exponential growth phase - t_4) in the absence or presence of different types of mannans - konjac glucomannan or carob galactomannan (Fig. 3.1). The goal of the experiment was to determine the transcriptional organization of the cluster, but also analyse the response at the transcriptional level to the presence of different heteromannans and distinct growth stages. The obtained RT-PCR results are illustrated in Fig. 3.1.

The RT-PCR results illustrated in figure 3.1, show a mRNA amplification pattern that is identical for RNA extracted from cells grown in all tested conditions (data not shown). No differences were detected in the intensity of the amplicons bands obtained when cells were grown in the absence of sugars or in the presence of the heteromannans (glucomannan or galactomannan). Similarly, identical patterns and amplicon band intensity were observed with RNA extracted from cells during exponential growth and early stationary growth phase.

Transcriptional organization of the region *Bbsp2924-2929* of *gam* cluster is illustrated in Fig. 3.2. The results indicated that the cluster *Bsp2924-2929* is not organized as an operon. Genes *Bbsp2926-Bbsp2927* appears to be co-transcribed, and likewise *Bbsp2928-Bbsp2929* constitutes a bicistronic transcriptional unit. On the other hand, *Bbsp2924* and *Bbsp2925* are monocistronic. Transcription of all genes comprised in the cluster *Bsp2924-2929* occurs in the absence of sugars.

A.



B.

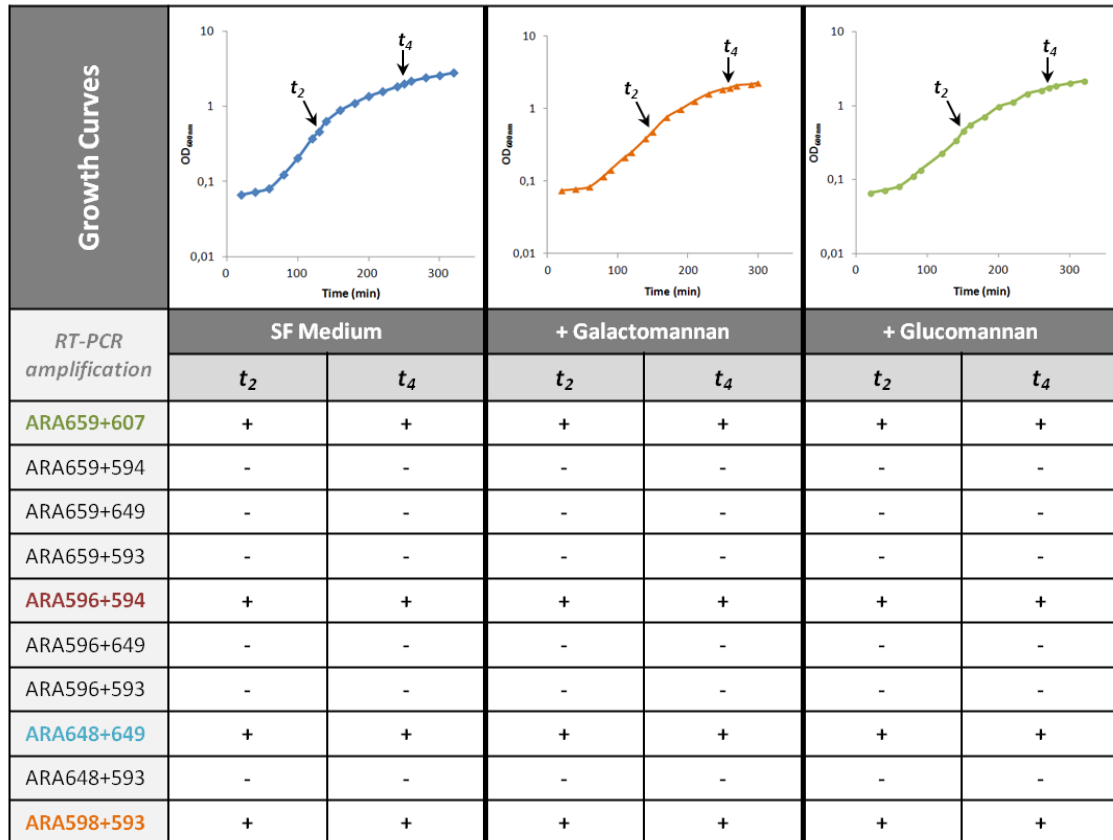


Figure 3.1 – RT-PCR results obtained with RNA extracted from cells grown in different conditions and using distinct pairs of primers. (A) The position of each primer used in the experiments is indicated above (reverse primer) or below (forward primer) the genes. (B) Cells were grown in SF medium in the absence of sugars (SF Medium) in the presence of galactomannan (+Galactomannan), or in the presence of glucomannan (+Glucomannan) (top panel from left to right, respectively). Arrows indicated the time at different growth phases where cells were harvested and RNA extracted (t_2) exponential growth phase and (t_4) late exponential-early stationary growth phase. The pairs of primers used in RT-PCR experiments are indicated in the left column. (+) indicates positive amplification and (-) no detection of product/amplification.

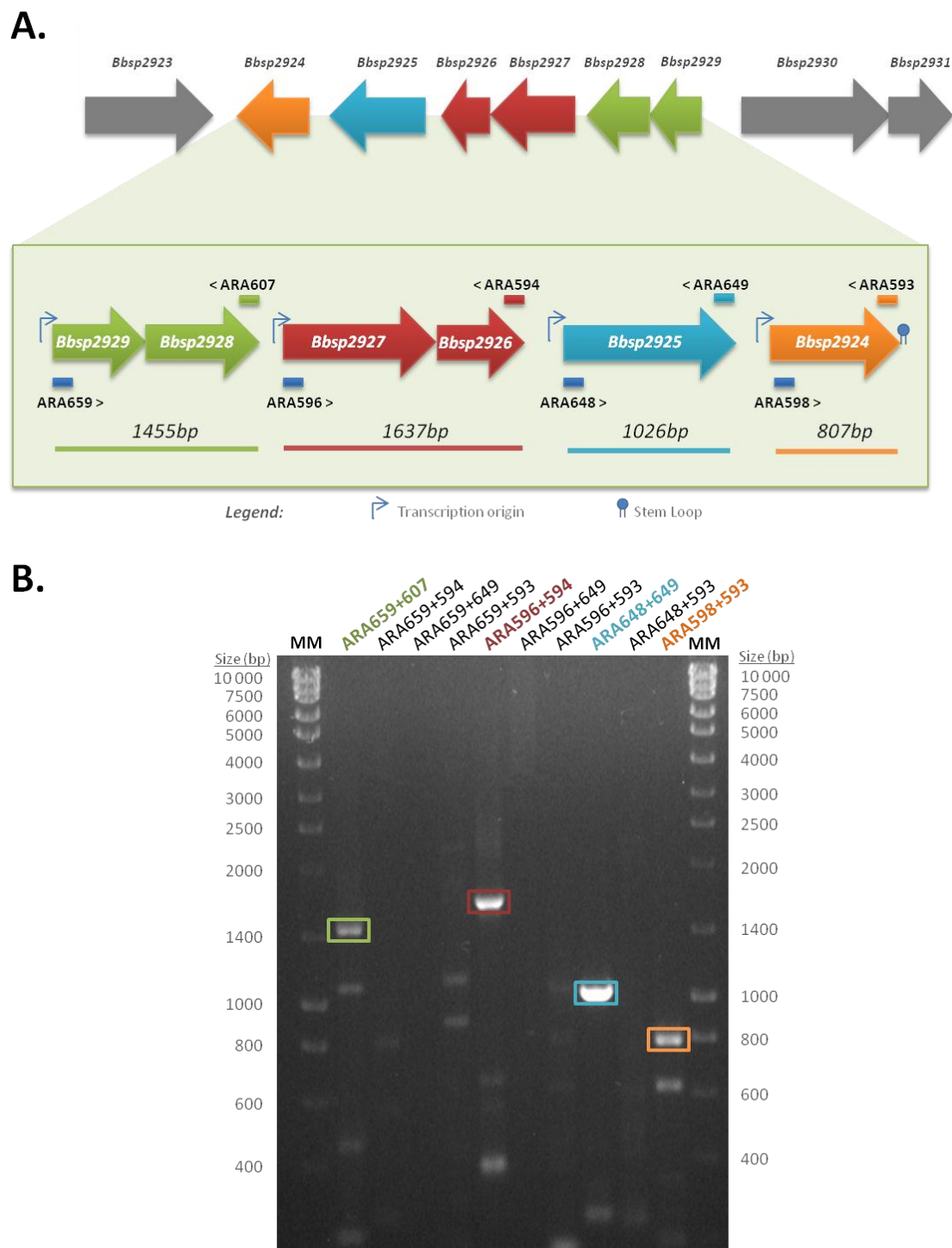


Figure 3.2 – Transcriptional organization of the region *Bbsp2929-2924* of *gam* cluster. (A) The transcripts detected by RT-PCR, and respective size of the amplicons are depicted below the cluster of genes. The position of each primer used in the experiments is indicated above (reverse primer) or below (forward primer) the genes. (B) Example of an obtained RT-PCR amplification pattern. In this example amplification products were obtained using RNA extracted from cells in exponential growth phase (t_2) grown in the presence of Konjac glucomannan. In each lane are indicated the pair of primers used and their sequence is available in Table 2.4. Amplification products were visualized under ultraviolet light in a 1% (w/v) electrophoresis agarose gel, buffered with 1xTAE and stained with GreenSafe Premium (NZYTech, Lda.). NZYDNA Ladder III (NZYTech, Lda.) was the used molecular marker (MM). Band size was determined using ChemiDoc™ XRS Gel Documentation system and Quantity One® software (Bio-Rad).

3.1.2. Sequence Analysis of the *gam* cluster

As described in previous sections, Schyns G., Serra C.R., Henriques, A.O. and co-workers (unpublished data) reported that the *gam* cluster encodes all the necessary elements for the metabolization of the galactoglucomannan, including a β -Mannanase (Bbsp2923), an intracellular α -Galactosidase (Bbsp2930), a β -Glucosidase (Bbsp2925), an N-acetylmannosamine kinase (Bbsp2931), an ABC transportation system (Bbsp2926-29) and a CcpA-like regulatory protein (Bbsp2924).

An update of the predictions was performed in the databases, ORFs forecast was confirmed using DNA Strider version 1.4f6 (Douglas, 1995) and the correspondent gene products compared against available databases using BLASTp search (Basic local alignment search tool using a protein query) (Altschul *et al.*, 1990). Ribosome binding sites (RBS) location and translation initiation were manually defined by complementarity with *B. subtilis* 16S rRNA sequence and canonical distance observed between RBS and start codon in bacteria [Snyder & Champess (1997); Hager & Robinowitz (1985)]. The putative location of RBS suggests start codons for *Bbsp2923*, *Bbsp2926*, *Bbsp2927* and *Bbsp2929*, which are distinct from the ones annotated in the available sequence of the *gam* cluster (Schyns, G., Serra C.R., Henriques, A.O. *et al.*, unpublished results). Table 3.1 summarizes the results.

Two distinct DNA regions of genes *Bbsp2923* and *Bbsp2927* were resequenced due to possible errors detected in the available nucleotide sequence. The primer pair ARA608-609 was design to amplify and sequence *Bbsp2927* and primers ARA610-611 to gene *Bbsp2923*. In the DNA region of *Bbsp2923* no mistakes were detected, however the DNA sequence of *Bbsp2927* presented a wrongly annotated additional base (T) in a thymine rich zone leading to a frame shift. The corrected sequence is presented in appendix 6.1.

The update of the BLAST search for proteins that are homologous to those deduced from the nucleotide sequence, confirms that the *gam* cluster encodes a complete putative degradation system for galactoglucomannan, including degradation, uptake and catabolism.

Bbsp2923 encodes a putative mannanase similar to *Bacillus* sp. JAMB750 extracellular β -mannanase, responsible for random cleavage of endo- β -1,4 mannosidic bonds that constitute mannans main chain (Dhawan & Kaur, 2007). Bbsp2923 presents a potential signal peptide suggesting that might be secreted by the Sec protein translocation machinery (Yamane *et al.*, 2004). Function, cloning and expression of Bbsp2923 will be addressed on section 3.3.

Table 3.1 – Predicted signals for translation initiation in *gam* cluster sequence.

<i>Gene</i>	<i>Orientation</i> ^a	<i>RBS and start codon (5'-3')</i> ^b	ΔG ^c (kcal/mol)
<i>Bbsp2923</i>	+	AAGGAGGAACTGGTTG	-17.8
<i>Bbsp2924</i>	-	AGGAGCGAGGATTTG	-11.6
<i>Bbsp2925</i>	-	AAGGAGTGCAAAGAAATG	-12.8
<i>Bbsp2926</i>	-	AAGGTGAAACGGATG	-9.6
<i>Bbsp2927</i>	-	AGGGAGTAATTCGATG	-9.8
<i>Bbsp2928</i>	-	GGAGGGAAAAGTGGAGTG	-14.4
<i>Bbsp2929</i>	-	AAGCGGGGTGAAAATG	-15.8
<i>Bbsp2930</i>	+	AAGGAGTGGTGAATAGTAATG	-12.8
<i>Bbsp2931</i>	+	AAGGGGGATACCTATG	-15.8
<i>B. subtilis</i> 16S rRNA consensus sequence ^d		3' UUCUCCA – N ₅₋₁₀ – <i>Start codon</i>	

^a Genomic orientation. (+) Same sense. (-) Anti-sense.

^b Consensus elements in RBS are in boldface and translation start site is underlined.

^c Free energy of binding for the interaction between the 3' end of the *B. subtilis* 16S rRNA sequence and the putative RBS, calculated by the method of Tinoco *et al.* (1973).

^d *B. subtilis* 16S rRNA consensus sequence adapted from Snyder & Champess, (1997) and Hager & Robinowitz (1985). N denotes any nucleotide.

Bbsp2925 is 72% identical to a β -glucosidase from *Clostridium* sp. DL-VIII (Table 3.2). β -glucosidases are necessary for the hydrolysis of the β -1,4 linkage between D-glucose and D-mannose residues that constitute galactoglucomannan backbone (Shallom & Shoham, 2003).

Bbsp2930 is most probably an α -galactosidase 76% identical to one found in *Geobacillus stearothermophilus* (Table 3.2). α -galactosidases cleave the α -1,6 bond linking galactose with galactose, glucose or mannose residues (Shallom & Shoham, 2003).

Bbsp2931 might code for a sugar kinase because it shares homology to proteins of the ROK (Repressor, Open reading frame, Kinase) superfamily (Larion *et al.*, 2007). This enzyme uses ATP to phosphorylate the target compound and could play a role in entrapping oligosaccharides in cell cytosol or helping solubilise them. However, ROK superfamily encompasses others mischaracterized proteins, and according to Moreira & Filho (2008) for the degradation of a galactoglucomannan polymer, an N-acetyl-mannosamine esterase would be necessary to hydrolyze the acetyl substitutions and not a kinase. Thus, the determination of the exact function of *Bbsp2931* needs more studies.

Table 3.2 – Comparison of *gam* cluster products to other proteins in available databases using BLASTp tool.

Gene	Size [aa (kDa)]	Homolog protein ^a	Database entry	Identity (%)	Amino acid overlap
Bbsp2923	687 (75.5)	Beta-Mannanase [<i>Bacillus</i> sp. JAMB750]	AB128831	57	486
Bbsp2924	339 (38.4)	LacI family transcriptional regulator [<i>Bacillus cellulosilyticus</i> DSM 2522]	ADU29634.1	64	332
Bbsp2925	454 (52.1)	6-phosphobeta-glucosidase [<i>Clostridium</i> sp. DL-VIII]	EHI97801.1	72	449
Bbsp2926	205 (23.9)	unspecified sugar ABC transport ATP-binding protein [<i>Mycoplasma canis</i> UFG4]	EIE41424.1	31	109
Bbsp2927	422 (46.4)	sugar ABC transporter substrate- binding protein [<i>Bacillus clausii</i> KSM-K16]	BAD62861.1	45	418
Bbsp2928	293 (32.9)	sugar ABC transporter permease [<i>Bacillus clausii</i> KSM-K16]	BAD62860.1	67	273
Bbsp2929	294 (33.4)	sugar ABC transporter permease [<i>Bacillus clausii</i> KSM-K16]	BAD62859.1	73	295
Bbsp2930	746 (86.5)	alpha-galactosidase [<i>Geobacillus stearothermophilus</i>]	AAF70204.1	76	742
Bbsp2931	322 (35.0)	ROK family protein [<i>Geobacillus</i> sp. C56-T3]	ADI25265.1	57	301

^a Homolog protein found in database search and respective strain indicated between square brackets. Search performed with default settings.

The Bbsp2926 to Bbsp2929 proteins present homology to all elements necessary to assemble an ABC-type transporter system. ABC or ATP-binding cassette transporters are ubiquitous membrane proteins that couple the transport of several substrates across the cellular membrane to the hydrolysis of ATP (Hollenstein *et al.*, 2007). This system is organized in two transmembranar domains (TMDs), which provide the translocation pathway and two cytoplasmic nucleotide-binding domains (NBDs) that bind and hydrolyze ATP molecules. In bacterial importers an additional external binding protein (substrate binding protein, SPB) that delivers captured substrates to the transporter is required. The ABC domains are generally encoded as separate polypeptides, however they may be fused together in any of the possible combinations (Higgins, 1992). In the *gam* cluster, *Bbsp2927* encodes a solute-binding protein (SBP), *Bbsp2926* an ATPase NBD and both *Bbsp2929* and *Bbsp2928* integral membrane proteins TMD.

Bbsp2924 most probably encodes a transcription factor belonging to the LacI family of transcriptional regulators (Weickert & Adhya, 1992).

The galactoglucomannan metabolization system encoded by *gam* cluster appears to be under the control of the regulatory mechanism of carbon catabolite repression (CCR) (Deutscher *et al.*, 2002; Fujita, 2009). CCR is a common type of regulation for the use of alternative carbon sources, such as arabinose (Inácio *et al.*, 2003). *In silico* analysis revealed the presence of several catabolic responsive elements (CREs) along the *gam* cluster sequence (Table 3.3). CcpA, a master regulator of CCR in gram-positive bacteria, functions as a repressor that in the presence of glucose binds to DNA operator sequences (CRE) usually located in the promoter region of a target gene, preventing or blocking transcription (Deutscher *et al.*, 2002; Fujita, 2009). As indicated in Table 3.3, CREs were found in the promoter region of genes *Bbsp2923*, *Bbsp2929* and *Bbsp2930* (Fig. 3.3).

Table 3.3 – Identified *cre* sequences and location.

Gene	<i>cre</i> sequence ^a	Distance to TSS (bp) ^b
<i>Bbsp2923</i>	TGAAGCGCTTCA	42
<i>Bbsp2929</i>	TGTAAGCGTTTTCT	292
<i>Bbsp2930</i>	TGTAAGTGCTTTCT	54
<i>B. subtilis</i> consensus sequence ^c	TGWNANCGNTNWCA	

^a *cre* sequences were identified with DBTBS search tool using a p-value=5% (threshold). Conserved positions are in boldface type.

^b Distance in base pairs to the TSS (Translational start site) of the respective gene.

^c *B. subtilis* *cre* consensus sequence (Deutscher *et al.*, 2002). W= A or T, N= any base.

Transcription initiation within the *gam* cluster was analysed using the DBTBS search tool to identify putative promoter sequences, namely the -35 and -10 promoter regions that interact with sigma (σ)-factors, allowing the RNA polymerase to bind to DNA. DBTBS is a database of transcriptional regulation in *Bacillus subtilis* that contains intergenic conservation information (Sierro *et al.*, 2008). The results are summarized in Table 3.4.

Six putative promoter sites for σ^A binding were identified upstream from *Bbsp2923*, *Bbsp2924*, *Bbsp2925*, *Bbsp2927*, *Bbsp2929*, and *Bbsp2930*. The σ^A is the primary vegetative cell sigma-factor in *B. subtilis*. However, prokaryotic transcriptional regulation is frequently accomplished by alternative σ -factors acting to alter promoter utilization, leading to changes in expression pattern (Helmann & Moran, 2002).

Table 3.4 – Typical *B. subtilis* sigma-factor (σ) binding sequences identified in some *gam* genes promoters.

σ	Gene	Promoter sequences ^a			Distance to TSS (bp) ^b
		-35	Spacer	-10	
σ^A	<i>Bbsp2923</i>	TTTACTTTATTTATTAAGCAACTAAATATATAAT			56
	<i>Bbsp2924</i>	TTGACAAAAGATTTTATTAGCCATAAAAAT			47
	<i>Bbsp2925</i>	TTGAGGTAAGCCTGCAATGTTATTT			28
	<i>Bbsp2927</i>	ATGACCTTTTCTGGAATCTTATCTATATTTT			33
	<i>Bbsp2929</i>	TAGACATTTAATTCTGGTGAATATATAAT			225
	<i>Bbsp2930</i>	TTGTCAAAAAATAAATAGTTTATTTT			68
	Consensus sequence ^c	TTGaca	N ₁₇	TAtaat	
σ^K	<i>Bbsp2923</i>	ACATCAAACCAATTCTCTCCATATGCT			366
	Consensus sequence ^c	AC	N ₁₆₋₁₈	CATANNNT	
σ^D	<i>Bbsp2924</i>	TTAATGATTTGCCATACTCCGCCTATAT			204
	Consensus sequence ^c	TAAA	N ₁₄₋₁₆	gCCGATAT	

^a Putative -35 and -10 promoter sequences identified using DBTBS search tool, using p-value=5% (threshold). Consensus elements are in boldface type.

^b Distance in base pairs, from the -10 promoter sequence beginning to the TSS (Translational start site).

^c Consensus sequences for the putative sigma-factor interaction (Helmann & Moran, 2002). Capital letters indicate highly conserved positions. Lowercase letters indicate less conserved positions. N= any base.

In *B. subtilis* there are secondary σ -factors that bind to RNA polymerase instead of the primary σ -factor, directing the enzyme to recognize alternative promoter sites (Helmann & Moran, 2002). Two alternative sequences for secondary σ -factors recognition, σ^K and σ^D , were identified in *Bbsp2923* and *Bbsp2924* promoter region, respectively. σ^K factors are active in the forespore during sporulation, while σ^D factors are related to flagella, chemotaxis or autolysins (Helmann & Moran, 2002). Chemotaxis is a mechanism used by some organisms to direct their movements towards some chemicals available in the environment (Garrity & Ordal, 1995).

<p>Bbsp2923 promoter region</p> <p>^{-35 (σ^K)} <u>CCTACTGTGCACATCAAACCAAATTCTCTCC</u> ^{-10 (σ^K)} <u>ATATGCTGAAAGCTGTTAGATGCAGTGAATGGCCTTTTC</u></p> <p>GTC AATTTGTCTTTTTTTAAAAAAGGCCTATCTAAAAGCTCGGGAAAGAAATGCTTATCTTTGCTACACT</p> <p>GATCATATTTTCATAAACGTGCTTTATCTCATTAAACAGCCAAGGACACTCTCCCTCGGACATTCTGTG</p> <p>TGTAGAATTTTTACAACCGCCTTCTGTTCAAACAAGAAATTTAGTTTATTACAGAAAACACATCTTTTT</p> <p>AATTTATCGAAAACCTTAAAAATTAATTTTT ^{-35 (σ^A)} <u>TTTACTTTATTTATTAAGCAACTAAATATATAATATAAAAT</u> ^{-10 (σ^A)}</p> <p><i>cre</i> TGAAGCGCTTTCAACATTAATGATTAAGGAGGAACTGGTTT ^{RBS} <u>GCTGAAAAAAGGATGGATTCTTTT</u></p> <p>Bbsp2923 – L L K K G W I S F</p>
<p>Bbsp2924 promoter region</p> <p>ATTTAAAACGTTGATAATTAAT ^{-35 (σ^D)} <u>GATTGGCCATACTCCGCCTATAT</u> ^{-10 (σ^D)} <u>CCATTGTTTATATTAACAAACAGTGAA</u></p> <p>TATAGGCTTTAATTGTTCCGCTCATACTATTGCTGCTGTTGTATTAATAATTTACGTACCAATCTGTGCAC</p> <p>GCTCTCGGGGTGGATTTAGTAAGCCTCATT ^{-35 (σ^A)} <u>TGACAAAAGATTTATTAGCCATAAAATTAAGTTATCGAT</u> ^{-10 (σ^A)}</p> <p>TAACAATAATT ^{RBS} <u>CAGAGAGGAGCGAGGATTTGAAA</u>ACTAATATTACCATGCGAGATATTGCAGAAAAGC</p> <p>Bbsp2924 – L K T N I T M R D I A E K</p>
<p>Bbsp2925 promoter region</p> <p>GTAATACAACATAAGATTTCTTTTTTCGCTTAAACATACGCCTGAAAGAATGTTCT ^{-35 (σ^A)} <u>TGAGGTAAGCCTG</u></p> <p>CAATGTTATTT ^{-10 (σ^A)} <u>CACAATTAAGGAGTGCAAAGAAATGAATAATAAGAATATTATCCCCGAGGGTTTTTTCT</u> ^{RBS}</p> <p>Bbsp2925 – M N N K N I I P E G F F</p>
<p>Bbsp2926 promoter region</p> <p>GATTTCGAAAGAATATAACCTGAAATAATAACGAAAAGGTAAGTACTTGTTAAATCAAGTACCTTATTTCAAG</p> <p>^{RBS} <u>AAGGTGAAACGGATGAATAAAAAGTACGATTTAGACAGCAAGGTTTTCTGGATAATGGATTTGGTCTT</u></p> <p>Bbsp2926 – M N K S T Y L D S K V F W I M D L V</p>
<p>Bbsp2927 promoter region</p> <p>GTATTTCATAGAAGGTATCGCTACTACTGGTGAAAGTAATAGCTAGGTTTGATTAGTAAGGGTCACGC</p> <p>TAGATAAT ^{-35 (σ^A)} <u>GACCTTTCTGGAATCTTATCTATATTTTT</u> ^{-10 (σ^A)} <u>ACTAAATAAAT</u> ^{RBS} <u>AGGGGGAGTAATTCGATGTTTA</u></p> <p>Bbsp2927 – M F</p>

Figure 3.3 – Annotated sequence of the putative promoter regions of *gam* cluster. Nucleotide sequence indicated in 5'-to-3' direction. Each promoter region (-35 and -10) is defined by putative sigma-factors binding sites (corresponding σ factor indicated above underlined sequence); the putative ribosome-binding sites (RBS) are also represented; the identified catabolic repression elements (*cre*) sequences are shaded; putative start codon is underlined; the beginning of the predicted primary structure of each protein is given in a single-letter code below the nucleotide sequence.

<p>Bbsp2928 promoter region</p> <p style="text-align: center;">RBS</p> <p>ACTATGAATATGACGAATAAGTGAGGGAGGGAAAGTGGAGTGGAAATCGCAAAGTAAAAACCAAAG</p> <p style="text-align: center;">Bbsp2928 – V E I A K V K T K</p>
<p>Bbsp2929 promoter region</p> <p style="text-align: center;"><i>cre</i></p> <p>AATTAGAATTAATCAATATTGTAAGCGTTTCTGTTTTATTTTTGAAAATTTCAAAAAGTATAGACAT^{-35 (σ^A)}</p> <p style="text-align: center;">^{-10 (σ^A)}</p> <p>TTAATTCTGGTGAATATATAATTATTTTATTAAGTTATCGATTAACGATTGTTAAGTTAAGTTAATAATT</p> <p>AAGTAACTGATTAGTGATTTAATTTTTTAAAAGGGAGTGAAAAACAAAGAATTTTTGAAAAGTAT</p> <p>CTAAAGGGTTTGAAGTTTTTCGAGAATCCCAGGATATTTGATAAAGTCATGTGGAGTTCCTTATTA</p> <p style="text-align: center;">RBS</p> <p>AATTAGATAAGGCGGGTGGAAAATGCATAAATCGAAAACATTTTTGTTCTACATGTTTCATCTTACCT</p> <p style="text-align: center;">Bbsp2929 – M H K S K T F L F Y M F I L P</p>
<p>Bbsp2930 promoter region</p> <p style="text-align: center;">^{-35 (σ^A)}</p> <p>TTAGTTTAATTTAAGTTTATATTGTCAAAAATAAATAGTTTATTTTGATAATTTTGTATAATTTAAGT^{-10 (σ^A)}</p> <p style="text-align: center;"><i>cre</i></p> <p>CGCTTGCAACTTCTTTTTATTTCTCTATTATTGTAAGTGCTTCTAAAATAAGGGTTTATTATTAGTA</p> <p style="text-align: center;">RBS</p> <p>AGGAGTGGTGAATAGTAATGCCGATTAAGGTCAATGAGGAAACGAAAGAATTCATCTTTCAACCGA</p> <p style="text-align: center;">Bbsp2930 – M P I K V N E E T K E F H L F N N</p>
<p>Bbsp2931 promoter region</p> <p style="text-align: center;">RBS</p> <p>TGCGTTGTA AAAAGGGGATACCTATGAATAATTACATGGTTTTTGATATAGGCGGCAGTGCGATTAA</p> <p style="text-align: center;">Bbsp2931 – M N N Y M V F D I G G S A I</p>

Figure 3.3 (continuation) – Annotated sequence of the putative promoter regions of *gam* cluster genes. See previous page legend.

RNA is synthesized until RNA polymerase recognizes a termination signal. Typically, this signal is a stem-loop structure comprising a CG-rich inverted repeat sequence followed by a short run of U residues (Helmann & Moran, 2002). A single putative stem-loop structure was identified in *gam* cluster. This hairpin (Fig. 3.4) is located downstream of *Bbsp2924* and possesses a free Gibbs-energy (ΔG) of -9.5kcal/mol.

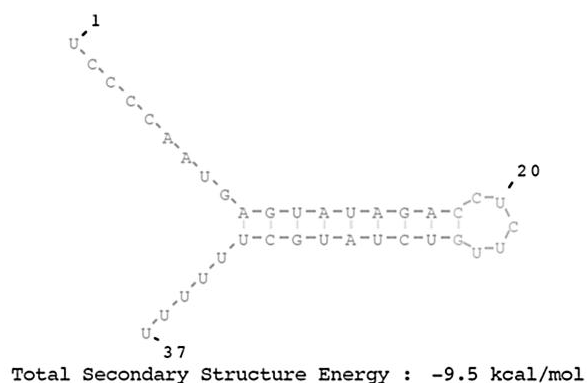


Figure 3.4 – Stem-loop structure of a putative transcription terminator found downstream of *Bbsp2924*. Free Gibbs-energy (ΔG) of hairpin formation is -9.5kcal/mol. ΔG was calculated using DNASIS-Mac v.3.7 software (Hitachi Solutions, Ltd., Tokyo, Japan).

The *in silico* analyses of the *gam* cluster concerning gene orientation, transcription promoters and stem-loop location as well as putative function of encoded proteins are summarized in figure 3.5.

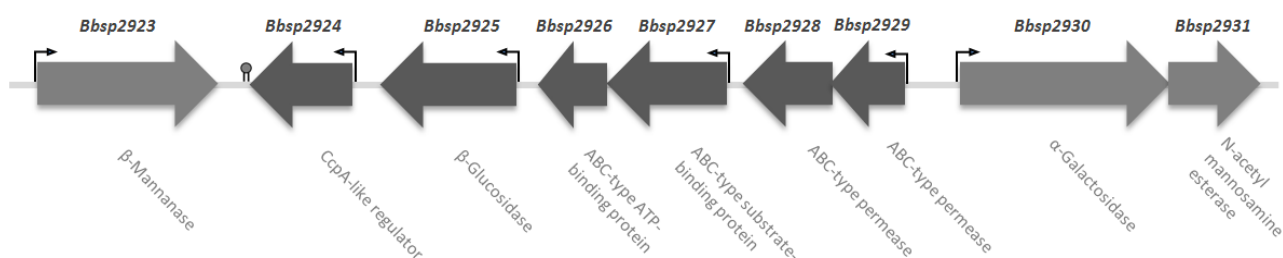


Figure 3.5 – Complete *gam* cluster scheme with identified promoter regions and hairpins. Putative function of the different encoded proteins based on BLAST results is also illustrated.

The bioinformatic promoter predictions are consistent with the transcriptional results obtained by RT-PCR concerning the cluster *Bbsp2924-Bbsp2929* (Fig. 3.2-A), corroborating the monocistronic nature of *Bbsp2924* and *Bbsp2925*, and suggesting that *Bbsp2929-Bbsp2928* and *Bbsp2926-Bbsp2927* constitute two distinct bicistronic operons.

3.2. Functional Characterization of the *gam* Cluster

As previously referred, hemicellulose is one of the most abundant heteropolymers present in nature, usually associated with cellulose and lignin in plant cell walls (Moreira & Filho, 2008). In hemicelluloses, heteroxylans and heteromannans are abundant polysaccharides. After xylanases, mannanases are the most important enzymes for the hydrolysis of hemicelluloses. Mannanases possess many biotechnological applications in different industries and are known to be produced by several *Bacillus* spp. as *B. subtilis*, *B. licheniformis* or *B. amyloliquefaciens* (Chauhan *et al.*, 2012).

In this section, a qualitative analysis of the hemicellulosic capabilities of *B. subtilis* BSP1 was performed and compared to other *Bacillus* spp. In addition, we also attempted to determine the signals that trigger *gam* cluster expression and whether its presence in BSP1 genome results in new hemicellulosic potential, when compared to strain 168.

3.2.1. Qualitative analysis of substrate hydrolysis

For qualitative comparison of hemicellulosic capabilities, between *B. subtilis* BSP1 and other *Bacillus* strains, cells were grown in solid medium containing different substrates and hemicellulolytic potential evaluated by the dimension of the clear halos of hydrolysis surrounding cell growth. It was also tested if the presence of *gam* cluster in BSP1 strain confers a powerful hydrolytic activity against mannan rich substrate.

The chosen substrates were locust bean gum (Fig. 3.6-B), a galactomannan based sugar, and birchwood xylan (Fig. 3.6-C), a xylose rich substrate. In the same plate were compared strains BSP1, 168 and 168T⁺, or BSP1, *B. licheniformis* and *B. amyloliquefaciens*. BSP1 contains the *gam* cluster and *gmuA-G* operon, while 168 and 168T⁺ contain only the *gmuA-G* operon. *B. licheniformis* and *B. amyloliquefaciens* were also tested because of their known ability to produce and secrete a great amount of hydrolytic enzymes (Ferrari *et al.*, 1993).

The results shown in figure 3.6-A indicate that BSP1 possesses the highest hydrolytic activity when compared to other *Bacillus* species and other *B. subtilis* strains. This is observed towards both tested substrates and not only towards mannan rich media. Thus, the higher hydrolytic capability shown by strain BSP1 in the galactomannan rich medium, when compared to strains 168 and 168T⁺, may be due to its natural capacity to produce and secrete hemicellulases and not due to the additional presence of the *gam* cluster in its genome.

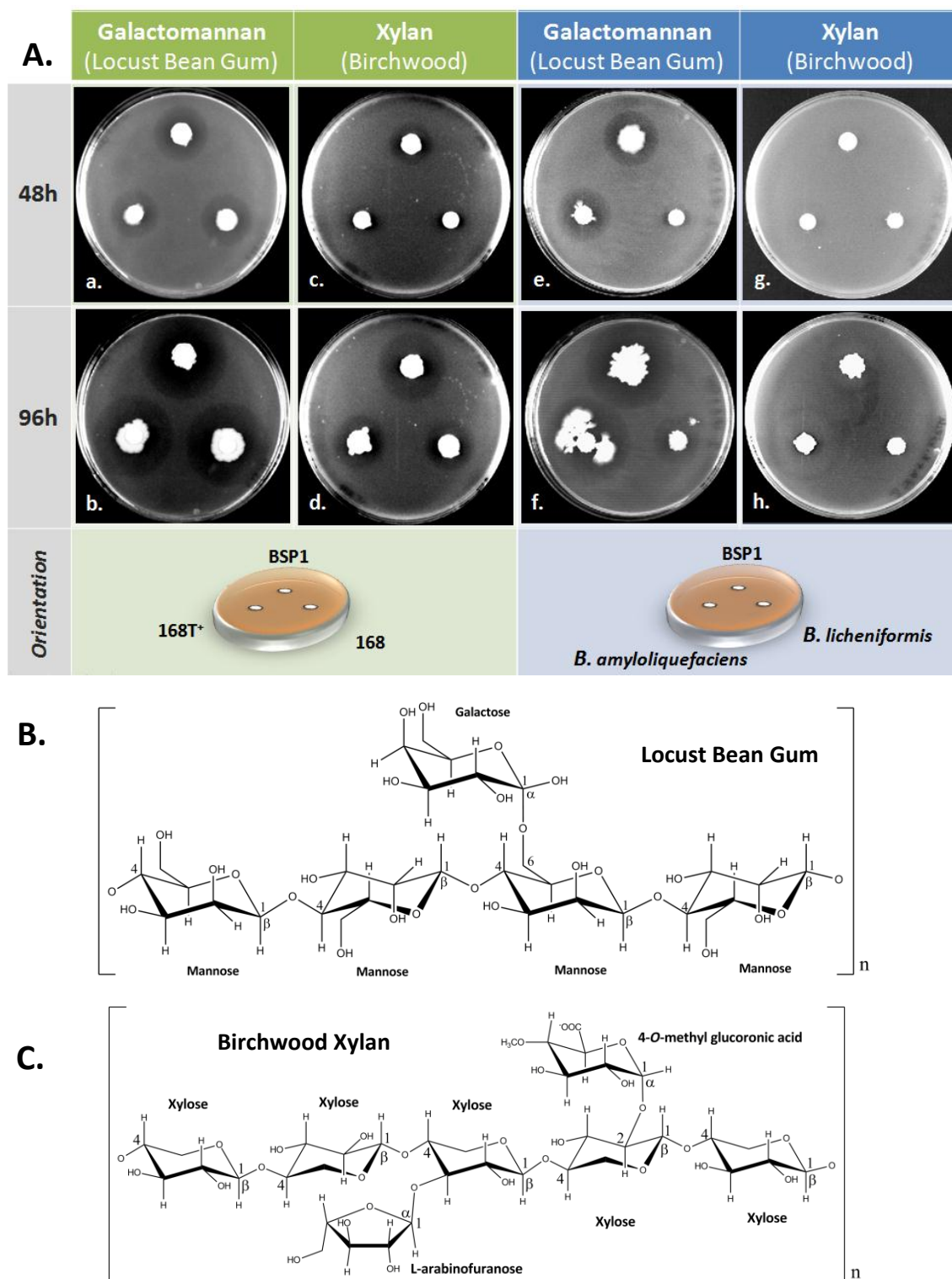


Figure 3.6 – Hydrolysis test on Locust Bean Gum (LBG) and Birchwood Xylan. (A) Solid SFA medium plates supplemented with the indicated polysaccharide. Plates were incubated at room temperature ($\approx 20^{\circ}\text{C}$) during 2 or 4 days. The clear halo indicates hydrolysis of the complex polysaccharide. In the same plate were inoculated for comparison strains BSP1, 168 and 168T⁺ (green) or BSP1, *B. licheniformis* and *B. amyloliquefaciens* (blue). (B) LBG structure: composed by a backbone of β -1,4 linked mannose units with a side-chain α -1,6-galactose residue in every four mannose units. (C) Birchwood xylan structure: formed by a backbone of xylose units linked through β -1,4-glycosidic bonds and 4-O-methyl-D-glucuronic acid linked to the xylose units by α -1,2-glycosidic bonds. In addition, there are L-arabinofuranose residues attached to the main chain through α -1,2 and/or α -1,3-glycosidic linkages.

3.2.2. Hydrolytic activity assays

To identify signals that trigger the *gam* genes expression and whether or not the presence of the *gam* cluster contributes to a higher hemicellulosic capability, the activity of the extracellular hydrolytic enzymes – β -mannanases, produced by both the *gam* cluster and the *gmuA-G* operon were tested. As source of enzymes, supernatants of cell cultures in early stationary phase of *B. subtilis* BSP1 (containing both *gam* cluster and *gmuA-G* operon) and 168 (possessing only *gmuA-G*), grown in the absence of sugars or in the presence of konjac glucomannan or carob galactomannan were used.

The extracellular hydrolytic activity was assayed towards different substrates, konjac glucomannan, carob galactomannan and birchwood xylan, and determined by the Nelson-Somogyi reducing sugars content method (Somogyi, 1952). Konjac glucomannan and carob galactomannan are mannan rich substrates that possess in their chemical structure some of the bonds present in galactoglucomannan, the polysaccharide proposed to be degraded by *gam* cluster, while birchwood xylan is a xylose polymer, used as a negative control in this experiment. Table 3.5 summarizes the obtained data.

Table 3.5 – Extracellular hydrolytic activity of *B. subtilis* BSP1 and 168 strains towards distinct substrates.

Growth ^a	Assay ^b	Activity ($\mu\text{g}\cdot\text{mL}^{-1}\cdot\text{min}^{-1}\cdot\text{OD}_{600\text{nm}}^{-1}$)	
		BSP1	168
SF	Gal	9.59	9.88
SF	Glu	3.63	4.02
SF	Xyl	0.53	0.54
Gal	Gal	42.92	28.74
Gal	Glu	49.43	29.04
Gal	Xyl	3.10	0.80
Glu	Gal	27.50	18.03
Glu	Glu	33.15	13.67
Glu	Xyl	1.62	1.65

^a Supernatants of cultures used as enzyme source were collected at early stationary phase (*t*₄) of cells grown in the absence of sugar (SF) or in the presence of carob galactomannan (Gal) or konjac glucomannan (Glu).

^b Hydrolytic activity of the supernatants was assayed towards the following substrates: carob galactomannan (Gal), konjac glucomannan (Glu) and birchwood xylan (Xyl).

^c The values of enzymatic activity represent the average of triplicates and were determined as described in Materials and Methods section 2.7.

At a first glance the obtained results (Table 3.5), suggest that galactomannan and glucomannan induce the production of β -mannanases in both strains BSP1 (containing both *gam* cluster and *gmuA-G* operon) and 168 (possessing only *gmuA-G*), and that BSP1 possesses a higher mannanolytic activity towards the D-mannan rich substrates, when compared to strain 168. The later could be due to the presence of both operons in BSP1, which resulted in the production of two different β -mannanases, while in 168 only a single β -mannanase is produced. However, an increase in the xylanolytic activity (negative control) of BSP1 grown in the presence of galactomannan when compared to cells grown in the absence of sugars was observed. Furthermore, when some of the experiments were repeated the reproducibility of the data was questionable and this methodology was abandon.

3.3. Cloning and Preliminary Characterization of the Bbsp2923

β -mannanases or endo-1,4- β -mannanases (EC 3.2.1.78, mannan endo-1,4- β -mannosidase), are hydrolases that catalyze the random cleavage of β -1,4 mannosidic linkages within the backbone of different mannans and heteromannans like glucomannans, galactomannans and galactoglucomannans (Dhawan & Kaur, 2007). Mannans are major components of hemicellulose, showing a wide spread distribution in plant tissue (Moreira & Filho, 2008).

Over the past years, β -mannanases isolated from fungi, actinomycetes or bacteria have found many biotechnological applications. Microbial mannanases are mainly extracellular enzymes that can act in wide range of pH and temperature, which have contributed to pulp and paper, pharmaceutical, food and feed, oil and textile successful industrial applications (Chauhan *et al.*, 2012).

The present section reports the cloning of *Bbsp2923* gene, from *B. subtilis* BSP1 strain, that encodes a putative β -mannanase belonging to glycosyl hydrolase family 26 (GH26), classified according to Carbohydrate-Active EnZymes (CAZy) database (Cantarel *et al.*, 2009). Over-expression and production of *Bbsp2923* in *Escherichia coli*, as well as a preliminary characterization of its hydrolytic activity are also described.

3.3.1. Sequence analysis of Bbsp2923

The *Bbsp2923* gene comprises an open reading frame of 2067bp, encoding a putative protein of 688aa. Using the online Signal Peptide Prediction tool (SIG-Pred – Bradford, 2001) a N-terminal signal peptide was identified suggesting an extracellular localization for this protein. The predicted signal sequence, with positively charged amino acids at the N-terminus followed by a hydrophobic core, displays two alternative processing sites functional in gram-positive bacteria, between positions 26/27 or 28/29, while *E. coli* only recognises position 26/27 as possible processing site (see Appendix 6.3).

A BLASTp search (Altschul *et al.*, 1990) of the deduced Bbsp2923 primary structure against NCBI database showed that the predicted polypeptide had a moderate identity to other known mannanases from *Bacillus* sp. strain JAMP750 (AB128831), 57% identity in 640aa overlap, and *Cellulomonas fimi* (2BV4) 49% identity in 406aa overlap. According to BLASTp results, Bbsp2923 presents a modular structure characteristic of β -mannanases, where a glycosyl hydrolase catalytic domain from family 26 (GH-26) is appended, by a flexible linker sequence, to a carboxyl binding module of family 11 (CBM-11), as showed in figure 3.7.

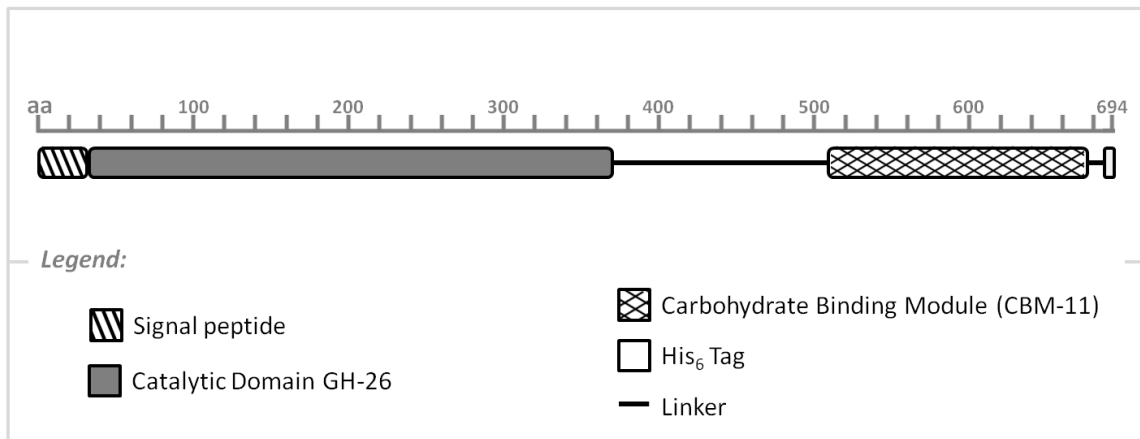


Figure 3.7 – Illustration of modular organization of Bbsp2923, adapted from BLASTp analysis.

Members of the same family of glycosyl hydrolases (CAZy) display a common fold and catalytic mechanism (Cantarel, 2009). Endo-acting β -mannanases are located within GH-5 and GH-26 families, whereas GH-2 contains β -mannosidases (Gilbert, 2010). Both types of enzymes belong to clan GH-A, characterized by a $(\beta/\alpha)_8$ barrel fold, hydrolysis of the glycosidic bond by a general acid/base-catalyzed double displacement mechanism, and the presence of two

catalytic glutamate residues, the acid/base and the nucleophile, at the C terminus of the β -strands 4 and 7, respectively (Bolam *et al.*, 1996).

Differences in conserved amino acid patterns can be used to distinguish different GH families. According to Braithwaite *et al.* (1995), one of the most highly conserved regions in all the GH-26 family members is an aromatic amino acid-rich region with the consensus sequence WFWWG.

The N-terminal sequence of Bbsp2923 and seven other well-characterized β -mannanases belonging to GH-26 family were aligned using ClustalW2 (Thompson *et al.*, 1994). The secondary structure of GH-26 characteristic motifs was depicted above the sequence alignment, using ESPript (Gouet *et al.*, 1999), in order to identify the referred characteristic motifs in Bbsp2923 (Fig. 3.8). For comparison, were selected sequences of *Cellulomonas fimi* (2BV4) and *Bacillus* sp. JAMB750 (AB128831), identical to Bbsp2923, as well as other typical mannanases from *B. amyloliquefaciens* (ABS75922), *B. licheniformis* (YP090363), *B. subtilis* W3 (Q3S4E4) and *Cellvibrio japonicus* (formerly *Pseudomonas fluorescens* subsp. *cellulose*) (1J9Y). It was also included the β -mannanase of the *gmuA-G* operon in *B. subtilis* 168, GmuG (O05512) (Sadaie *et al.*, 2008).

Bbsp2923 presents the classic $(\beta/\alpha)_8$ barrel fold of GH-A clan (Fig. 3.8). The two catalytic glutamate residues, the acid/base and the nucleophile, in the β -strands 4 and 7, respectively, are also visible (Fig. 3.8, marked with a star). The highly conserved region WFWWG is noticeable right after β_4 strand. All these evidences, classifies Bbsp2923 as an endo-1,4- β -mannanase. Secondary structure prediction obtained for Bbsp2923 alone, using PSIPred (Buchan, 2010), is presented in appendix 6.6.

As previously referred, the CBM is a non-catalytic polysaccharide-recognizing module of many glycosyl hydrolases (Gilbert, 2010). According to CAZy's classification, CBMs can be grouped into families CBM 1-39, based on amino acid sequence similarity. They can also be classified into three categories, based on ligand recognition and binding sites topologies: "surface-binding" type A; "glycan-chain-binding" type B; and "small sugar-binding" type C (Boraston *et al.*, 2004). CBM-11 is a type B glycan-chain binder, which appears generally associated with GH-26 or GH-5. It is also known that type B CBMs can bind to a great diversity of ligands and that its specificity greatly depends on the presence of aromatic residues in the binding cleft (Boraston *et al.*, 2004). The work of Viegas *et al.* (2008) showed that a CBM-11 of *Clostridium thermocellum*, binds strongly the glucose units of cellotetraose and cellohexaose. However, no studies were performed using mannans. The CBMs domains by binding the

substrate can facilitate the action of the catalytic domain. Thus, in general, the ligand specificity of CBMs reflects the substrate degraded by the linked catalytic domain. So it is suggested by the presence of CBM-11 in Bbsp2923, the putative β -mannanase binding to linear and heteromannans.

It must be noticed that GmuG, the β -mannanase of glucomannan utilization operon in *B. subtilis* 168, is a non-modular enzyme, presenting only a single module with GH-26 homology. So, the presence of CBM-11 in Bbsp2923 can contribute to greater substrate detection, giving BSP1 an adaptive advantage. No type 11-CBMs exist in the *B. subtilis* 168 genome, thus the presence of a CBM-11 in the *gam* cluster supports horizontal gene transfer.

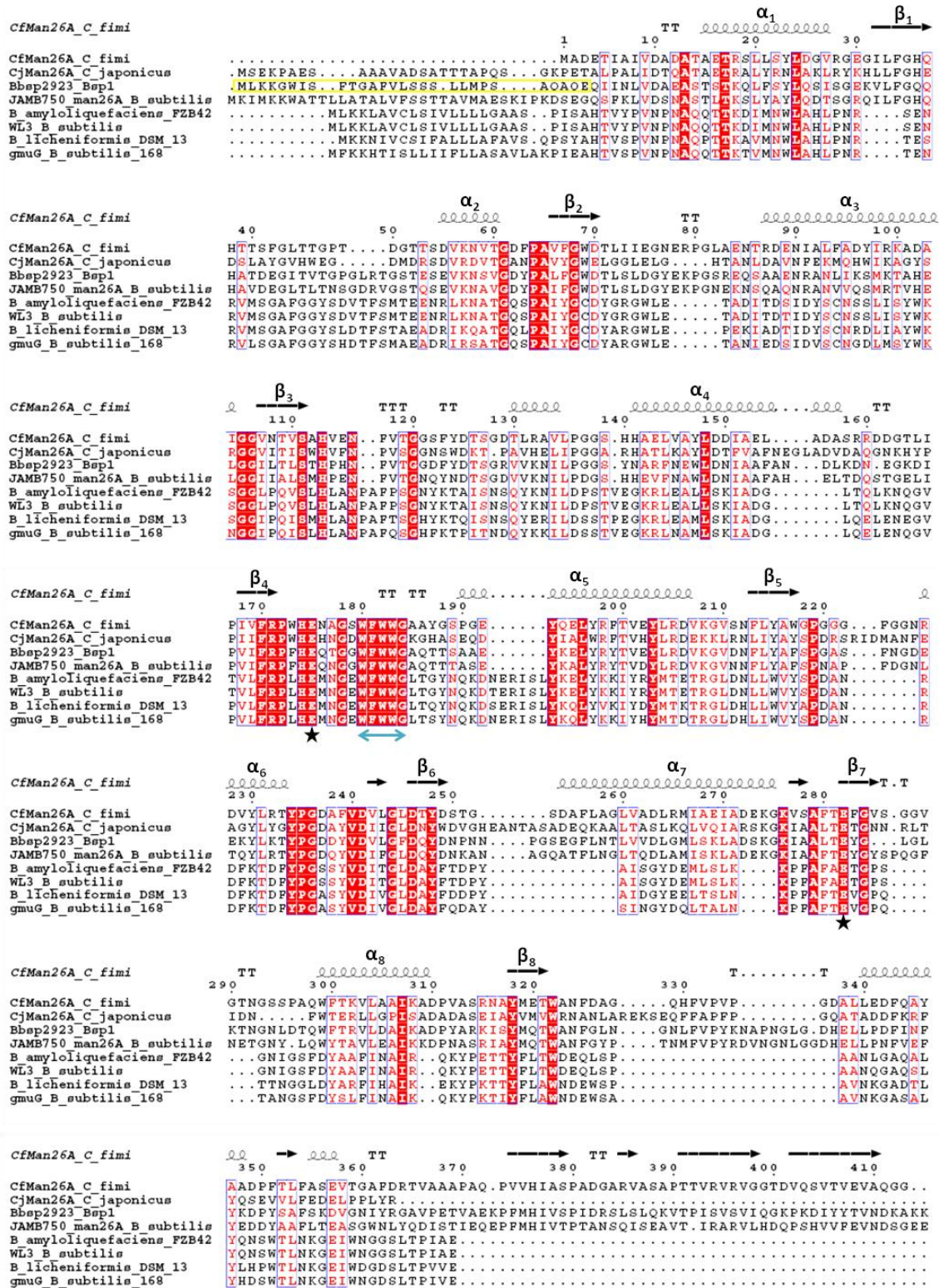


Figure 3.8 – N-terminal sequence alignment of Bbsp2923 and other GH-26 members correlated with secondary structure prediction information. Sequence alignment was performed with ClustalW2 (Thompson *et al.*, 1994). Strictly conserved residues are highlighted in red, and conservatively substituted residues are boxed. The secondary structure elements, helices (α) and strands (β), are shown above the alignment. Double T's indicate β turns and triple T's, strict α turns. Putative catalytic residues are marked with a star. Conserved WFWWG region is marked with a two-direction arrow. Bbsp2923 signal peptide is yellow boxed. GenBank or PDB sequences accession number is given along the text. The scheme was designed using ESPrnt (Gouet *et al.*, 1999).

3.3.2. Expression of *Bbsp2923* in *E. coli*

For the expression and production of *Bbsp2923*, the pVC1 vector was constructed (Appendix 6.4), by inserting the *Bbsp2923* complete gene into the pET30a(+) expression vector. pVC1 allows the production of a recombinant version of *Bbsp2923* with the native signal peptide at the N-terminus and bearing an His₆-tag at the C-terminus, with a total length of 694aa and 76.4kDa of molecular mass. A small scale induction of the expression protocol of *E. coli* BL21 pLysS (DE3) strain harboring pVC1 was performed. However, after SDS-PAGE analysis of the soluble and insoluble fraction of the cell extracts, no production was obtained (data not showed). Thus, overproduction was attempted using an auto-induction protocol, as described in section 2.8 of Materials and Methods, but again no protein was detected (data not shown).

The analysis of the *Bbsp2923* sequence revealed the presence of several rare codons in *E. coli*, namely ACA, GGA and TCA that encode fifteen threonines, twenty-one glycines and ten serines, respectively. Thus, it was hypothesized that failure of production was due to codon bias of *E. coli*. To overcome this problem, pVC1 was transformed in *E. coli* Rosetta™ (DE3) pLysS. This strain possesses the pRARE plasmid, which encodes tRNA genes for all the rarely used codons that can impede translation in *E. coli* (Novy *et al.*, 2001). Overproduction of *Bbsp2923* in *E. coli* Rosetta™ (DE3) pLysS strain harboring pVC1 was attempted using small scale and auto-induction protocols without success (data not shown).

To attain *Bbsp2923* overproduction an alternative strategy was followed by constructing pVC2 (Appendix 6.5), a pET30a(+) derivative carrying a truncated *Bbsp2923* gene without the signal peptide sequence, encoding a 666aa polypeptide of 73.4kDa molecular mass. *E. coli* BL21 pLysS (DE3) and *E. coli* Rosetta™ (DE3) pLysS strains, both harbouring pVC2 were subjected to a small scale induction of the expression protocol for the production of the recombinant version of *Bbsp2923* without the signal peptide.

As showed in figure 3.9, the SDS-PAGE analysis of the cell extracts revealed the presence of *Bbsp2923* (approximately 73 kDa) in the soluble fraction of both *E. coli* BL21 pLysS (DE3) and *E. coli* Rosetta™ (DE3) pLysS strains. Noteworthy, the production of recombinant *Bbsp2923*-His₆ is extremely enhanced in *E. coli* Rosetta™ (DE3) pLysS harboring pVC2 due to the presence of pRARE.

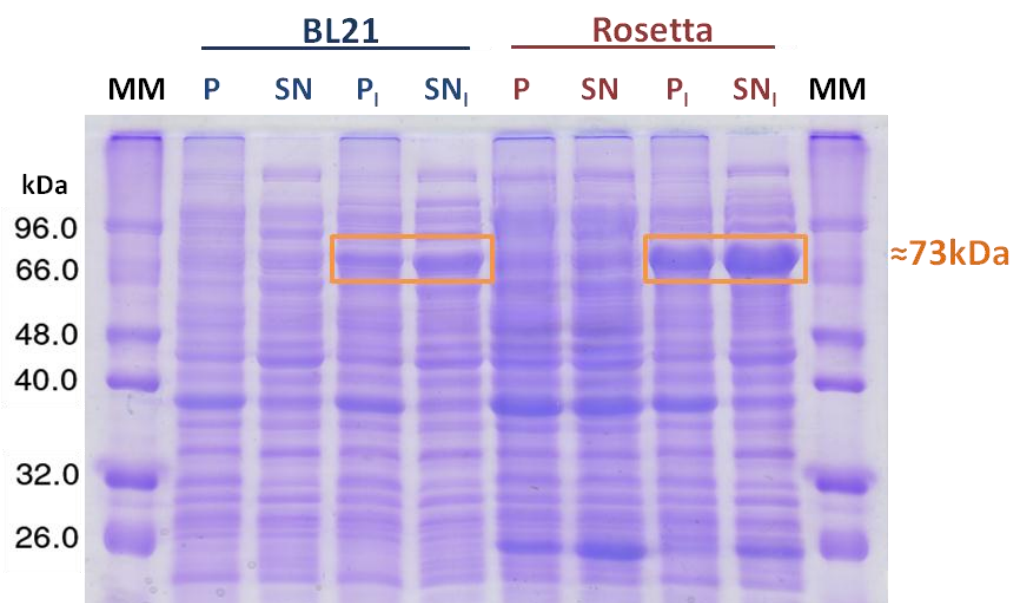


Figure 3.9 – Overproduction of recombinant Bbsp2923-His₆ in *E. coli* BL21 pLysS (DE3) and *E. coli* Rosetta™ (DE3) pLysS strains harboring pVC2. Samples of 8 μL of cell extract fraction plus 2 μL of loading buffer were resolved on a 12.5% SDS-PAGE gel, stained with Coomassie Blue. (MM) NZYTech LMW Protein Marker (NZYTech, Lda.); (S) Soluble fraction; (P) Insoluble fraction; (I) Induced samples using IPTG 1mM; Orange boxes highlight Bbsp2923.

3.3.3. β-Mannanase activity assay

Due to lack of time, a preliminary characterization of Bbsp2923-His₆ enzymatic activity towards different substrates was performed, using Nelson-Somogyi reducing sugars content method (Somogyi, 1952).

The analyses were conducted using the soluble fraction of induced *E. coli* Rosetta™ (DE3) pLysS harboring pVC2 cell extracts. The soluble fraction of induced *E. coli* Rosetta™ (DE3) pLysS carrying pET30a(+) cell extracts was used as negative control.

As showed in table 3.6, recombinant Bbsp2923 was found to be active towards linear mannan (1,4-β-D-mannan), carob galactomannan and konjac glucomannan, but inactive towards lupin galactan.

Table 3.6 – Activity rates of recombinant Bbsp2923, against several tested substrates.

Substrate	Enzymatic activity ^a ($\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$)	
	pVC2 (Bbsp2923-His ₆)	pET30a(+)
1,4-β-D-Mannan	8.21	ND
Carob Galactomannan	6.21	ND
Konjac Glucomannan	3.25	ND
Lupin Galactan	0.09	ND

^a Activity of *E. coli* Rosetta™ (DE3) pLysS strain harboring pVC2 or empty pET30a(+) vectors was determined as described. ND - No detected activity.

These results indicate that Bbsp2923 displays endo-1,4- β -mannanase activity, since all tested substrates are 1,4- β -polymers, except lupin galactan, that is formed by a 1,6- α -linked galactose backbone (Fig. 3.10).

The rate of hydrolysis towards glucomannan is lower than that observed for mannan and galactomannan. This result could reflect the higher amount of 1,4- β linkages available for recognition and cleavage in mannan and galactomannan (Fig. 3.10).

Yoon and co-workers (2008) tested a β -mannanase from *B. subtilis* WL3 and obtained an identical result – the relative activity in locust bean gum (LBG) (a synonym for carob galactomannan) was higher than the obtained in konjac glucomannan. However, linear mannan was not tested.

The results clearly show that Bbsp2923 is a β -mannanase. Purification of the recombinant enzyme is required for future biochemical and physical characterization.

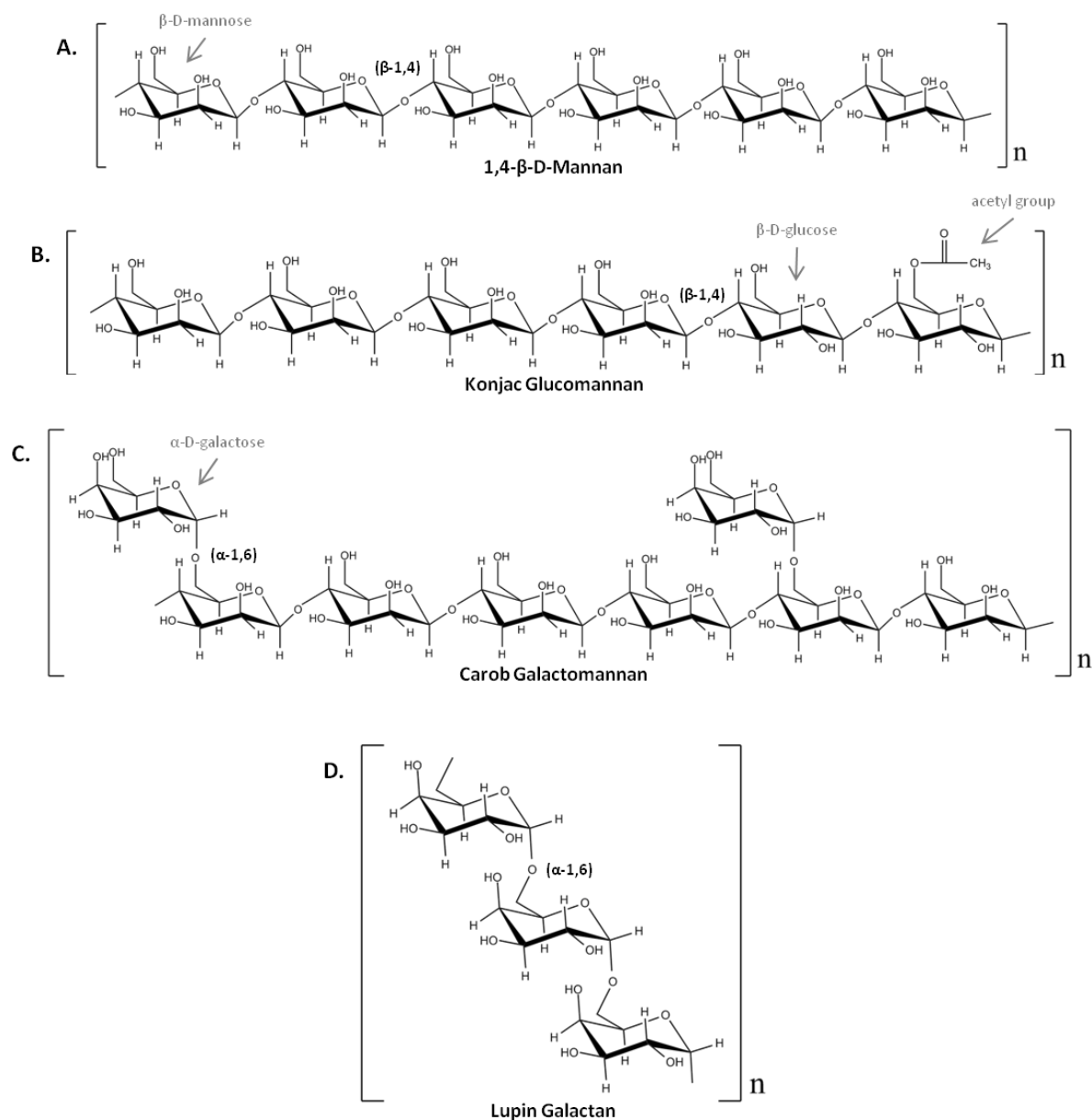


Figure 3.10 – Chemical structure of used substrates. (A) 1,4-β-D-Mannan structure, a main chain of β-1,4 linked β-D-mannose residues; (B) Konjac glucomannan structure, a main chain of β-1,4 linked β-D-mannose and β-D-glucose residues in a 4:2 ratio. Some glucose residues might be acetylated; (C) Carob galactomannan structure, a main chain of β-1,4 linked β-D-mannose residues with α-1,6 linked α-D-galactose residues attached to some mannose residues in a 1:4 ratio; (D) Lupin galactan structure, a main chain of α-1,6 linked α-D-galactose residues; (n) Depicted sugars formed by several repetitions of illustrated unites. Adapted from Dhawan & Kaur (2007) and Moreira & Filho (2008).

Chapter 4

Concluding Remarks and Future Perspectives

4. Concluding remarks and future perspectives

Bacillus subtilis is one of the popular microorganisms for industrial biotechnology, extensively used for enzyme secretion and fine chemicals/metabolites production. Besides that, *B. subtilis* is preferred as gram-positive model for fundamental research. Elucidation of microbial physiology and genetic organization is crucial to modulate and control metabolic pathways, leading interest features towards new commercial applications.

Several features distinguish laboratory and undomesticated strains of *Bacillus subtilis*. Recently, the genome sequence of an undomesticated strain of *B. subtilis*, BSP1, was determined (Schyns, G., Serra C.R., Henriques, A.O. *et al.*, unpublished results). In addition to the eight-gene *gmuA-G* operon for glucomannan utilization present in strain 168, the undomesticated BSP1 strain, has a second cluster of genes most likely specifying a system for the degradation and utilization of a different mannan polymer. This system comprises a second putative extracellular β -mannanase. The galactoglucomannan *gam* cluster is found at about 176° in the genome of BSP1 where it replaces a region of eleven genes, in the chromosome of strain 168 (Schyns, G., Serra C.R., Henriques, A.O. *et al.*, unpublished results, and Fig. 1.4). The *gam* cluster (Fig. 1.4) codes not only for a predicted extracellular β -mannanase, but also for an intracellular α -galactosidase, a β -glucosidase, and an N-acetyl mannosamine esterase. The presence of these genes suggests that the polysaccharide substrate is a heteromannan, possibly a galactoglucomannan, formed by a main chain of β -1,4 linked mannose and glucose residues, with α -galactose attached to the O-6 position of some mannose residues through an α -1,6 bond (Schyns, G., Serra C.R., Henriques, A.O. *et al.*, unpublished results). The proposed model for the degradation of galactoglucomannan in *B. subtilis* BSP1 is illustrated in figure 4.1.

In the presence of galactoglucomannan rich substrates, a pre- β -mannanase with a signal peptide (Bbsp2923) is produced and further secreted by the Sec system. The extracellular β -mannanase randomly cleaves the β -1,4 bonds within galactoglucomannan, releasing oligosaccharides. These smaller saccharides are then captured by a binding protein (Bbsp2927) and delivered to the other components of the ABC transporter (Bbsp2926, Bbsp2928 and Bbsp2929). This system is responsible for oligo-mannans translocation across cell membrane. Remaining enzymes will act depending on translocated oligomer. The N-acetyl-mannosamine esterase (Bbsp2931) hydrolyzes acetyl substitutions in D-mannose units. The α -galactosidase (Bbsp2930) can cleave the α -1,6 linkage between galactose and galactose, glucose or

mannose residues, while β -glucosidase (Bbsp2925) separates D-glucose and D-mannose residues linked through a β -1,4 bond (Fig. 4.1).

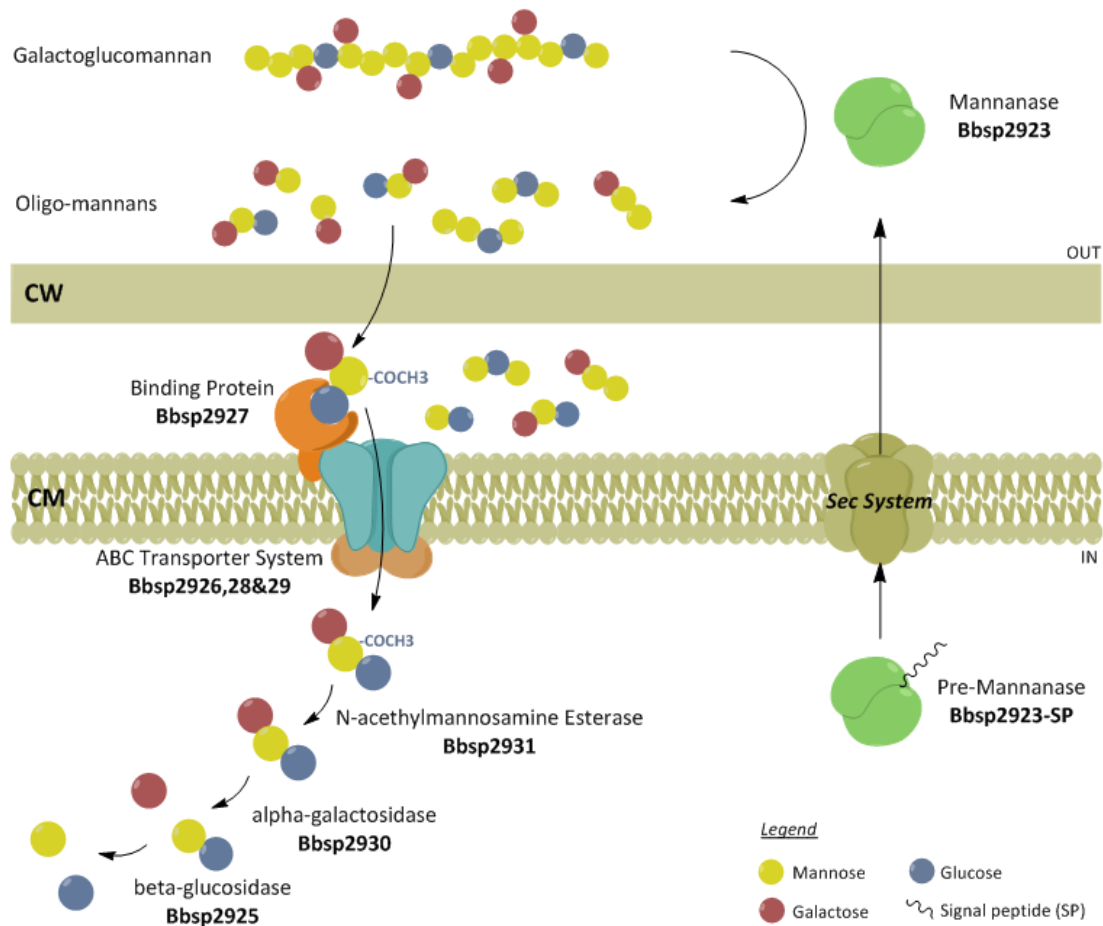


Figure 4.1 – Proposed model for the degradation of galactoglucomannan in *B. subtilis* BSP1. Extracellular β -mannanases (Bbsp2923) randomly cleave the galactoglucomannan backbone formed by β -1,4-linked mannose and glucose residues. The oligo-mannans are transported to the cytoplasm by the combined action of a specific binding protein (Bbsp2927) and an ABC Transporter system (Bbsp2926, Bbsp2928 and Bbsp2929). N-acetyl-mannosamine esterases (Bbsp2931) hydrolyzes the acetate groups (-COCH₃) substituting O-2 or O-3 mannose units, while α -galactosidases (Bbsp2930) cleave the α -1,6-linked galactose side chains. β -glucosidases (Bbsp2925) will hydrolyse the β -1,4 bond between D-glucose and D-mannose residues.

The genetic organization and function of the *gam* cluster, and whether its presence in BSP1 strain results in new hemicellulolytic capabilities, compared to those of the laboratory strain, was address in this work.

The transcriptional organization of the cluster of genes *Bbsp2924-Bbsp2929* within the *gam* cluster was analyzed by RT-PCR and the results revealed that it does not constitute an operon. Instead of a single transcriptional unit of six cistrons, genes *Bbsp2924* and *Bbsp2925* are monocistronic. On the other hand *Bbsp2926-Bbsp2927* appear to be co-transcribed, and likewise *Bbsp2928-Bbsp2929* constitute a bicistronic transcriptional unit. Further studies should address transcription analysis of the remaining genes of the *gam* cluster *Bbsp2923*, *Bbsp2930* and *Bsp2931*, utilizing the same methodology and/or Northern blot analysis. Furthermore, the putative promoters identified *in silico* need to be confirmed by primer extension analysis or 5'RACE (Rapid Amplification of 5'cDNA ends).

RNA extracted from cells grown in the absence of sugars and in the presence of different types of mannan was analyzed by RT-PCR, however the identification of signals that trigger the expression of the *gam* genes *Bbsp2924-Bbsp2929* remains elusive. In addition to specific regulation in response to substrate availability, genes involved in extracellular degradation of polysaccharides respond to global regulatory circuits, including those responsible for delaying the expression of certain genes until cells enter the stationary phase of growth. Thus, total RNA from cells in different stages of growth, exponential growth phase and early stationary phase, was examined by RT-PCR but no differences were detected and no conclusions were drawn. To elucidate these questions a different methodology, such as transcriptional fusions and/or real-time PCR, which constitutes a more accurate technique for mRNA quantification, should be used in future studies.

Mannans (linear mannan, glucomannan, galactomannan and galactoglucomannan) are the major constituents of the hemicellulose fraction in softwoods and show widespread distribution in plant tissues (Chauhan *et al.*, 2012). Microbial mannan degrading enzymes, in particular β -mannanases, have found many applications in the food, feed, pharmaceutical, and pulp and paper industries. In the last years, many groups have isolated mannanases from several *Bacillus* spp. but without further characterization of the genetic regulation. The study of Sadaie and co-workers (2008), who characterized a putative glucomannan utilization operon in *B. subtilis* 168, is an exception.

A qualitative analysis of the hemicellulolytic potential of BSP1 was assessed towards mannans and xylan and compared to other *Bacillus* spp. strains. It is clear that the undomesticated strain BSP1 possesses augmented extracellular hydrolytic capabilities. To elucidate the specific contribution of the *gam* cluster to the enhanced ability to hydrolyse mannans, further analysis is required, such as gene disruption and growth kinetics

experiments. In addition, determination of the intracellular activity of two enzymes α -galactosidase (present in the *gam* cluster) and β -galactosidase (absent) should be tested for comparison.

In this work, cloning, expression, and production of Bbsp2923 in *Escherichia coli* was achieved. Preliminary characterization showed that the enzyme is indeed a β -mannanase, displaying higher activity towards 1,4- β -D-mannan and carob galactomannan. Full characterization of the enzyme requires the protein in a pure form for further biochemical analyses on its activity and substrates specificity, and additional determination of physico-chemical properties and kinetic parameters. Activity towards manno-oligosaccharides, such as mannobiose and mannotriose, will be necessary to exclude the hypothesis of exo-acting activity, as documented by Cartmell *et al.* (2008) who showed that subtle changes in the active site region of β -mannanase (CJMan26A) from *Cellvibrio japonicus*, could confer exo-mode of action.

The high hemicellulolytic capacity of the undomesticated *B. subtilis* BSP1, demonstrated in this study, suggests potential to be used in the food and feed industries and modification of this strain by genetic engineering constitute an appealing line for future investigation.

Chapter 5

References

5. References

- Altschul, S.F., Gish, A., Miller, W., Myers, E.W. and Lipman, D.J.** 1990. Basic local alignment search tool. *J Mol Biol.* **215** (3): 403-410.
- Anagnostopoulos, C. and Spizizen, J.** 1961. Requirements for transformation in *Bacillus subtilis*. *J Bacteriol.* **81**: 741-746.
- Barbosa, T.M., Serra, C.R., La Ragione, R.M., Woodward, M.J. and Henriques, A.O.** 2005. Screening for *Bacillus* Isolates in the Broiler Gastrointestinal Tract. *Appl Environ Microbiol.* **71**(2): 968-978.
- Bolam, D.N., Hughes, N., Virden, R., Lakey, J.H., Hazlewood, G.P., Henrissat, B., Braithwaite, K.L. and Gilbert, H.J.** 1996. Mannanase A from *Pseudomonas fluorescens* ssp. *cellulosa* is a retaining glycosyl hydrolase in which E212 and E320 are the putative catalytic residues. *Biochemistry.* **35**(50): 16195-204.
- Boraston, A.B., Bolam, D.N., Gilbert, H.J. and Davies, G.J.** 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J.* **382**(Pt 3): 769-81.
- Bradford, J.R.** 2001. *In silico* Methods for Prediction of Signal Peptides and their Cleavage Sites, and Linear Epitopes. MRes in Bioinformatics, SBMB, University of Leeds (UK).
- Braithwaite, K.L., Black, G.W., Hazlewood, G.P., Ali, B.R. and Gilbert, H.J.** 1995. A non-modular endo-beta-1,4-mannanase from *Pseudomonas fluorescens* subspecies *cellulosa*. *Biochem J.* **305** (3): 1005-10.
- Buchan, D.W., Ward, S.M., Lobley, A.E., Nugent, T.C., Bryson, K. and Jones, D.T.** 2010. Protein annotation and modelling servers at University College London. *Nucleic Acids Res.* **38**(Web Server issue): W563-8.
- Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V. and Henrissat, B.** 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* **37**(Database issue): D233-8.

- Cartmell, A., Topakas, E., Ducros, V.M., Suits, M.D., Davies, G.J. and Gilbert, H.J.** 2008. The *Cellvibrio japonicus* mannanase CjMan26C displays a unique exo-mode of action that is conferred by subtle changes to the distal region of the active site. *J Biol Chem.* **283**(49): 34403-13.
- Casula, G. and Cutting, S. M.** 2002. *Bacillus* Probiotics: Spore Germination in the Gastrointestinal Tract. *Appl Environ Microbiol.* **68**: 2344-2352.
- Chauhan, P.S., Puri, N., Sharma, P. and Gupta, N.** 2012. Mannanases: microbial sources, production, properties and potential biotechnological applications. *Appl Microbiol Biotechnol.* **93**(5): 1817-30.
- Chen, X.H., Koumoutsis, A., Scholz, R., Eisenreich, A., Schneider, K., Heinemeyer, I., Morgenstern, B., Voss, B., Hess, W.R., Reva, O., Junge, H., Voigt, B., Jungblut, P.R., Vater, J., Süßmuth, R., Liesegang, H., Strittmatter, A., Gottschalk, G. and Borriss, R.** 2007. Comparative analysis of the complete genome sequence of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. *Nat Biotechnol.* **25**(9): 1007-14.
- Davidson, A.L. and Chen, J.** 2004. ATP-binding cassette transporters in bacteria. *Annu Rev Biochem.* **73**: 241-68.
- Deutscher, J., Galinier, A. and Martin-Verstraete, I.** 2002. Carbohydrate uptake and metabolism. *In* Sonenshein, L., Hoch, J. and Losick, R. (ed.), *Bacillus subtilis* and its Closest Relatives: from Genes to Cells. ASM Press, Washington, DC.
- Dewan, S.S.** 2012. Global Markets for Enzymes in Industrial Applications. *BCC Research* (Online report – code: BIO030G)
- Dhawan, S. and Kaur, J.** 2007. Microbial Mannanases: An Overview of Production and Applications. *Crit Rev Biotechnol.* **27**: 197-216.
- Douglas, S.E.** 1995. DNA Strider, an inexpensive sequence analysis package for the Macintosh. *Mol Biotechnol.* **3**(1): 37-45.
- Ferrari, E., Jarnagin, A.S. and Schmidt, B.F.** 1993. Commercial Production of Extracellular Enzymes. *In* Sonenshein, L., Hoch, J. and Losick, R. (ed.), *Bacillus subtilis* and other Gram-positive Bacteria. ASM Press, Washington, DC.

Fujita, Y. 2009. Carbon catabolite control of the metabolic network in *Bacillus subtilis*. *Biosci Biotechnol Biochem.* **73**(2): 245-59.

Garrity, L.F. and Ordal, G.W. 1995. Chemotaxis in *Bacillus subtilis*: how bacteria monitor environmental signals. *Pharmacol Ther.* **68**(1):87-104.

Gilbert, H.J. 2010. The Biochemistry and Structural Biology of Plant Cell Wall Deconstruction. *Plant Physiol.* **153**(2): 444–455.

Gouet, P., Courcelle, E., Stuart, D. I. and Metz, F. 1999. ESPript: multiple sequence alignments in PostScript. *Bioinformatics.* **15**: 305-8.

Hager, P.W., Rabinowitz, J.C. 1985. Translational specificity in *Bacillus subtilis*. In Dubnau D.A. (ed.), *The Molecular biology of the bacilli*, Volume II. Academic Press, Inc.

Harwood, C.R. 1992. *Bacillus subtilis* and its relatives: molecular biological and industrial workhorses. *Trends Biotechnol.* **10**(7): 247-56.

Helmann, J.D. and Moran, C.P. 2002. RNA Polymerase and sigma factors. In Sonenshein, A. L., Hoch, J. A. and Losick, R. (ed.), *Bacillus subtilis* and its closest relatives: from genes to cells. ASM Press, Washington, DC.

Higgins, C.F. 1992. ABC Transporters: From microorganisms to Man. *Annu Rev Cell Dev Biol.* **8**: 67-113.

Higgins, C.F. 2001. ABC transporters: physiology, structure and mechanism – an overview. *Res Microbiol.* **152**: 205-210.

Hollenstein, K., Dawson, R.J. and Locher, K.P. 2007. Structure and mechanism of ABC transporter proteins. *Curr Opin Struct Biol.* **17**(4): 412-8.

Inácio, J.M., Costa, C. and de Sá-Nogueira, I. 2003. Distinct molecular mechanisms involved in carbon catabolite repression of the arabinose regulon in *Bacillus subtilis*. *Microbiology.* **149**(9):2345-55.

- Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V. et al.** 1997. The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature*. **390**: 249-256.
- Larion, M., Moore, L.B., Thompson, S.M. and Miller, B.G.** 2007. Divergent evolution of function in the ROK sugar kinase superfamily: role of enzyme loops in substrate specificity. *Biochemistry*. **46**(47): 13564-72.
- McNeil, M., Darvill, A.G., Fry, S.C. and Albersheim, P.** 1984. Structure and function of the primary cell walls of plants. *Annu Rev Biochem*. **53**: 625-663.
- Miller, J.H.** 1972. *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Moreira, L.R.S. and Filho, E.X.F.** 2008. An overview of mannan structure and mannan-degrading enzyme systems. *Appl Microbiol Biotechnol*. **79**: 165-178.
- Nishinari, K., Takemasa, M., Zhang, H. and Takahashi, R.** 2007. Storage Plant Polysaccharides: Xyloglucans, Galactomannans, Glucomannans, p. 613–652. *In* Kamerling, J.P. (ed.), *Comprehensive Glycoscience - from Chemistry to Systems Biology 2: Analysis of Glycans; Polysaccharide Functional Properties*. Elsevier.
- Novy, R., Drott, D., Yaeger, K. and Mierendorf R.** 2001. Overcoming the codon bias of *E. coli* for enhanced protein expression. *Newsletter of Novagen, Inc.* **12**
- Postma, P.W., Lengeler, J.W. and Jacobson, G.R.** 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev*. **57**(3): 543-94.
- Priest, F.** 1993. Systematics and Ecology of *Bacillus*. *In* Sonenshein, A. L., Hoch, J. A. and Losick, R. (ed.), *Bacillus subtilis and other Gram-positive Bacteria*. ASM Press, Washington, DC.
- Sadaie, Y., Nakadate, H., Fukui, R., Yee, L.M. and Asai, K.** 2008. Glucomannan utilization operon of *Bacillus subtilis*. *FEMS Microbiol Lett*. **279**(1): 103-9.
- Sambrook, J. and Russell, D. W.** 2001. *Molecular Cloning: a Laboratory Manual*, 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Molecular Cloning: a Laboratory Manual, 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Schallmeyer, M., Singh, A. and Ward, O.P. 2004. Developments in the use of *Bacillus* species for industrial production. Can J Microbiol. **50**(1): 1-17.

Schyns, G., Serra, C.R., Henriques, A.O. et al. Unpublished results.

Schyns, G., Pereira-Leal, J., Serra, C., Fickers, P., Potot, S., Barbosa, T., Chevreux, B., de Saizieu, A., Perkins, J.B., Wyss, M., Henriques, A.O. 2009. The genome of a gut strain of *Bacillus subtilis*: molecular insights into a bi-phasic life cycle and the mechanisms of probiosis (Presentation). MicroBiotec09.

Shallom, D. and Shoham, Y. 2003. Microbial hemicellulases. Curr Opin Microbiol. **6**: 219–228.

Sierro, N., Makita, Y., de Hoon, M.J.L. and Nakai, K. 2008. DBTBS: a database of transcriptional regulation in *Bacillus subtilis* containing upstream intergenic conservation information. Nucleic Acids Res. **36** (Database issue):D93-D96.

Snyder, L. and Champness, W. 1997. Introduction to macromolecular synthesis: Gene expression. In Snyder, L. and Champness, W. (ed.), Molecular genetics of Bacteria, 3rd ed. ASM Press, Washington, DC.

Somogyi, M. 1952. Notes on sugar determination. J Biol Chem. **195**: 19–23.

Spizizen, J. 1958. Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. Proc Natl Acad Sci USA. **44**: 1072-8.

Studier, F.W. 2005. Protein production by auto-induction in high density shaking cultures. Protein Expr Purif. **41**: 207–234.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185**: 60–89.

Tam, N.K., Uyen, N.Q., Hong, H.A., Duc le, H., Hoa, T.T., Serra, C.R., Henriques, A.O. and Cutting, S.M. 2006. The intestinal life cycle of *Bacillus subtilis* and close relatives. J Bacteriol. **188**(7): 2692-700.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**(22): 4673-80.

Várnai, A., Huikko, L., Pere, J., Siika-Aho, M. and Viikari, L. 2011. Synergistic action of xylanase and mannanase improves the total hydrolysis of softwood. *Bioresour Technol.* **102**(19): 9096-104.

Veith, B., Herzberg, C., Steckel, S., Feesche, J., Maurer, K.H., Ehrenreich, P., Bäumer, S., Henne, A., Liesegang, H., Merkl, R., Ehrenreich, A. and Gottschalk, G. 2004. The complete genome sequence of *Bacillus licheniformis* DSM13, an organism with great industrial potential. *J Mol Microbiol Biotechnol.* **7**(4): 204-11.

Viegas, A., Brás, N.F., Cerqueira, N.M., Fernandes, P.A., Prates, J.A., Fontes, C.M., Bruix, M., Romão, M.J., Carvalho, A.L., Ramos, M.J., Macedo, A.L. and Cabrita, E.J. 2008. Molecular determinants of ligand specificity in family 11 carbohydrate binding modules: an NMR, X-ray crystallography and computational chemistry approach. *FEBS J.* **275**(10): 2524-35.

Weickert, M.J. and Adhya, S. 1992. A family of bacterial regulators homologous to Gal and Lac repressors. *J Biol Chem.* **267**(22):15869-74.

Westers, L., Westers, H. and Quax, W.J. 2004. *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochim Biophys Acta.* **1694**(1-3): 299-310.

Yamane, K., Bunai, K. and Kakeshita, H. 2004. Protein traffic for secretion and related machinery of *Bacillus subtilis*. *Biosci Biotechnol Biochem.* **68**(10): 2007-23.

Yoon, K.H., Chung, S. and Lim, B.L. 2008. Characterization of the *Bacillus subtilis* WL-3 mannanase from a recombinant *Escherichia coli*. *J Microbiol.* **46**(3): 344-9.

Chapter 6

Appendices

6. Appendices

Appendix 6.1 – Corrected sequence of Bbsp2927, after resequencing. Start codon is underlined; wrong base pair (T) is bold and strikethrough.

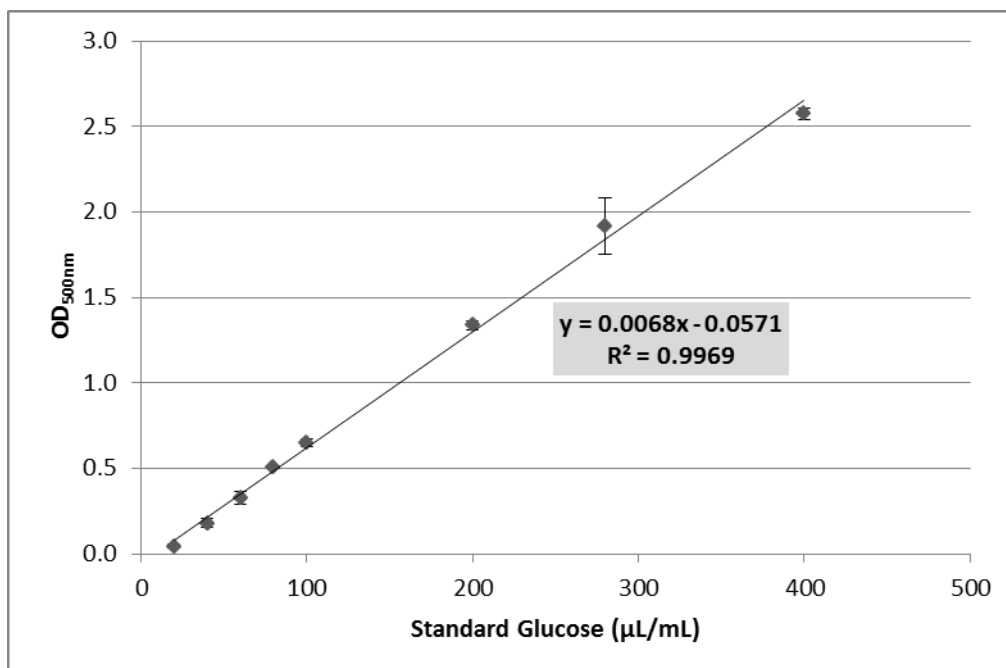
```

atgtttaagaagaaaggattaggtggttttccttttaacccttggtgctggcctttatcagcct
M F K K K G L G V F L L T L V L A L S A
tttttgtcaggatggttctggatctgatacgtcaagtggggactcaaagaatggtgtgacg
F L S G C S G S D T S S G D S K N G V T
ctccgtatattagtttgaataacaatccagaaggtacaaaactggaaaatgaaatcttt
L R I L V W N N N P E G T K L E N E I F
aaagcatttgagaaagaaaatcctgggtgcaatggtgaaatggtttatgctccatagat
K A F E K E N P G V N V E M V Y A P Y D
aagtttaacgacaagttcctaacaatgtctgccggcggtgaccagccagacttagtttg
K F N D K F L T M S A G G D Q P D L V W
atccaaccatctgcctttggacagtttagtatctaaaaatgtattgatggatctatctgat
I Q P S A F G Q L V S K N V L M D L S D
aaagagatcaataaagatgaataacttgcctaacattccttgaaatgggtaagggtgatggc
K E I N K D E Y L P N I L E M G K V D G
aagcaatacgcgctgattcgtgatgcagcagacattccagatgggctataacaaagatag
K Q Y A L I R D A A A T F Q M G Y N K D M
tttgatgaagcaggagtgccatccttcagatgattggacttgggatgacttccttaat
F D E A G V P Y P S D D W T W D D F L N
gcagcgaagaaattaacaaaagttgaaggtggaaaaacagtacagtttgaattgaaaac
A A K K L T K V E G G K T V Q F G I E N
tactatactggggaaaattttggttcaaaacgggggatcatacataagtcaagatggcaag
Y Y T G E I L V Q N G G S Y I S Q D G K
acagtgacagtgactcgccggaacaattgaagcaatccagtttggatccgatcttatc
T V T V D S P E T I E A I Q F G S D L I
aacaatatataagttcaaccaacttcagctcaagcacagggcatgtctaatttgttccta
N K Y K V Q P T S A Q A Q G M S N L F L
actggaaaagcgggcatgaaattaatgggccatgggactgggagatacagagaaaaac
T G K A G M K L M G P W D W A D T E K N
gtttcctttgagtgggatgtggcaccaatgccagcaggaaaagccgggcaatgtatctgca
V S F E W D V A P M P A G K A G N V S A
gcttcttatctgccaatcgggtattggaaaagggacaaaacatcctgaagaagcatttgag
A S Y L P I G I G K G T K H P E E A F E
ttgctgaagttcctatcaaccggtaaagggcaggatatccagattgatacaattaacgca
L L K F L S T G K G Q D I Q I D T I N A
attccagtggtaaagcgtcatgcagataaaaataactactatgaaaaacgcgccggagaat
I P V V K R H A D K I T T M K N A P E N
gcagcttcattaggaacgcttcttcaagaaggcaaaaactggtatgaatgcaccgtataca
A A S L G T L L Q E G K T V M N A P Y T
cctgattatgccgagattgcaaataaaaatccagccgggtcattgacaatattaatttgaag
P D Y A E I A N K I Q P V I D N I N L K
gatctcgatgcggaaaaagaattgaaaaaaattgcggatcagattcgaaagaataaac
D L D A E K E L K K I A D Q I R K E Y N
ctgaaataa
L K -

```

Appendix 6.2 – Calibration curve for D-Glucose equivalents calculation in Nelson-Somogyi's assays. Duplicates were made and the average used for graphics plotting.

Standard Glucose ($\mu\text{L}/\text{mL}$)	OD _{500nm}		
	1	2	Average
600	2.462	2.774	2.618
480	2.770	2.766	2.768
400	2.550	2.598	2.574
280	1.802	2.032	1.917
200	1.320	1.354	1.337
100	0.664	0.632	0.648
80	0.512	0.501	0.507
60	0.353	0.299	0,326
40	0.197	0.162	0,180
20	0.054	0.037	0,046



Appendix 6.3 – Sequence of the *Bbsp2923* gene cloned in pET30a(+), coding for a putative β -mannanase, with a signal peptide – pVC1 insert (using primers pair ARA612/ARA613), and without it – pVC2 insert (using primers pair ARA664/ARA613). Legend: ATG, start codon (double underlined); Signal peptide (highlighted in grey); possible cleavage positions in gram-positive bacterium (|); possible cleavage positions in *E. coli* (+); GAG, last codon before histidine tag (dashed); *His₆-Tag* (italic); TGA, stop codon (*).

```

<gaggaactcatatgctgaaaaaagg> ARA612 Fwd
1  atgctgaaaaaaggatggatttcttttacgggtgctttcgttttatcttctctctctctt
M L K K G W I S F T G A F V L S S S L L
      <gcgcgcccagcatatgattattaatcttg> ARA664 Fwd
61  atgccttccgcacaagcccaggagcagattattaatctttagatgctgaagcatcaacg
M P S A Q A I + Q E I Q I I N L V D A E A S T
121 tcgacaaaacagctgttttcttacttgcaaagtattagcggcgaaaaagttttgctcggc
S T K Q L F S Y L Q S I S G E K V L F G
181 cagcagcacgcaacggatgaggggaattaccggttaccggaccggctctcggacaggttct
Q Q H A T D E G I T V T G P G L R T G S
241 accgaatctgaagtgaaaaactcggtcggtgactatcccgcctcttttgctgggatacg
T E S E V K N S V G D Y P A L F G W D T
301 ctcagctcttgatggctatgaaaagccgggatcacgtgaacaatctgccgcgaaaaatcga
L S L D G Y E K P G S R E Q S A A E N R
361 gccaatctgatcaaactgatgaaaaccgcacatgaacttggtggcatcttaacgcttagt
A N L I K S M K T A H E L G I L T L S
421 actcatccgcataactttgtaaccggcggcgacttttatgacacaagcggcctgtagtc
T H P H N F V T G G D F Y D T S G R V V
481 aaaaatatcctgccgggcggaagctataatgctagattcaatgaatggcttgacaacatt
K N I L P G G S Y N A R F N E W L D N I
541 gctgcttttgcaaatgatttgaaagacaacgaaggcaaatatctctgtgattttccgc
A A F A N D L K D N E G K D I P V I F R
601 ccgttccacgagcaaacccggcggctggttctggtggggagctcaaaccctcggctgca
P F H E Q T G G W F W W G A Q T T S A A
661 gaatacaaagaactttaccgatatactgtagaatacttgctgatgtaaaagcgctcgat
E Y K E L Y R Y T V E Y L R D V K G V D
721 aactttttatacgtttttcaccaggtgcatcttttaaccggagatgaagagaataccta
N F L Y A F A G S P G A S F N G D E E K Y L
781 aaaacctatcctggagatgattatggtttaggttttaggatttgatcagatgatacca
K T Y P G D D Y V D V L G F D Q Y D N P
841 aacaatccggggtctgaagggttttttaataactctcgtagtagatcttggcatgctctca
N N P G S E G F L N T L V V D L G M L S
901 aaattagcagactcgaagggaatactgctttaaactgaatattggacttgattaaaa
K L A D S K G K I A A L T E Y G L G L K
961 acaaacggaaaccttgacacacaatggttcacgagagtgcttgatgccatcaaagcagat
T N G N L D T Q W F T R V L D A I K A G
1021 ccctatgcaagaaaaatttcttatatgcagacctgggctaacttcgggcttaattggcaac
P Y A R K I S Y M Q T W A N F G L N G N
1081 ttggttccatacaaaaatgcaccaaattgggctcggagatcacgagttacttctctgac
L F V P Y K N A P N G L G D H E L L P D
1141 tttatcaatttttataaaagaccctattctgctttctcgaaagatggttgggaatatatac
F I N F Y K D P Y S A F S K D V G N I Y
1201 agaggagcagtagccgaaacagttgcagaaaagccggttatgcatattgcttcgccaatt
R G A V P E T V A E K P F M H I V S P I
1261 gaccgtagtctttcgtccaaaaagttacaccaatcagcgttagtgatgattcagggtaaa
D R S L S L Q K V T P I S V S V I Q G K
1321 cccaaagatatttattataccgtgaatgacaaaagcgaagaagtaccctcttgcaaaaggt
P K D I Y Y T V N D K A K K Y P L V K G
1381 gatggttattattatgaagggtctgctgcttttaaaaggtgacaaaagcaactattcatgta
D G Y Y Y E G S A A L K G D K A T I H V
1441 acagccgaatttgctgatggaacctctcaaaaacaaacgattaagggtgattttaaaagag
T A E F A D G T S Q K Q T I K V Y L K E
1501 ccagaaaaacagccaccaacagttgttgatacatttgagacatattatggcgatgatgaa
P E K Q P P T V V D T F E T Y Y G D D E

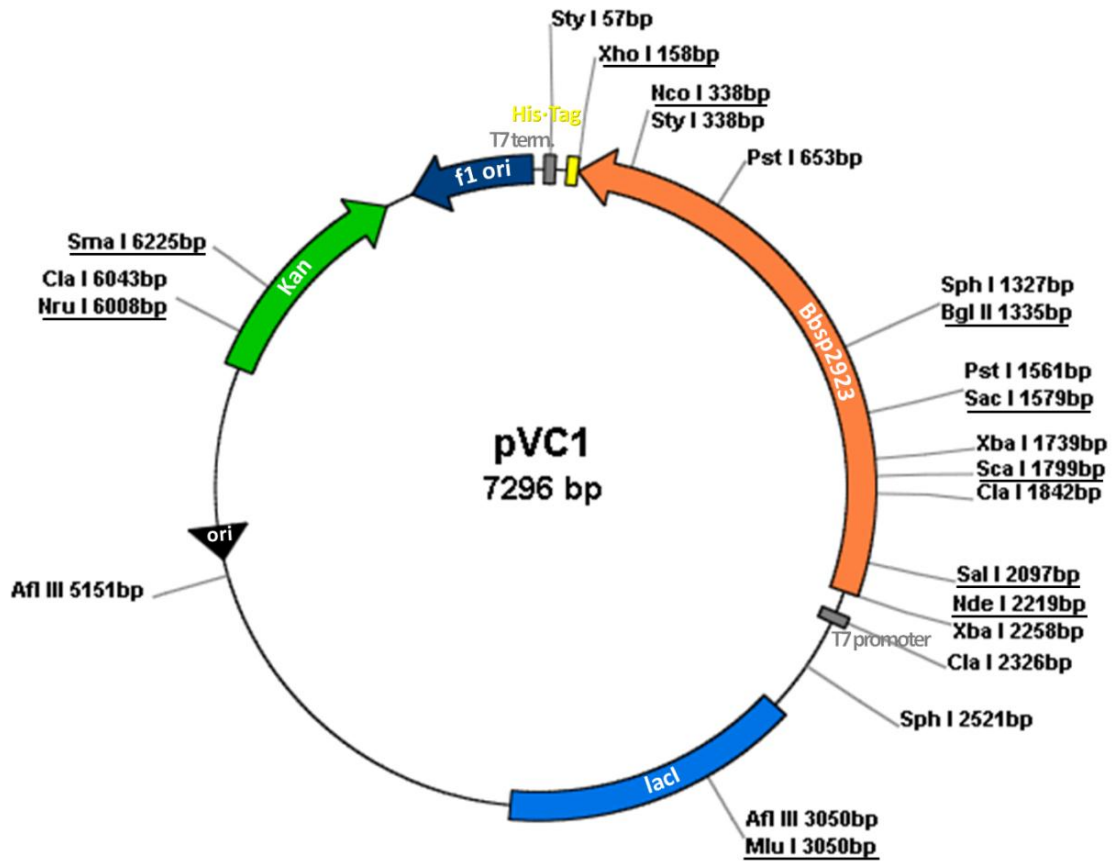
```

Appendix 6.3 (continuation) – See legend above.

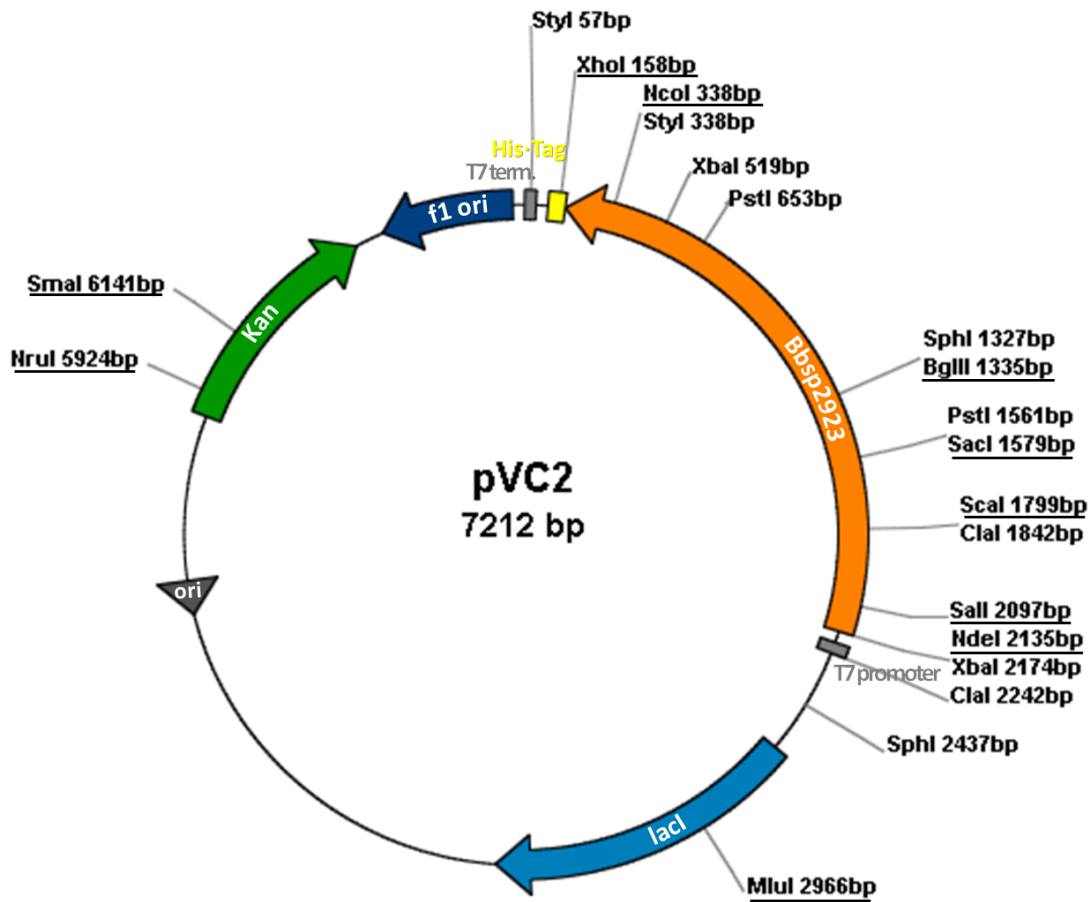
1561 cagctgcaggctgcctttgctactcaaggtgatcccttaaagctatcgtaaccagtagc
Q L Q A A F A T Q G D P L K L S L T S S
1621 gtaaaaacggaagggttcctatgcattaaaatgatgactattcactagctggagcaggttat
V K T E G S Y A L K Y D Y S L A G A G Y
1681 actggcatgaccaaactctctagatagtggttgactgggctgaagcagattcaattgatttc
T G M T K S L D S V D W A E A D S I D F
1741 tggctgaaagcagacggtggaaatcaaaaaatgggttggtcaactcaatgccggaggaatc
W L K A D G G N Q K M V V Q L N A G G I
1801 gccttcgagggcgtatccgtcacttgcttctaaaacagccagcgagggtatcaatcccattt
A F E A Y P S L A S K T A S E V S I P F
1861 tccgaatttaaacggcccatgggaatctgccgaacggcaaaaagaaacattaacagct
S E F K P A P W E S A E R Q K E T L T A
1921 gaacacttgaaaagtgaagggcgttttctatattatgtaaatgctacgggatcagaacct
E H L K S V R A F S I Y V N A T G S E P
1981 gttaacggcacgatttaccttgataatattcggtgatcatcttctggaaaaacagacttt
V N G T I Y L D N I R V S S S G K T D F
2041 accgatataccggccagcctc**gag**caccaccaccaccaccactga
T D I P A S L E H H H H H H H *

<ggccagcctcgaggcacatatata> Ara613 Rev

Appendix 6.4 – pVC1 plasmid map. Plasmid pVC1 constructed to express Bbsp2923 gene from *B. subtilis* BSP1(#200). *Bbsp2923* codes for a putative extracellular mannanase. The gene was amplified using the Ara612-Fwd and Ara613-Rev primers. The amplicon was digested with *Nde*I and *Xho*I restriction enzymes. The insert was cloned into pET30a(+) vector previously digested with the same enzymes. The gene sequence was fused with a C-terminal His-Tag. Expression is under control of T7 phage promoter, induced by IPTG. Underlined restriction sites are unique.



Appendix 6.5 – pVC2 plasmid map. Plasmid pVC2 constructed to express Bbsp2923 gene from *B. subtilis* BSP1(#200), without the signal peptide at the N-terminal. *Bbsp2923* codes for a putative mannanase. The gene was amplified using the Ara664-Fwd and Ara613-Rev primers. The amplicon was digested with *NdeI* and *XhoI* restriction enzymes. The insert was cloned into pET30a(+) vector previously digested with the same enzymes. The gene sequence was fused with a C-terminal His-Tag. Expression is under control of T7 phage promoter, induced by IPTG. Underlined restriction sites are unique.



Appendix 6.6 – Secondary structure prediction of Bbsp2923 using the online Protein Structure Prediction Server, PSIPred (Buchan, 2010).

