

Characterization of sialidase enzymes of *Gardnerella* spp.

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Master of Science
In the Department of Veterinary Microbiology
University of Saskatchewan
Saskatoon

By

SHAKYA PRASHASTHI KURUKULASURIYA

© Copyright Shakya P. Kurukulasuriya, April 2020. All rights reserved.

PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Veterinary Microbiology

University of Saskatchewan

Saskatoon, Saskatchewan S7N 5B4 Canada

Or

Dean

College of Graduate and Postdoctoral Studies

University of Saskatchewan

116 Thorvaldson Building, 110 Science Place

Saskatoon, Saskatchewan

S7N 5C9

Abstract

Bacterial Vaginosis (BV) is a condition that occurs when the healthy, *Lactobacillus* spp. dominated vaginal microbiota is replaced by BV related bacteria. BV is highly prevalent in women in their reproductive age and known to be associated with preterm delivery and increased susceptibility to sexually transmitted diseases. An abundance of *Gardnerella* spp. is often found in cases of symptomatic BV, although they are also found in healthy women without manifesting any signs of BV. Recently, the description of *Gardnerella vaginalis* was amended and three new species were defined within the genus *Gardnerella*: *G. leopoldii*, *G. swidsinskii*, and *G. piotii*. Sialidase activity is recognised as an important virulence factor that contributes to degradation of vaginal mucus and is only found in *G. piotii* and *Gardnerella* genome species 3. For years, *nanH1* was assumed to be the gene responsible for sialidase activity, but the intact open reading frame (ORF) is also found in sialidase negative strains. Our lab group discovered a gene (*nanH3*) in *Gardnerella piotii* and *Gardnerella* genome sp. 3 that is predicted to encode a cell wall attached, extra cellular sialidase. Interestingly, the ORF of *nanH3* contains a homopolymeric tract of about 12 cytosine residues. Genomic regions that contain short, homogenous or heterogenous repeats are susceptible for slipped-strand mispairing (SSM) and may change the length of the repeat region at each replication. Here we attempt to characterize *nanH3*, determine if the homopolymer region of *nanH3* varies in length in *G. piotii* and localize sialidases encoded by *nanH1* and *nanH3*.

Since previous attempts to express the entire *nanH3* failed, a truncated version of *nanH3* was expressed as a GST fusion protein (GST+TN3) in *E. coli*. Although expression of GST+TN3 was successful, the catalytic activity of the recombinant protein was not confirmed due to its poor solubility. The length of the homopolymer region of *nanH3* varied from 8-14 cytosine residues within and among strains *Gardnerella piotii*, W11, VN014, VN015 and NR032 indicating that the expression of *nanH3* may be regulated by SSM. Sialidase activity was more associated with the intact cells and the sonicated cell pellet than the respective supernatants. This suggests that sialidase activity of the four strains of *Gardnerella piotii* is more likely to be localized in the cell wall.

The results of this study contribute to knowledge of characteristics that differentiate *Gardnerella* spp. and to the future development of preferable diagnostics for identifying high risk microbiomes.

Acknowledgements

I wish to express my sincere gratitude to my supervisor, Dr. Janet Hill for providing me invaluable guidance throughout my research. Her clear vision, immense knowledge and dedication towards work have deeply inspired me. Without her persistent support the goals of this project would not have been accomplished.

Beside my supervisor, I would like to thank the rest of my Research Advisory Committee: Dr. Antonio Ruzzini, Dr. Joseph Rubin, Dr. Vikram Misra for their insightful comments and assistance.

My heartfelt gratitude also goes to our Laboratory technician, Champika Fernando for sharing her knowledge and training me. Professional and moral support she renders to all graduate students makes the Hill Lab a great place to work. I would like to thank my fellow lab mates: Salahuddin Khan, Pashupati Bandari, Scott Dos Santos, Maodong Zhang and Laura Speerin for the stimulating discussions and the fun time we had working together as a group. I feel truly privileged to be a member of the Hill Lab.

I am thankful to Douglas Fansher from Dr. David Plamer's lab (Department of Chemistry) for helping me purify recombinant NanH1. I express my thanks to Noreen Rapin and Dr. Sonu Subudhi from Misra Lab for sharing their insights and helping me to trouble-shoot protein screening.

A special thanks goes to Linda Nemeth and Lana Abrey for their assistance throughout my graduate program.

This project would not have been possible unless for the funding sources: NSERC and WCVM.

I am forever indebted to my parents for their love, care, support and encouragement throughout my life to achieve my goals. I am very much thankful to my husband, Deshan Jayampath for being supportive in all possible ways to complete my research work and thesis. Last but not least, my heartfelt gratitude goes to my sister, Dr. Shanika Kurukulasuriya for being my role model and inspiration to pursue my master's degree at University of Saskatchewan. I would also like to thank my brother-in-law, Viraj Premaratne and my nephew Thiven.

Table of Contents

PERMISSION TO USE	i
Abstract	ii
Acknowledgements	iii
List of Tables.....	vi
List of Figures	vi
List of Abbreviations.....	ix
1. Introduction.....	1
2. Literature review	3
2.1 Vaginal microbiome and bacterial vaginosis.....	3
2.1.1 Bacterial Vaginosis	3
2.1.2 Diagnosis of BV.....	4
2.2 Gardnerella spp.....	6
2.2.1 Variation within the genus <i>Gardnerella</i>	7
2.2.2 Description of new species of genus <i>Gardnerella</i>	9
2.2.3 Virulence factors of <i>Gardnerella</i> spp.	9
2.3 Sialidases	10
2.3.1 Extra cellular sialidase of a closely related species, <i>Bifidobacterium bifidum</i>	11
2.3.2 Genes encoding sialidases in <i>Gardnerella</i> spp.	12
2.4 Phase variation in bacteria.....	16
2.4.1 Mechanisms of phase variation.....	16
2.4.2 Is nanH3 a candidate for phase variation?	17
3.0 Objectives.....	20
4.0 Methods.....	21
4.1 Protein domain identification and sequence alignment.....	21
4.2 DNA and protein sequencing	21
4.3 Bacterial strains and culture conditions.....	21
4.4 DNA manipulation	23
4.4.1 PCR amplification of putative functional domain of <i>nanH3</i> / truncated <i>nanH3</i> and DNA ligation.....	23
4.4.2 Transformation of <i>E. coli</i>	24
4.5 Gene expression.....	27

4.5.1	Expression of <i>nanH1</i> , <i>nanH3</i> sialidase domain and Truncated <i>nanH3</i>	27
4.5.2	Determination of optimal conditions for expression of sialidase domain encoding region of <i>nanH3</i>	28
4.5.3	Large-scale expression of <i>nanH1</i> and sialidase domain encoding region of <i>nanH3</i>	29
4.5.4	Solubility of expressed NanH3 sialidase domain in Tris-HCl buffer (pH 7.4)	29
4.6	Purification of proteins	29
4.6.1	Purification of soluble His-tagged recombinant protein	29
4.6.2	Purification of insoluble protein	30
4.7	Sialidase assay	31
4.8	Amplification of homopolymer region of <i>nanH3</i>	32
4.9	Production of polyclonal antibodies against NanH1	33
4.10	Western blotting	34
5.0	Results	37
5.1	Comparison of sialidases	37
5.1.1	Comparison of NanH3 of <i>Gardnerella</i> sp. strain W11 with exo- α -sialidase of <i>Bifidobacterium bifidum</i>	37
5.1.2	Comparison of putative cell-wall tethered sialidases of <i>Gardnerella</i> spp.: NanH3 of strain W11, NanH2 and NanH3 of JCP8151B.....	39
5.2	Expression of NanH3 sialidase domain.....	40
5.2.1	Expression of NanH3 sialidase domain with an N-terminal 6 \times His tag (His-SiaN3)....	40
5.2.2	Expression of the NanH3 sialidase domain as a GST fusion protein	42
5.3	Expression of Truncated <i>nanH3</i> with a GST tag.....	50
5.3.1	Amplification and cloning of Truncated <i>nanH3</i> into pET-41a.....	50
5.3.2	Expression of GST fused TN3	52
5.3.3	Determining the solubility of GST+TG2	52
5.3.4	Low temperature expression of GST tagged TN3	53
5.4	Evidence of slipped strand mispairing in <i>nanH3</i>	54
5.4.1	Amplification and screening of homopolymer region of <i>nanH3</i> in <i>Gardnerella piovii</i> strain W11	55
5.4.2	Cloning PCR products by ligating into pGEM-T vector	56
5.4.3	Length of the homopolymer of <i>nanH3</i> in <i>Gardnerella piovii</i> strain W11	57
5.4.4	Length of the homopolymer of <i>nanH3</i> in other <i>Gardnerella piovii</i> strains	59
5.5	Localization of sialidases in <i>Gardnerella</i> strains	63

5.6 Localization of sialidase activity in <i>Gardnerella</i> spp.....	65
6.0 Discussion	69
6.1 Putative sialidase gene, <i>nanH3</i> was successfully expressed in <i>E. coli</i> , but the fusion proteins lacked sialidase activity.	69
6.2 The homopolymer region of <i>nanH3</i> varied in length within and between four isolates of <i>Gardnerella piotii</i> (cpn-60 subgroup B).....	71
6.3 Sialidase activity in <i>Gardnerella piotii</i> isolates is associated with the cell wall.....	72
6.4 Future prospects.....	74

List of Tables

Table 4.1: <i>Gardnerella</i> strains used in this study.....	22
Table 4.2: Primers used in this study	25
Table 4.3: All constructs generated in this study	28
Table 5.1: Concentration of protein eluted by Ni-NTA column purification.	48
Table 5. 2: Homopolymer variants in <i>Gardnerella piotii</i> (cpn60 subgroup B) isolates	62

List of Figures

Figure 2.1: Gram-stained vaginal smears from women with normal vaginal flora (A), intermediate vaginal flora (B), or bacterial vaginosis (C).	5
Figure 2.2: The extracellular exo- α -sialidase (SiaBb2) of <i>Bifidobacterium bifidum</i> ATCC 15696.	12
Figure 2.3: Association of PCR presence of <i>nanH1</i> and <i>nanH3</i> with sialidase activity in 112 <i>Gardnerella</i> spp. isolates.	14
Figure 2.4: Predicted domain of putative sialidases encoded by <i>nanH1</i> and <i>nanH3</i>	14
Figure 2.5: Predicted domain structure of NanH1, NanH2 and NanH3 of <i>Gardnerella</i> strain JCP8151B.	15
Figure 2.6: Mechanism of slipped-strand mispairing.....	19

Figure 2.7: Illustration of the homopolymer region in ORF of <i>nanH3</i> in <i>Gardnerella</i> strain W11.	19
Figure 3.1: Illustration of a <i>Gardnerella</i> sp. cell based on our hypothesis.	20
Figure 4.1: Binding sites and directions of primers for <i>nanH3</i> (<i>Gardnerella</i> strain W11) constructs.	26
Figure 4.2: Expression vectors used for this study.....	26
Figure 4.3: Rabbit immunization protocol for NanH1 protein.....	34
Figure 4.4: Recombinant protein constructs built in this study.....	36
Figure 5.1: Predicted domain structure of sialidases of <i>Bifidobacterium bifidum</i> and <i>Gardnerella</i> spp.....	38
Figure 5.2: Sequence alignment of sialidase domains of <i>Gardnerella</i> sp. strain W11 NanH3 and <i>B. bifidum</i> extracellular sialidase.	38
Figure 5. 3 : Predicted domain structures of NanH2, NanH3 of JCP8151B and NanH3 of W11.	39
Figure 5.4: Amplification and cloning of the predicted sialidase domain encoding region of <i>nanH3</i> with an N-terminal His-tag.	40
Figure 5.5: SDS PAGE separated protein profiles of <i>E. coli</i> (pQE-80L+ SiaN3) cell lysates and protein products extracted by Ni-NTA column purification.	41
Figure 5.6: Amplification and cloning of the predicted sialidase domain encoding region of <i>nanH3</i> with an N-terminal GST tag.....	42
Figure 5.7: SDS PAGE separated protein profiles of BL21 <i>E. coli</i> with GST tagged NanH3 sialidase domain.....	43
Figure 5.8: Protein profiles of cell lysates of <i>E. coli</i> BL21 (pET-41a + SiaN3).....	45
Figure 5.9: Protein profiles of <i>E. coli</i> BL21 (DE3) cultures (pET-41a + <i>nanH3</i> sialidase domain) induced at 10 °C and 25 °C.....	46
Figure 5.10: Protein profiles of purification products of glutathione bead purification.	47
Figure 5. 11: Protein profiles of purification products of Ni-NTA column purification.	48
Figure 5. 12: Filter spot test to detect sialidase activity of 2 nd Eluate of Ni-NTA column purification. NC1 – Substrate only, NC2 – buffer only, PC – Recombinant NanH1	49
Figure 5.13: Filter spot test to detect sialidase activity of 2 nd eluate of Ni-NTA column purification.	50

Figure 5.14: PCR amplification of the truncated version of <i>nanH3</i> (TN3).	51
Figure 5.15: PCR screening of transformants after cloning of TN3 with an N-terminal GST tag.	51
Figure 5.16: SDS PAGE separated protein profiles of <i>E. coli</i> (pET-41a+TN3) cell lysates and purified protein products extracted Ni-NTA column purification.	52
Figure 5.17: Solubility of GST+TN3. Protein profiles of <i>E. coli</i> BL21 (pET-41a + TN3) cell pellet and the supernatant indicated that GST fused TN3 was largely insoluble.	53
Figure 5.18: Protein profiles of <i>E. coli</i> BL21 (DE3) culture (pET-41a+TN3) induced at 20 °C. GST fused TN3 (114 kDa) was associated with the pellet.	54
Figure 5.19: Electropherograms showing the poly-C region of <i>nanH3</i> in three colonies of W11.	56
Figure 5.20: PCR amplification of plasmid (pGEM-T+ homopolymer) extractions.	57
Figure 5.21: Electropherograms of the homopolymer regions in two W11 colonies.	58
Figure 5.22: Study design of detecting the length of the homopolymer region of <i>nanH3</i>	60
Figure 5.23: Frequencies of the homopolymer variants in four <i>Gardnerella piovii</i> isolates.	63
Figure 5.24: Western blot analysis of polyclonal antibodies raised in rabbit A and B.	64
Figure 5.25: Western blot analysis of polyclonal antibodies raised in rabbit.	65
Figure 5.26: Localization sialidase activity of <i>Gardnerella</i> spp. isolates.	67
Figure 5.27: Localization sialidase activity of <i>Gardnerella</i> spp. isolates.	68

List of Abbreviations

ARDRA	Amplified ribosomal DNA restriction analysis
ANI	Average nucleotide identity
ATCC	American type culture collection
A, T, G, C	Adenine, Thymine, Guanine, Cytosine
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
BV	Bacterial vaginosis
CAGEF	Centre for the Analysis of Genome Evolution and Function
cpn60	Chaperonin 60
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotides
EMBL_EBI	European Molecular Biology Laboratory_ European Bioinformatics Institute
GST	Glutathione-S-transferase
IPTG	Isopropyl β -D-1-thiogalactopyranoside
KAN	Kanamycin
LB	Lysogeny broth
MALDI-TOF	Matrix assisted-laser desorption ionization time-of-flight
MUNANA	2'-(4-methylumbelliferyl)- α -D-N-acetylneuraminic acid sodium salt hydrate
NanH	N-acetyl neuraminic acid / sialidase
NC	Negative control
NCBI	National Center for Biotechnology Information
NTC	No template control
OD	Optical density
ORF	Open reading frame
PCR	Polymerase chain reaction
PI	Iso-electric point
PVDF	Polyvinylidene difluoride

RFU	Relative fluorescence units
rRNA	Ribosomal ribonucleic acid
SDS-PAGE	Sodium dodecyl sulfate – Poly-acrylamide gel electrophoresis
SiaN3	Sialidase domain of NanH3
SSM	Slipped strand mispairing
TEM	Transmission electron microscopy
UV	Ultra violet

1. Introduction

Large and diverse populations of bacteria, viruses, and fungi occupy almost every surface of the human body (Rup, 2012). Host-microbiota interactions of these native microbial communities are mostly mutually beneficial with the exception of opportunistic pathogens (Ma, Li, & Gotelli, 2019).

The human vaginal microbiota acts as a first line defense by preventing infection by invasive/pathogenic species (Ma et al., 2019). Most healthy vaginal microbial communities are dominated by *Lactobacillus* species which ferment carbohydrates to lactic acid yielding an acidic pH that is intolerable for many pathogenic microbes (Boskey, Cone, Whaley, & Moench, 2001).

Bacterial Vaginosis (BV) is a condition that is characterized by altered microbial composition of the vaginal microbiota and occurs when the healthy microbiota is replaced by BV related bacteria (Hillier et al., 1996). This shift occurs from *Lactobacillus* spp. abundant community to a diverse vaginal flora that includes *Gardnerella* spp., *Bacteroides* spp., *Mobiluncus* spp., and *Mycoplasma hominis* (Nugent, Krohn, & Hillier, 1991; Chaban et al., 2014).

BV is associated with infertility, preterm delivery (Hillier et al., 1996) cervical cancer (Lu et al., 2015) and increased susceptibility to sexually transmitted diseases (Sewankambo et al., 1997). The symptoms of BV also have a negative impact on quality of life, and an associated burden on the healthcare system.

Although the specific microbial etiology of BV is debatable, past studies have confirmed that *Gardnerella* spp. are associated with formation of vaginal BV associated biofilms in patients with persistent BV (Swidsinski et al., 2005).

Despite its clinical significance, many important unanswered questions remain about BV, including why a significant proportion of women with “abnormal” vaginal microbiota do not have clinical signs of BV. One previous study reported that 84% of women meeting the microbiological definition of BV were asymptomatic (Koumans et al., 2007). Part of the explanation for this observation may be phenotypic diversity of *Gardnerella* spp.. What was once considered a single species, *G. vaginalis*, has now been recognized as at least four different species that may have different virulence potential and may play different roles in establishing and maintaining dysbiosis. Describing the phenotypes and understanding the roles of *Gardnerella* spp. in the vaginal

microbiome is necessary in order to better recognize microbial communities that lead to negative health outcomes.

Sialidase is recognised as an important virulence factor found in some *Gardnerella* strains that could possibly degrade the mucus layer (Hardy et al., 2017). Sialidase (also called neuraminidase) is an enzyme that can react with the mucins present in the vaginal fluid by cleaving the glycosidic linkages of sialic acids from terminal glycans present in the vaginal fluid. For years it was assumed that gene responsible for the sialidase activity observed in *Gardnerella* spp. was *nanH1* (Sialidase A/ Gene 1-according to the previous nomenclature in our lab) (Hardy et al., 2017; W. G. Lewis, Robinson, Gilbert, Perry, & Lewis, 2013). Previous work in our lab has shown that *nanH1* was present in not only sialidase positive strains but also in sialidase negative strains. A second sialidase gene, *nanH3* (previously called as Gene 2 in our lab) was later identified, and its presence strongly correlated with extracellular sialidase activity (Patterson, 2018).

Characterization of sialidases in *Gardnerella* spp., will improve understanding of the roles of these diverse bacteria in the vaginal microbiome, and contribute to developing preferable diagnostics for identifying high risk microbiomes and informing treatment decisions.

2. Literature review

2.1 Vaginal microbiome and bacterial vaginosis

The Human Microbiome Project estimated that 90% of the cells of a healthy individual are microorganisms that inhabit surfaces and cavities exposed or connected to the external environment (Rup, 2012). These body sites include skin, mouth, nose, digestive tract, and vagina where ecological communities of microbial species exist as commensals (Ravel et al., 2011).

The microbiota that the vagina harbors, plays a vital role in first line defense by preventing colonization and proliferation of organisms that give rise to various urogenital diseases: bacterial vaginosis, yeast infection, urinary tract infections and sexually transmitted diseases (Hickey, Zhou, Pierson, Ravel, & Forney, 2012). Healthy vaginal microbial communities are usually dominated by *Lactobacillus* species (*Lactobacillus crispatus*, *Lactobacillus iners*, *Lactobacillus gasseri*, *Lactobacillus jensenii*) which ferment carbohydrates to lactic acid yielding an acidic pH that is intolerable for many pathogenic microbes (Boskey et al., 2001).

2.1.1 Bacterial Vaginosis

Bacterial Vaginosis (BV) is a condition that is characterized by altered microbial composition of the vaginal microbiota and occurs when the healthy microbiota is replaced by an overgrowth of mixed aerobic and anaerobic species (Hillier et al., 1996). This shift occurs from *Lactobacillus* spp. dominant community to a diverse vaginal flora with an intense increase (100-1000 fold in total bacterial load) of the commensal anaerobic bacteria including *Gardnerella* spp., *Bacteroides* spp., *Mobiluncus* spp., and *Mycoplasma hominis* (Nugent, Krohn, & Hillier, 1991; Chaban et al., 2014).

BV is associated with infertility, preterm delivery (Hillier, Sharon L; Nugent, Robert P.; Eschenbach, David A.; Krohn, Marijane A; Gibbs, 1996), cervical cancer (Lu et al., 2015) and increased susceptibility to sexually transmitted diseases like HIV AIDS, gonorrhoea, trichomoniasis, and herpes simplex virus type 2 infections (Sewankambo et al., 1997). The prevalence of BV among reproductive-aged women (14 to 49-year-old) in the United States has been estimated to be 29.2%. This accounts 21.7 million women with BV but 84% of cases were reported to be asymptomatic; microbiologically BV positive and clinically negative (Koumans et al., 2007). The

prevalence of BV varies between countries and ethnic groups within the country. In fact, BV prevalence is low in North America with the exception of black women in United States and indigenous women in Canada (Kenyon, Colebunders, & Crucitti, 2013).

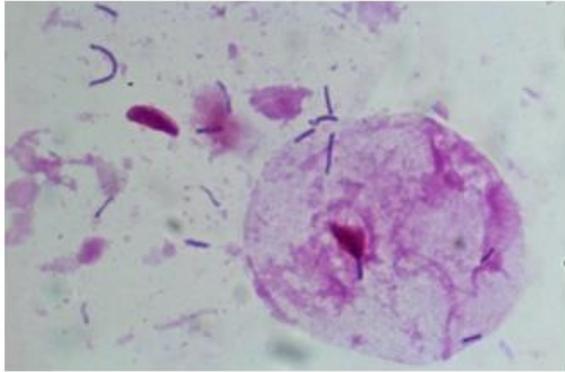
2.1.2 Diagnosis of BV

2.1.2.1 Clinical diagnosis of BV – Amsel’s criteria

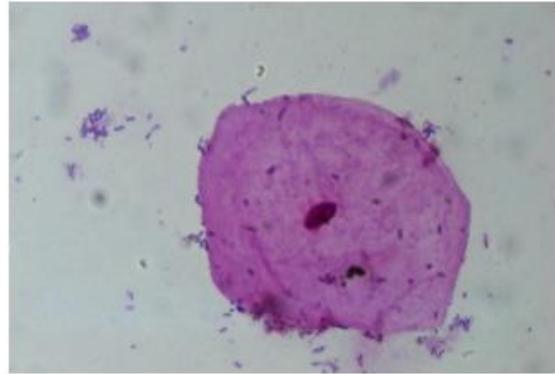
Enzymes produced by BV related bacteria degrade proteins and breakdown amino acids to produce amine compounds. These amines elevate the vaginal pH and give rise to a characteristic fishy odor (Gutman, Peipert, Weitzen, & Blume, 2005; Newton, 1997). Clinical diagnosis of bacterial vaginosis is made primarily based on following criteria; a thin homogeneous discharge, vaginal pH > 4.5, release of an amine odor from vaginal fluid after the addition of KOH, presence of epithelial cells coated with a biofilm of bacteria termed “clue cells”.

2.1.2.2 Microbiological diagnosis

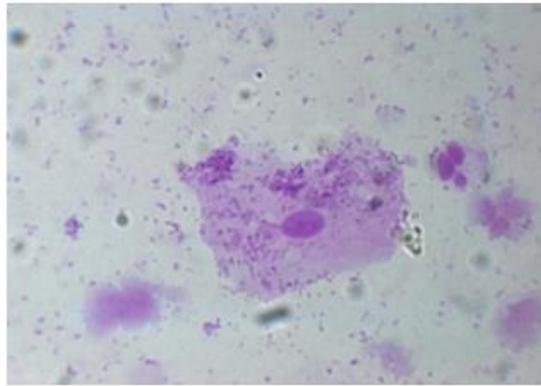
In the laboratory, BV is diagnosed by scoring bacterial morphotypes from a Gram stain (Nugent score) of vaginal fluid. The Nugent score is given based on the presence and the quantity of *Lactobacillus* morphotypes, *Gardnerella* and *Bacteroides* species, and curved, Gram-variable rods (Figure 2.1). A score of 0–3 is defined as not consistent with BV, 4–6 as intermediate, and 7–10 as BV (Nugent et al., 1991). Although this method is considered as the “gold standard” of diagnosis, it does not provide reliable identification and quantification of the specific bacteria involved with BV. In comparison to wet mounts, this method may not represent the actual microbial composition as a result of loss of morphotypes due to fixing and staining (Donders, Vereecken, Dekeersmaecker, & Bulck, 2000). Gram staining does not distinguish other polymicrobial vaginal disorders like aerobic vaginitis and the clinical significance of intermediate scores is not well understood (Hickey et al., 2012).



(A)



(B)



(C)

Figure 2.1: Gram-stained vaginal smears from women with normal vaginal flora (A), intermediate vaginal flora (B), or bacterial vaginosis (C).

(A) 4+ lactobacillus morphotypes, no small Gram-negative or Gram-variable rods (score = 0); (B) 3+ lactobacillus morphotypes and 3+ small Gram-variable rods (score = 4); (C) no lactobacilli and 4+ Gram-negative rods and curved rods (score = 10) (Nugent et al., 1991).

(Photographed by Teenus Jayaprakash)

2.1.2.3 Detection of sialidase activity

Sialoglycoproteins (glycoproteins with terminal sialic acid moieties) are major constituents of mucus and act as a physical barrier, preventing close contact of pathogens with epithelial cells (A. L. Lewis & Lewis, 2012). The presence of mucolytic enzymes like mucinases and sialidases has been highly associated with the adverse health outcomes BV as these enzymes could potentially degrade vaginal mucus (Mcgregor et al., 1994). Sialidases are secreted from BV- associated Gram-negative/Gram-variable rods; *Bacteroides*, *Gardnerella* and *Prevotella* species (Olmsted, Meyn, Rohan, & Hillier, 2003; Von Nicolai, Hammann, Salehnia, & Zilliken, 1984). Smayevsky et al. used filter paper spot test to detect the sialidase activity in non-pregnant women and found that this method was a very good tool to detect BV with 81% sensitivity and 94% specificity (Smayevsky, Canigia, & Lanza, 2001).

The BVBlue system (Gryphus Diagnostics, L.L.C.) is a rapid chromogenic diagnostic test that detects sialidase activity in vaginal fluid specimens. The detection limit of this test is 7.8 units (1 unit = amount of enzyme required to liberate 1 nmol of substrate/ml/minute at 37°C) of sialidase (Myziuk, Romanowski, & Johnson, 2003).

Although BVBlue is a quick and easy method to detect BV, it could give a positive result for a mixed vaginal infection or a yeast infection. Since sialidase activity in the vaginal fluid can also be seen in women who are infected with *Chlamydia trachomatis* and yeast species (Mcgregor et al., 1994), suitability of this test as the sole diagnostic method for BV is questionable.

2.2 Gardnerella spp.

Gardnerella vaginalis was first identified by Leopold (Leopold, 1953) and named as “*Haemophilus vaginalis*” by Gardner and Duke in 1955. Leopold successfully isolated haemolytic, Gram-negative and micro-aerobic “*Haemophilus vaginalis*” from cervical swabs of women, and the organism was believed to be responsible for a characteristic vaginal discharge. Considering the lack of similarity to the genus *Haemophilus*, Zinnemann and Turner suggested removal of *Haemophilus vaginalis* from its genus and renaming as “*Corynebacterium vaginale*” in 1963. Reclassification of *Haemophilus vaginalis* under the genus *Corynebacterium* was solely based on morphological characteristics. In 1966, Reyn and colleagues carried out the first electron microscopy study of *Haemophilus vaginalis*. The authors suggested that the cell wall structure of *Haemophilus vaginalis*

resembled that of a typical Gram-positive organism (Reyn & Birch-Anderson, 1966). Later, Criswell et al. discovered that the microscopic structure and the biochemical properties of the cell wall is more comparable with Gram-negative organisms (Criswell, Marston, Stenback, Black, & Gardner, 1971). Greenwood and Pickett conducted a study to clarify the taxonomic uncertainty of *Haemophilus vaginalis*, using variety of methods: biochemical profiling, DNA hybridization and electron microscopy. The results revealed that “*Haemophilus vaginalis*” showed no correlation or possible clustering with established genera. Thus, they proposed a new genus “*Gardnerella*” and a new species within the genus, “*Gardnerella vaginalis*” (Greenwood & Pickett, 1980).

The genus *Gardnerella* was described as Gram-negative to Gram-variable, facultatively anaerobic, rod shaped bacteria with laminated cell walls that belong to the family of Bifidobacteriaceae. Acetic acid is produced as the end product of fermentation. *Gardnerella* spp. are found in human genital/urinary tract, where it causes vaginitis (Greenwood & Pickett, 1980).

Gardnerella vaginalis is the type species of genus *Gardnerella* ; a Gram-negative to Gram-variable, rod-shaped, anaerobic bacteria (Greenwood & Pickett, 1980; Joesoef, Hillier, Josodiwondo, & Linnan, 1991). Although *Gardnerella* spp. appear to be Gram negative, it was discovered that at the ultrastructural level the cell walls show Gram-positive composition. In 1989, Sadhu et al. provided microscopical and chemical evidence that *Gardnerella vaginalis* was undoubtedly Gram- positive (Sadhu et al., 1989).

2.2.1 Variation within the genus *Gardnerella*

Phenotypic diversity within *Gardnerella vaginalis* was first assessed by Piot et al.. They divided strains isolated from vaginal fluids into eight biotypes based on β -galactosidase, lipase, and hippurate hydrolysis reactions (Piot et al., 1984). In 1986, Benito et al. proposed a scheme of biotyping with respect to detection of hippurate hydrolysis, β -galactosidase and lipase, and fermentation of arabinose, galactose and xylose. They discovered 17 biotypes among 197 *Gardnerella* strains and some biotypes were more frequently found in BV patients than others (Benito, Vazquez, Berron, Fenoll, & Saez-Nieto, 1986).

Observations of phenotypic diversity inspired researchers to investigate genotypic diversity. Ingianni and colleagues analysed 34 strains of *Gardnerella vaginalis* with different ribotyping

methods and concluded that amplified ribosomal DNA restriction analysis (ARDRA) was the most suitable typing method to distinguish subtypes. The authors used ADRA to establish the presence of at least 3-4 genotypes within this species (Ingianni, Petruzzelli, Morandotti, & Pompei, 1997). The presence of four clusters based on partial chaperonin-60 (cpn60) sequences was observed by Hill et al. in 2005 indicating intraspecies variation within "*Gardnerella vaginalis*" (Hill et al., 2005). In 2011, Santiago et al. confirmed the existence of three genotypes by genotyping 134 *Gardnerella vaginalis* isolates using random amplified polymorphic DNA (RAPD) and ARDRA. They also assessed sialidase production and unveiled that only two of these distinct ARDRA genotypes produce sialidase (Lopes Dos Santos Santiago et al., 2011). Our research group has resolved that this species consists of four genotypically distinct groups (subgroup A – D) with potentially different virulence based on cpn-60 gene sequencing (Jayaprakash, Schellenberg, & Hill, 2012). Support for the definition of four subgroups within *Gardnerella* came from a whole genome sequencing study. Based on neighbour-grouping analyses using gene possession data and core gene allelic data from 17 isolates, Ahmed et al. inferred the presence of four genetically isolated groups (Ahmed et al., 2012). They showed that each group had a characteristic genome size, G/C (guanine to cytosine ratio) content and there was evidence of homologous recombination within the group. They also suggested that these four non-recombining clades of organisms with distinct genetic properties and possible ecological variations, could be treated as "separate species" (Ahmed et al., 2012). The *Gardnerella* clades were subsequently shown to correspond with the cpn60-defined subgroups (Schellenberg et al., 2016). More recently, by analysing 81 full genome sequences using digital DNA-DNA hybridization (dDDH) and average nucleotide identity (ANI) Vaneechoutte et al. confirmed the presence of at least 13 different taxa within the species *Gardnerella vaginalis*. They proposed an amendment of the description of *Gardnerella vaginalis* and also defined three new species within the genus *Gardnerella*: *Gardnerella leopoldii*, *Gardnerella piotii*, *Gardnerella swidsinskii* (Vaneechoutte et al., 2019).

2.2.2 Description of new species of genus *Gardnerella*

New species were defined under the genus *Gardnerella*, taking genotypic and phenotypic characteristics into account. The genus *Gardnerella* newly defined as a taxon that consists Gram-negative to Gram-variable, facultative anaerobic, rod-shaped organisms (Greenwood & Pickett, 1980) with a GC content ranges between 38% to 43.4% (Vaneechoutte et al., 2019).

Gardnerella vaginalis was redefined with modifications added to the original description by Gardner and Dukes in 1955 and Greenwood and Pickett in 1980 (Greenwood & Pickett, 1980). *G. vaginalis* is β -galactosidase positive and has a G+C content of 41-42.8% (Vaneechoutte et al., 2019). *Gardnerella leopoldii* is β -galactosidase and sialidase activity negative and has a G+C content of 41.9-43.2%. *Gardnerella piovii* is defined as group of organisms with sialidase activity and no β -galactosidase with a G+C content of 41.1-42.3%. *Gardnerella swidsinskii* does not show sialidase or β -galactosidase activity and has a G+C content of 41.1-42.3%. All these defined species can be unambiguously distinguished from each other based on digital DNA-DNA hybridization (dDDH), average nucleotide identity (ANI) and matrix assisted-laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry profiles, but not by 16S rRNA gene sequences (Vaneechoutte et al., 2019).

The alignment of cpn60 based subgroups and the newly defined species was demonstrated by Hill et al.. *G. piovii* and *G. vaginalis* corresponded to cpn60 subgroup B and C respectively. *G. leopoldii* and *G. swidsinskii* were previously grouped together as subgroup A based on cpn60 sequences of the isolates available at that time. Isolates belonging to cpn60 subgroup D corresponded to three distinct genome species: genome species 8, 9 and 10 (Hill & Albert, 2019).

2.2.3 Virulence factors of *Gardnerella* spp.

Although bacterial vaginosis is a polymicrobial disease, *Gardnerella* spp. are strongly associated with BV (Criswell, Ladwig, Gardner, & Dukes, 1969), suggesting that at least some *Gardnerella* spp. have pathogenic potential. Studies have shown that *Gardnerella* spp. are armed with various virulence factors that could possibly give rise to negative outcomes of BV: biofilm formation (Swidsinski et al., 2005), sialidase activity (Jayaprakash et al., 2012) and secretion of a pore-

forming toxin known as vaginolysin that lyses vaginal epithelial cells and human red blood cells (Cauci et al., 1993; Gelber, Aguilar, Lewis, & Ratner, 2008).

2.3 Sialidases

Sialidases or neuraminidases are enzymes that react with sialo-glycoconjugates and cleave the terminal sialic acid moieties.

The first study to identify the enzymatic activity of sialidase was a virology-related study carried out by Hirst in 1941 (Hirst, 1941). The enzyme was later discovered by Burnet et al. in filtrates of *Vibrio cholerae* and *Clostridium perfringens*. The authors named this enzyme as “receptor-destroying enzyme”, since the crude extracts were able to destroy the binding sites on red blood cells for influenza virus (Burnet, McCrea, & Stone, 1946). In 1956, Heimer and Meyer demonstrated the release of sialic acid, by hydrolysis of mucus using pneumococcal extracts and suggested the name “sialidase” for this enzyme (Heimer & Meyer, 1956). A year later, Gottschalk proposed the name “neuraminidase” for the same enzyme (Gottschalk, 1957) and both names have been used thenceforth. In 1959, Ada and French were able to purify sialidase enzyme from *V. cholerae* by chromatography (Ada & French, 1959).

Sialidases are produced by many pathogenic bacterial species indicating a significant role in pathogenicity. However, sialidase is also present in non-pathogenic bacteria, fulfilling a nutritional role (A. P. Corfield, Wagner, Clamp, Kriaris, & Hoskin, 1992).

Sialidases are produced by a wide spectrum of pathogens, along with their other virulence factors (T. Corfield, 1992). Extracellular sialidase activity may be secreted or cell bound. May and Brown observed secreted sialidase activity in pathogenic canine *Mycoplasma* spp.: *M. cynos* and *M. canis* (May & Brown, 2009). *Streptococcus pneumoniae* and *Bifidobacterium bifidum* produce sialidases that are tethered to the cell wall (Camara, Boulnois, Andrew, & Mitchell, 1994; Nishiyama et al., 2017). *Clostridium perfringens* produces a “small” sialidase molecule (41 kDa) that is not secreted along with a larger (71 kDa), phylogenetically unrelated, secreted sialidase (Roggentin, Kleineidam, & Schauer, 1995).

Although the precise roles of sialidase enzyme in pathogenesis are yet to be clarified, it has been suggested that sialidases act as generalised virulence determinants by breaking down sialylated glycoproteins like mucins and exposing cryptic carbohydrate binding sites for other enzymes or

invading organisms. Sialidase secreted by *Vibrio cholerae*, plays a subtle but a distinct role in enabling the cholerae toxin to access the receptors of epithelial cells (Galen et al., 1991). Cell-bound sialidase in *Bifidobacterium* is involved in surface adhesion and carbohydrate assimilation (Nishiyama et al., 2017). Furthermore, production of sialidase may be important in liberating sialic acids from glycoproteins and providing carbohydrate sources for the bacterial cells to uptake. (Roggentin et al., 1995).

2.3.1 Extra cellular sialidase of a closely related species, *Bifidobacterium bifidum*

Bifidobacteria are health-promoting bacteria, abundantly colonized in the intestine and also found in the oral cavity and vagina (Mattarelli, 2006). *Bifidobacterium bifidum* and *Gardnerella* spp. belong to the family of Bifidobacteriaceae indicating the phylogenetic relatedness of the two genera. Genes encoding sialidases have been detected in genomes of some *Bifidobacterium* strains, including *B. longum*, *B. breve* and *B. bifidum* (Egan, Motherway, Ventura, & Sinderen, 2014; Sela et al., 2008; Turrone et al., 2010).

Turrone et al. identified two genes predicted to encode exo- α -sialidase in the genome of *Bifidobacterium bifidum* PRL 2010 (Turrone et al., 2010). In 2011, Kiyohara et al. cloned the exo- α -sialidase (siabb2) gene from *Bifidobacterium bifidum* JCM1254, which encoded an 835-amino-acid protein (SiaBb2) with a molecular mass of 87 kDa (Kiyohara et al., 2011). SiaBb2 consists of an N-terminal signal peptide, sialidase catalytic domain, alpha-galactosidase domain and a C-terminal transmembrane domain suggesting that the protein is extracellular and anchored to the cell wall (Figure 2.2) (Kiyohara et al., 2011). The role of this protein as suggested by Kiyohara and colleagues, is to metabolize sialylated human milk oligosaccharides. In 2017, Nishiyama et al. demonstrated that SiaBb2 in *Bifidobacterium bifidum* was involved not only in acquisition of nutrients but also adhesion to the mucosal surface (Nishiyama et al., 2017).

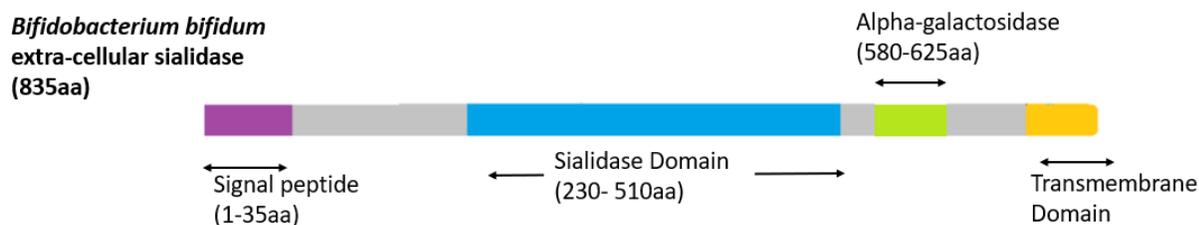


Figure 2.2: The extracellular α -sialidase (SiaBb2) of *Bifidobacterium bifidum* ATCC 15696.

Amino acid numbering indicates the initiation and termination of codons of the domains predicted by InterPro. SiaBb2 consists of a signal peptide (purple box), sialidase domain (blue box), alpha-galactosidase domain (green box) and a transmembrane domain (orange box). (Nishiyama et al., 2017).

2.3.2 Genes encoding sialidases in *Gardnerella* spp.

Sialidase activity is identified as a virulence factor of *Gardnerella* spp. that contributes to degradation of the vaginal mucus in BV patients (Briselden, Moncla, Stevens, & Hillier, 1992) (W. G. Lewis et al., 2013). A sialidase enzyme from *Gardnerella* spp. was first discovered and purified by Von Nicolai and colleagues, in 1984. They detected sialidase activity in the bacterial sediment and the culture medium. The enzyme was liberated by sonication and purified by 60-80% ammonium sulfate precipitation and column chromatography (Von Nicolai et al., 1984). Isolates of *Gardnerella* *piotii* and genome species 3 (cpn60 subgroup B) can liberate sialic acid from sialylated glycoproteins by expressing sialidase. All other species of *Gardnerella* are sialidase negative. In sialidase positive strains the catalytic activity is mostly cell-associated (W. G. Lewis et al., 2013).

Santiago et al. identified a putative sialidase gene, sialidase A (Gene1/*nanHI*) as the gene responsible for the sialidase activity observed in *Gardnerella* strains. The gene *nanHI* was originally annotated in strain ATCC 14019 (Lopes Dos Santos Santiago et al., 2011). When the authors investigated the presence of the putative sialidase gene (*nanHI*), they could not correlate observed sialidase activity with the *nanHI*. In 2016, Schellenberg and colleagues from the Hill lab investigated the distribution of this putative sialidase gene (*nanHI*) sequence and sialidase activity among the four cpn60 defined-subgroups (subgroup A-D). Although all isolates of subgroup B, C and D were *nanHI* positive, sialidase activity was observed only in subgroup B isolates and 9% of the subgroup C isolates (Schellenberg et al., 2016). *NanHI* (Accession RDW98445.1) appears to

be present in all sialidase positive strains but, also present in sialidase negative strains raising the question whether *nanHI* accounts for the sialidase activity observed in *Gardnerella* spp.. Our lab group analyzed 39 whole genome sequences and found that *Gardnerella* spp. potentially produces two sialidases encoded by two different genes; *nanHI* and *nanH3* (Accession RDW96829.1, previously known as Gene 2). In contrast to *nanHI*, the presence of *nanH3* is strongly correlated with sialidase activity (Figure 2.3) (Patterson, 2018).

The predicted protein encoded by *nanH3* of *Gardnerella* sp. strain W11 consists of a signal peptide, a sialidase domain and a trans-membrane domain (Figure 2.4) showing a structure similar to the exo- α -sialidase (SiaBb2) encoded by closely related species, *Bifidobacterium bifidum*. The protein encoded by *nanHI* has a predicted lectin binding domain and a sialidase domain (Figure 2.4) (Patterson, 2018). Lack of a signal peptide or any extracellular domain, suggest that product of *nanHI* is an intracellular sialidase.

		nanH1	
		PCR	
		+	-
Activity	+	36	0
	-	41	35

		nanH3	
		PCR	
		+	-
Activity	+	35	1
	-	0	74

(A)
(B)

Figure 2.3: Association of PCR presence of *nanH1* and *nanH3* with sialidase activity in 112 *Gardnerella* spp. isolates.

3(A); Out of 112 isolates, 36 were *nanH1* positive and sialidase activity positive, 41 were *nanH1* positive but sialidase activity negative, 35 were *nanH1* negative and sialidase activity negative. 3(B); Out of 112 isolates, 35 were *nanH3* positive and sialidase activity positive, 1 isolate was *nanH3* negative but sialidase activity positive, 74 isolates were *nanH3* negative and sialidase activity positive (Patterson, 2018).

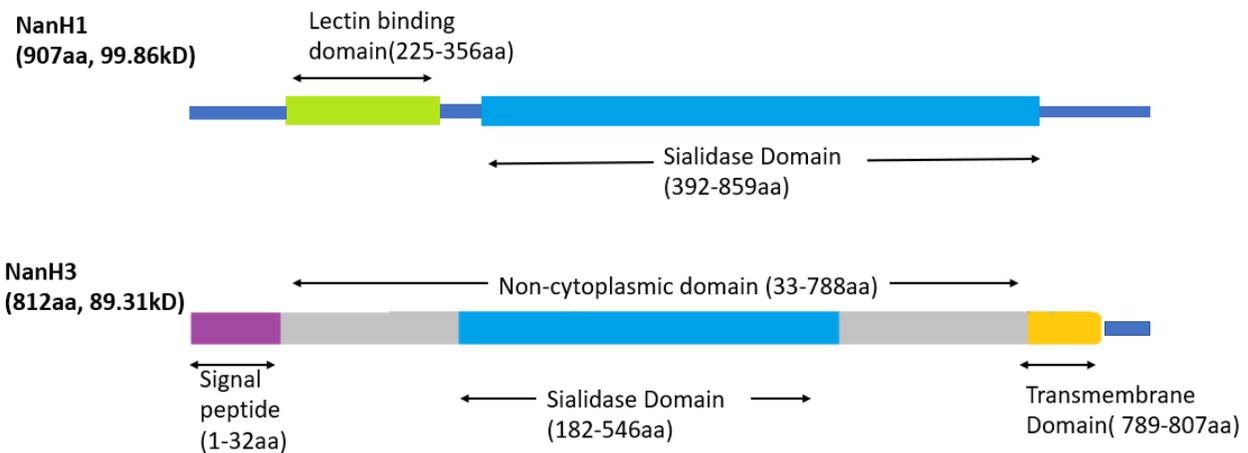


Figure 2.4: Predicted domain of putative sialidases encoded by *nanH1* and *nanH3*.

The predicted protein encoded by *nanH1* of *Gardnerella* strain W11 (Accession RDW98445.1) consists of a lectin binding domain and a sialidase domain. The predicted protein encoded by *nanH3* of *Gardnerella* strain W11 (Accession RDW96829.1) consists of a signal peptide, sialidase domain and a transmembrane domain.

Recently Robinson et al. identified two putative sialidase genes (*nanH2* and *nanH3*) in *Gardnerella vaginalis* in addition to *nanH1*. A BLASTp search using *Bifidobacterium longum* ATCC 15697 *nanH2* as a query sequence revealed three sialidase homologs in a sialidase-positive *Gardnerella* strain, JCP8151B (*G. piovii*, cpn60 subgroup B). The catalytic activities of NanH2 and NanH3 expressed in *E. coli*, were higher than NanH1 against the fluorogenic substrate 4-Methylumbelliferyl-N-acetylneuraminic acid. Either or both Genes *nanH2*, and *nanH3* were detected in 15 sialidase positive strains. The authors concluded that NanH2 and NanH3 are the proteins responsible for sialidase activity shown in BV patients (Robinson, Schwebke, Lewis, & Lewis, 2019). The *nanH3* sequence of *Gardnerella* strain JCP8151B shows the highest similarity to *nanH3* in strain W11 but it lacks a signal peptide. Although the authors assumed that NanH2 and NanH3 were extra cellular, membrane bound sialidases, the cellular localization of the sialidases was not demonstrated.

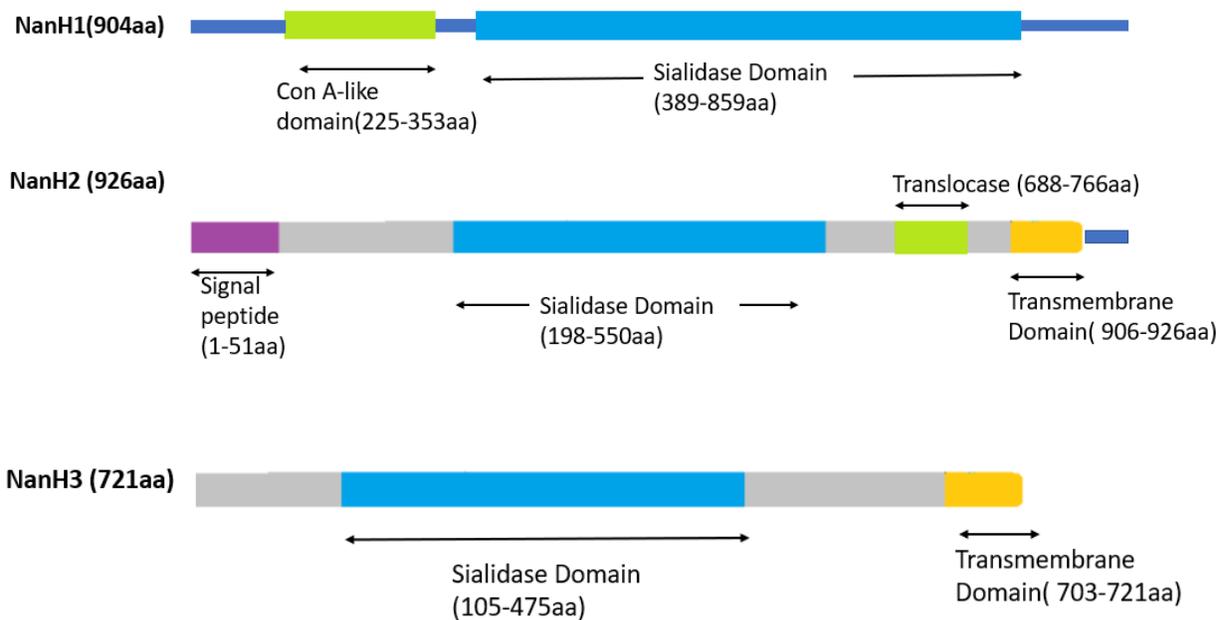


Figure 2.5: Predicted domain structure of NanH1, NanH2 and NanH3 of *Gardnerella* strain JCP8151B.

The putative catalytic domains are indicated in blue boxes. ConA – Concanavalin A. NanH1 consists of a ConA domain and a sialidase domain. NanH2 consist of a signal peptide, a sialidase domain, a translocase domain and a transmembrane domain. NanH3 consists of a sialidase domain and a transmembrane domain (Robinson et al., 2019).

2.4 Phase variation in bacteria

Pathogenic microorganisms undergo enormous pressure in infecting and colonizing their hosts. The surface receptors and molecules they use to bind and anchor themselves to the host can also be detrimental to them as the host immune system recognizes these antigens and eradicates the pathogens. As a survival strategy, microorganisms may express hypervariable surface molecules to hide from the host immune system or develop mechanisms to avoid being destroyed by the host, once recognized. Microorganisms from different lineages have evolved similar survival strategies for immune evasion. One such strategy is phase variation (Deitsch, Lukehart, & Stringer, 2009).

Phase variation is a form of phenotypic variation where the expression of two phenotypes (phases) alternates between “on” state and “off” state in a heritable and a reversible manner. This makes one phase variant differ from another by presence or absence of a certain cell surface marker (Deitsch et al., 2009). This switch is a stochastic event and results a diverse clonal population. After a clonal division, a majority of daughter cells will inherit the expression phase of the parent (template) and a minority will have the switched phase. The switching frequency can be modulated by external factors and the proportion of a phase variant depends on how favourable the phase is for survival in that environment (Woude & Baumler, 2004).

2.4.1 Mechanisms of phase variation

Common genetic mechanisms involved in phase variation are DNA inversion and slipped-strand mispairing (SSM) during replication. Phase variation in bacteria was first detected in 1922 (Andrewes, 1922) and more than 50 years later, Zieg et al. carried out the first study to identify the mechanism of phase variation in bacteria. They proposed that DNA inversion in a flagella gene was the phase-determinant of flagellar gene expression in *Salmonella* (Zieg, Silverman, Hilmen, & Simon, 1977). Phase variation has also been extensively studied in pathogenic *Neisseria* spp., *Neisseria meningitidis* and *Neisseria gonorrhoeae* (Hagblom, Segalt, & Billyardt, 1985; Seifert, Ajioka, Marchal, Sparling, & So, 1988) where the variation is generated through slipped-strand mispairing (SSM).

2.4.1.1 Phase variation through slipped-strand mispairing (SSM)

During replication, the template or the daughter strand loop-out causing a mismatch between the strands (Figure 2.6). Slipped-strand mispairing (SSM) can occur in genomic regions that contain short, homogenous or heterogenous repeats and as a result of it, the length of the repeat region at each replication may change. This could lead to consequent changes in the transcription or translation product of a gene depending on whether the SSM occurs within the open reading frame or extragenic regions such as promoter sequences (Woude & Baumler, 2004).

Phase variation through translational regulation was first described in *opa* genes that encode opacity surface proteins of *Neisseria* spp. (Stern, Brown, Nickel, & Meyer, 1983) The open reading frames (ORF) of *opa* genes contain CTCTT pentamer repeats within the signal peptide encoding region and the expression of the Opa protein is regulated by slipped-strand mispairing. With six, nine or twelve CTCTT repeats the initiation codon is in frame with the remaining *opa* gene translating the Opa protein. Four or eight coding repeats makes the initiation codon out of frame with rest of the *opa* codons. The phase “on” state allows them to adhere to specific surface receptors on host cells (Stern, Nickel, & Meyer, 1984).

The *Treponema pallidum* repeat (*tpr*) gene family encodes proteins that are predicted to be in the outer membrane. Some genes in this family have homopolymeric guanosine (G) repeats of lengths varying from 7 bp to 12 bp, immediately upstream of the transcriptional start site (Claire, Steven, & George, 1998) and the length of the homopolymer affects the expression of the *tpr* genes. When there are eight Gs or fewer, the genes are transcribed whereas with longer poly-G tracts, transcription is reduced by 95-100% as the promoter is not activated. Change in the length of the homopolymer is thought to occur through slipped-strand mispairing during replication (Giacani, Lukehart, & Centurion-lara, 2007).

2.4.2 Is nanH3 a candidate for phase variation?

Genes with short tandem repeat regions are potentially susceptible to SSM during replication. Interestingly, *nanH3* in the *Gardnerella* strain, W11 contains a poly-cytosine repeat within the open reading frame. This homopolymer lies between 73-84bp from the start codon and is approximately 12 bp long. Based on previous studies, the critical threshold length of the

homopolymer is 7-10 bp for any repeat region to undergo SSM (Dechering, Cuelenaere, Konings, & Leunissen, 1998). A search for other homopolymer sequences within the whole genome sequence of W11 identified no homopolymer in any other genes of this strain. The gene *nanH3* encodes a protein that is predicted to be extracellular. This raises the question if the expression of *nanH3* is regulated by slipped-strand mispairing. If this gene is subjected to SSM, different lengths of the homopolymer region would be observable within a population.

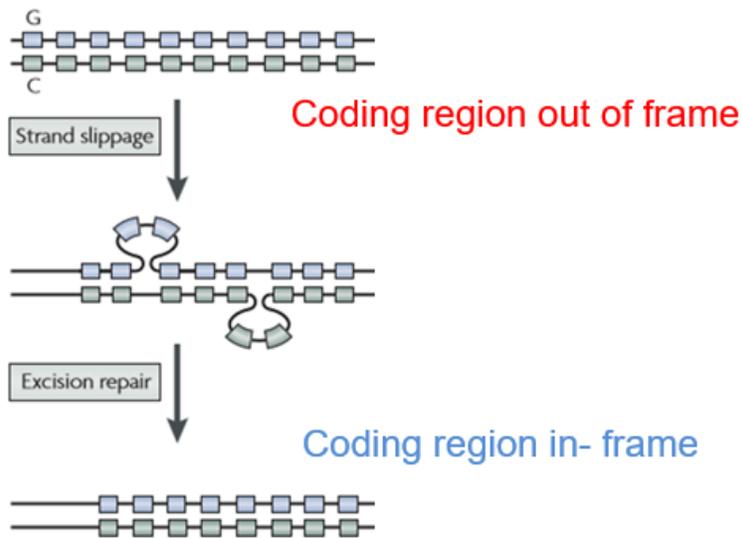


Figure 2.6: Mechanism of slipped-strand mispairing.
 Modified from Deitsch et al (Deitsch et al., 2009).

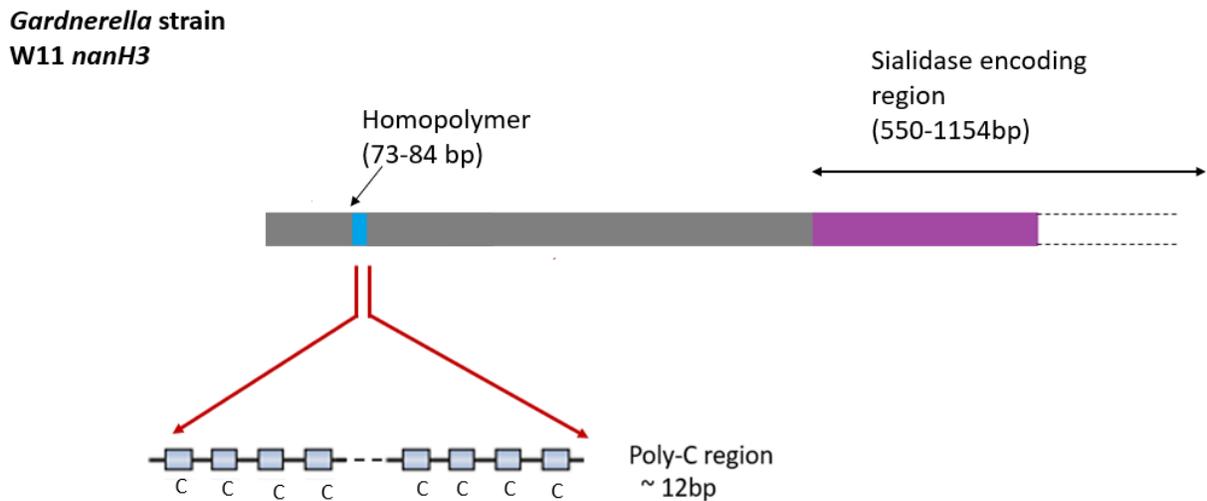


Figure 2.7: Illustration of the homopolymer region in ORF of *nanH3* in *Gardnerella* strain W11.
 The homopolymer has about 12 bp of cytosine repeats. Modified from (Deitsch et al., 2009).

3.0 Objectives

We hypothesize that *nanH1* encodes a sialidase with an intracellular activity and *nanH3* encodes an extracellular sialidase that could be subjected to phase variation.

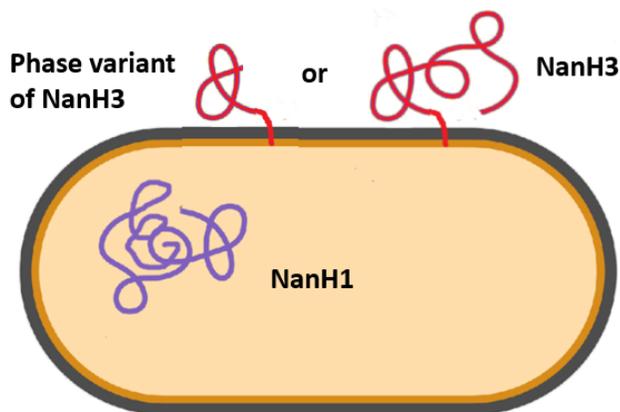


Figure 3.1: Illustration of a *Gardnerella* sp. cell based on our hypothesis.

Our objectives are

1. To determine if *nanH3* encodes a protein with sialidase activity.
2. Determine if *nanH3* in *Gardnerella* sp. strain W11 is regulated by slipped-strand mispairing.
3. To determine the localization of sialidase enzymes encoded by *nanH1* and *nanH3*.

4.0 Methods

4.1 Protein domain identification and sequence alignment

Conserved domains in protein sequences were identified by NCBI Conserved Domain search (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). DNA and protein sequence alignments were performed with Clustal Omega (EMBL_EBI) (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) and NCBI BLAST (basic local alignment search tool) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). InterPro protein sequence analysis and classification tool (<https://www.ebi.ac.uk/interpro/>) was used to predict the domains of amino acid sequences encoded by putative genes.

4.2 DNA and protein sequencing

Sanger sequencing of plasmid constructs and PCR products were carried out by Macrogen, Korea. Universal vector primers (T7 and Sp6) or gene specific primers (Table 4.2) were used for Sanger sequencing. Protein sequence identification was determined by coupled liquid chromatography-mass spectrometry (LC/MS) followed by comparison of the resulting peptide sequences to NCBI and Uniprot databases (performed by the Centre for the Analysis of Genome Evolution and Function (CAGEF), University of Toronto, ON).

4.3 Bacterial strains and culture conditions

Gardnerella strains (Table 4.1) from our -80 °C freezer stocks were used in the study. *G. piovii* isolate W11 was grown on Columbia Sheep Blood Agar (BBL, Becton, Dickinson and Company, Sparks, MD, USA) plates incubating at 37 °C for 48 hours with anaerobic BD GasPak EZ (Becton, Dickinson and Company, Sparks, MD, USA). For broth cultures a few colonies from the plate were collected with a 10 µl inoculation loop and used to inoculate NYC III (per litre: 2.4 g HEPES, 15 g Proteose peptone, 3.8 g Yeast extract, 5 g NaCl, 5 g Glucose). Broth cultures were incubated at 37 °C for 48 hours in anaerobic conditions.

E. coli Top 10 was used for cloning *nanH3* of W11 into pET-41a or pQE-80L and *E. coli* BL21 (DE3) was used for recombinant protein expression. *E. coli* DH5α cells were used to prepare clone libraries after ligation of *nanH3* homopolymer region of *Gardnerella* subgroup B strains with pGEM-T vector (Promega, Madison, WI, USA). All *E. coli* were grown in Lysogeny Broth (LB)

while shaking at 37°C, with antibiotic selection as required. For plasmid maintenance, ampicillin (100 µg/ml) or kanamycin (50 µg/ml) was used. Isopropyl β-D-1-thiogalactopyranoside (IPTG) methowas used to induce gene expression in *E. coli*.

For *E. coli* colony selection by blue-white screening, X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was used as the chromogenic substrate.

Table 4.1: *Gardnerella* strains used in this study

<i>Gardnerella</i> Strain	Species	Sialidase activity	<i>nanH1</i>	<i>nanH3</i>	cpn60 Subgroup
W11	<i>G. piovii</i>	+	+	+	Subgroup B
VN014	<i>G. piovii</i>	+	+	+	Subgroup B
VN015	<i>G. piovii</i>	+	+	+	Subgroup B
NR032	<i>G. piovii</i>	+	+	+	Subgroup B
VN001	<i>G. vaginalis</i>	-	+	-	Subgroup C
VN007	<i>G. vaginalis</i>	-	+	-	Subgroup C
NR038	<i>G. vaginalis</i>	-	+	-	Subgroup C
NR042	<i>G. vaginalis</i>	-	+	-	Subgroup C

4.4 DNA manipulation

4.4.1 PCR amplification of putative functional domain of *nanH3* / truncated *nanH3* and DNA ligation

Gardnerella strain W11 was cultured in NYC III media for 48 hours. The cells were harvested by centrifuging at $13\,000 \times g$ for 10 minutes. Genomic DNA was extracted from W11 cells using a modified salting out method (Martin-Platero, Valdivia, Maqueda, & Martinez-Bueno, 2007). Extracted genomic DNA was quantified by ND-1000 NanoDrop Spectrophotometer (NanoDrop Technologies, USA).

Primers were designed for the predicted sialidase domain of *nanH3* (from 550 to 1554 in the open reading frame) with restriction sites to facilitate ligation to pQE-80L vector (Figure 4.1). Forward primer JH0766 was designed with Sph I site and the reverse primer JH0767 was designed with KpnI site (Table 4.2).

To ligate the amplicon into the pET-41a vector, primers JH 0770 and JH 0771 (Table 4.2) were designed for the predicted sialidase domain of *nanH3* (from 550 to 1554 in the open reading frame) with restriction sites KpnI and SacI (Figure 4.1)

Primers JH0799 and JH0800 (Table 4.2) were designed with restriction sites to amplify *nanH3* without the signal peptide and transmembrane domain encoding sequences for the purpose of ligating the amplicon into the expression vector pET-41a.

Portions of the *nanH3* sequence were amplified by polymerase chain reaction (PCR) in a Mastercycler Pro 6321 (Eppendorf AG, Hamburg, Germany). The components of the PCR reaction mix (50 μ l per reaction) were added to achieve final concentrations of 1X High Fidelity PCR buffer (60 mM Tris-SO₄ (pH 8.9), 18 mM (NH₄)₂SO₄), 0.2 mM dNTP mix, 2 mM MgSO₄, 1U Platinum High Fidelity (Hi-Fi) proof-reading Taq polymerase (Invitrogen, Carlsbad, CA, USA) and less than 500 ng of genomic DNA template. The thermocycler was set to following parameters; initial denature at 94 °C for 2 minutes, 35 °C cycles of denature at 94 °C for 15 seconds, annealing at 55 °C for 30 seconds, extension at 68 °C for 2 or 3 minutes, and final extension for 5 minutes. PCR products were visualized by running on a 1% agarose gel along with GeneRuler ladder DNA mix (Thermo Scientific, Lithuania).

To ligate amplicon into expression vector pQE-80L or pET-41a (Figure 4.2), amplicons and the vectors were digested with restriction enzymes Sph I and Kpn I or Kpn I and Sac I respectively. Restriction double digestion was carried out by mixing 1 µg of DNA (PCR amplicon or plasmid), 10 units of each restriction enzyme (Sph I and Kpn I or Kpn I and Sac I), 1X 1.1 NEbuffer (10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂ 100 µg/ml, BSA pH 7.0). The digestion reaction (50 µl) was incubated in a 37 °C water bath for one hour. The digested DNA was purified using QIAquick PCR Purification Kit (Qiagen, Mississauga, ON). The purified PCR amplicon and plasmids with sticky ends were added in 4:1 molar ratio to the ligation master mix containing 1 µl of T4 DNA ligase (Promega, Madison, WI, USA), 1X ligation buffer (30 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM DTT and 1 mM ATP). The total volume of the ligation reaction mix was 10 µl and the amplicons and the plasmids were ligated overnight at 4 °C.

4.4.2 Transformation of *E. coli*

The ligation reaction mixture was then used to transform chemically competent One shot Top 10 *E. coli* (Invitrogen, Carlsbad, CA, USA). A 50 µl vial of One shot Top 10 chemically competent *E. coli* cells was thawed on ice. A volume of 5 µl of the ligation reaction was added directly into the vial of competent cells and mixed by tapping. The mixture was incubated on ice for 30 minutes. The vial was incubated in a 42 °C water bath for exactly 30 seconds and placed on ice. To the vial, 250 µl of prewarmed SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, and 20 mM glucose) was added and incubated at 37 °C for exactly 1 hour at 225 rpm in a shaking incubator. After 1-hour incubation, 50-200 µl of transformed culture was spread on each LB agar plate with ampicillin or kanamycin. The plates were inverted and incubated at 37 °C overnight.

Ten randomly selected colonies of each construct were sub-cultured on a “patch plate” and screened by PCR with amplicon specific primers (to assess the presence of the insert). A colony of each construct with the correct sized insert was sequenced to confirm if the insert was in-frame with the N-terminal 6xHistidine tag or GST-tag of the vector (Figure 4.2).

To facilitate expression of GST fusion protein, the plasmid constructs (pET41a+insert) were transferred into *E. coli*. BL21 (DE3). *E. coli* Top10 colony with the correct insert was used to inoculate 5 ml of LB media with kanamycin. Cells were grown overnight in a 37 °C shaking

incubator at 225 rpm. Overnight culture was centrifuged at $10000 \times g$ for 5 mins and harvested cell pellet was used to extract plasmids using QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany). A 50 μ l vial of chemically competent *E. coli* BL21(DE3) cells was thawed on ice. A volume of 5 μ l of the extracted plasmid was added directly into the vial of competent cells and mixed by tapping. The mixture was incubated on ice for 30 minutes. The vial was incubated in a 42 °C water bath for exactly 30 seconds and placed on ice. To the vial, 250 ml of prewarmed S.O.C medium was added and incubated at 37 °C for exactly 1 hour at 225 rpm in a shaking incubator. After 1-hour incubation, 50- 200 μ l of transformation vial was spread on each LB + kanamycin agar plate. The plates were inverted and incubated at 37 °C overnight. A selected transformant was used to inoculate LB + kanamycin broth and grown overnight at 37 °C, 225 rpm in a shaking incubator. The culture was mixed with LB +20% (v/v%) glycerol in 1:1 ratio and stored at -80°C.

Table 4.2: Primers used in this study

Target	Recipient vector	Restriction sites (underlined)	Primer name	Sequence	Amplicon size (bp)
<i>nanH3</i> sialidase domain (550-1554)	pQE-80L	SphI KpnI	JH0766 JH0767	<u>GGCATGCGT</u> GGAAGGACTAACTGGAAAACC GCGGT <u>ACCAGAAATTCTCGT</u> TATACATTATG	1197
<i>nanH3</i> sialidase domain (550-1554)	pET-41a	KpnI SacI	JH0770 JH0771	<u>GCGGTACCGT</u> GGAAGGACTAACTGGAAAACC CAGG <u>GAGCTCGAAATTCTCGT</u> TATACATTATG	1197
Truncated <i>nanH3</i> (121- 2340)	pET-41a	KpnI SacI	JH0799 JH0800	<u>GCGGTACCACTGA</u> ACCAGTAACATTTAC CAGG <u>GAGCTCATCATTCC</u> TCTATTCTTTG	2236
Homopolymer region of <i>nanH3</i> (2-310)	pGEM-T	NA	JH0780 JH0781	ATGATTGGAACAGCGCATAAAG GATTTCTCCACCTACAGTTACC	308

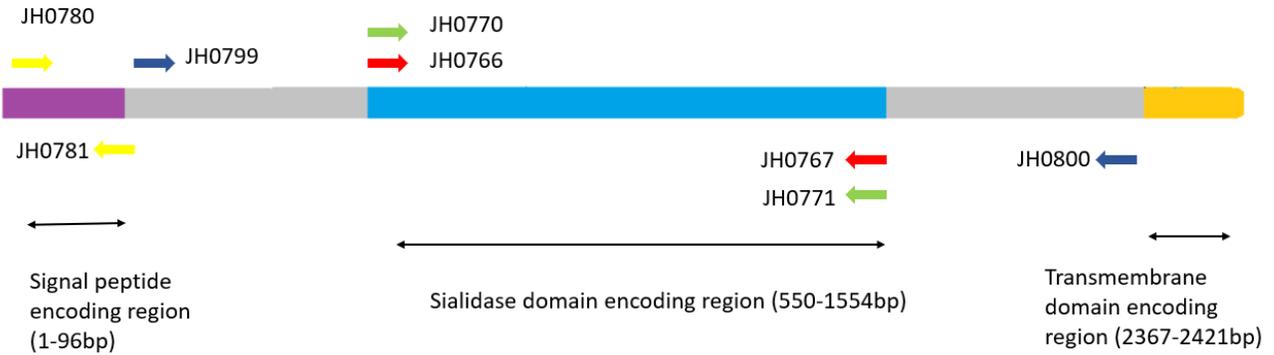


Figure 4.1: Binding sites and directions of primers for *nanH3* (*Gardnerella* strain W11) constructs.

Each primer pair used for amplification is indicated in the same color arrows (yellow/blue/red/green). Colored boxes represent domains of *nanH3*: signal peptide encoding region (purple), sialidase domain encoding region (blue), transmembrane domain encoding region (orange). Numbering is based on the *nanH3* sequence of *G. piovii* isolate W11.

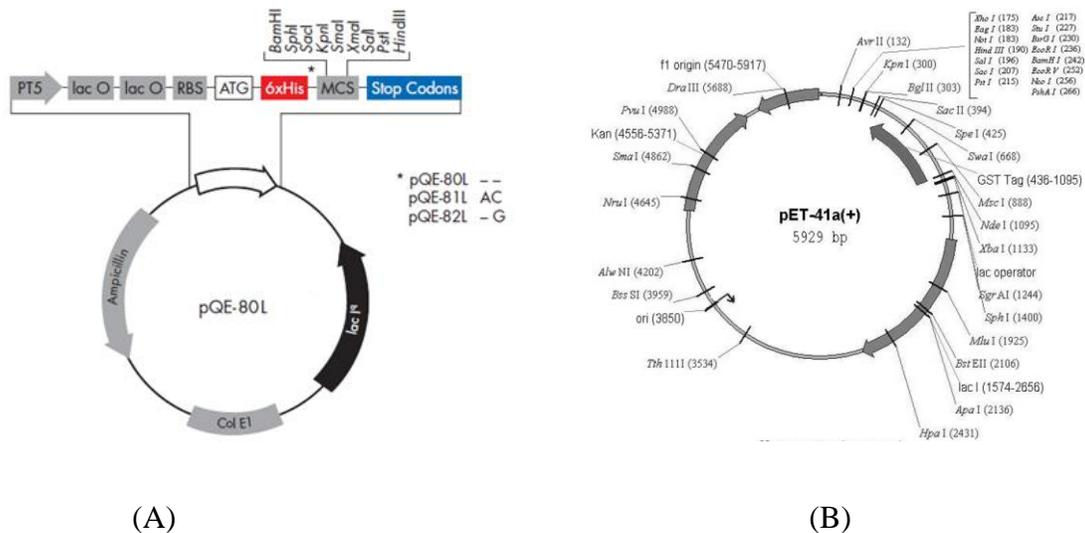


Figure 4.2: Expression vectors used for this study.

pQE-80L (A) includes an N-terminal 6xHistidine encoding region, SphI and KpnI restriction sites. pET-41a vector (B) includes an N-terminal GST tag, 6xHistidine tag and C-terminal 6xHistidine encoding region, KpnI and SacI restriction sites. (Sources: pQE-80L—QIAexpressionist Handbook, Qiagen and pEt-41a - <https://www.addgene.org/vector-database/2592/>).

4.5 Gene expression

4.5.1 Expression of *nanH1*, *nanH3* sialidase domain and Truncated *nanH3*

All the vector-insert constructs generated in this study are described in Table 4.3 and recombinant protein constructs are illustrated in Figure 4.4. Cultures from -80 °C freezer stocks were used to inoculate 5 ml of LB broth containing 100 µg/ml ampicillin or 50 µg/ml kanamycin and incubated overnight, shaking at 225 rpm. Overnight culture (1 ml) was then used to inoculate 25 ml of prewarmed LB broth containing 100 µg/ml ampicillin or 50 µg/ml kanamycin, which was then incubated at 37 °C in a shaking incubator (shaking at 225 rpm) until an optical density measured at 600 nm (OD_{600}) of 0.5-0.7 was reached. OD_{600} was measured using a Nanodrop Spectrophotometer 2000. A 5 ml aliquot was transferred to a snapped-cap tube and incubated in the shaking incubator (uninduced control). Protein expression was induced with the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM to the remaining culture and incubated for 4 hours with shaking at 37 °C. Induced and uninduced cultures were centrifuged at $13,000 \times g$ for 5 minutes and the resulting cell pellet was re-suspended in 200 µl sample buffer (62.5 mM 1.0 M Tris (pH 6.8), 0.465% (w/v) of 10% SDS, 20% (v/v) glycerol, 10% (v/v) 2-mercaptoethanol and 0.0025% (v/v) of bromophenol blue). Samples were boiled for 2 minutes and then sonicated for 3 seconds on ice. The resulting viscous lysate was clarified by centrifugation for 5 minutes at $13,000 \times g$.

Induced and uninduced cell lysates were screened by SDS-PAGE using miniVE Vertical Electrophoresis System (GE Healthcare Life Sciences, Pittsburg, PA, USA). Samples were loaded onto an 12% (w/v) polyacrylamide gel (0.39 M Tris pH 8.5, 0.1% SDS, 0.1% ammonium persulfate, 0.01% Tetramethylethylenediamine) to visualize the protein profile of the crude lysate by staining with Coomassie blue stain to determine if there was protein expressed from the cloned gene. The SDS-PAGE was run at 150 V for approximately 2 hours, or until the SDS sample buffer ran off the bottom of the gel. The gels were placed in Coomassie stain (0.1 % Coomassie Blue, 10% glacial acetic acid) for 1 hour on a rocking platform and then placed in Coomassie de-stain solution (30% methanol, 10% glacial acetic acid) and left on the rocking platform overnight.

Table 4.3: All constructs generated in this study

Vector-insert construct	Expression strain of <i>E. coli</i>	Affinity Tags	Size of the recombinant protein (kDa)	Solubility	Purification method/s
pQE-80L+ <i>nanH1</i>	Top 10	N-terminal 6x Histidine	101	Soluble	Ni-affinity column
pQE-80L+ <i>nanH3</i> Sialidase domain (pQE-80L+SiaN3)	Top 10	N-terminal 6x Histidine	41.5	Soluble	Ni-affinity column
pET-41a + <i>nanH3</i> Sialidase domain (pET-41a +SiaN3)	BL21(DE3)	N-terminal GST 6x Histidine, S-tag C-terminal 6x Histidine	68	Insoluble under the given conditions	Glutathione beads Ni-affinity column
pET-41a+ Truncated <i>nanH3</i> (pET-41a +TN3)	BL21(DE3)	N-terminal GST, 6x Histidine, S-tag C-terminal 6x Histidine	114	Insoluble under the given conditions	-

4.5.2 Determination of optimal conditions for expression of sialidase domain encoding region of *nanH3*

Optimal conditions were tested for expression of sialidase domain encoding region of *nanH3* (SiaN3) and Truncated *nanH3* (TN3) by changing incubation time, temperature and IPTG concentration. Protein expression was carried out as described above at 0.1 mM, 0.5 mM or 1 mM IPTG concentrations, at 37 °C for 4 hour or at 27 °C overnight.

4.5.3 Large-scale expression of *nanH1* and sialidase domain encoding region of *nanH3*

Five millilitres of LB + ampicillin broth was inoculated with Top 10 *E. coli* containing pQE-80L with *nanH1* (clone prepared by Mo Patterson) and grown overnight at 37 °C with shaking at 225 rpm. A 2 ml aliquot from the overnight culture was used to inoculate 2 l of LB + ampicillin broth. When the OD at 600 nm reached 0.6, protein expression was induced with 1 mM IPTG followed by incubation at same conditions for 4 hours. Cells were harvested by centrifugation at 4000 × g for 30 minutes.

A 500 ml LB + kanamycin culture of *E. coli* BL21 (DE3) containing pET-41a+SiaN3 was induced with 1 mM IPTG at OD₆₀₀ 0.5-0.7, followed by incubation at 37 °C with shaking (225 rpm) for 4 hours. Cells were harvested by centrifuging at 5000 × g for 20 minutes.

4.5.4 Solubility of expressed NanH3 sialidase domain in Tris-HCl buffer (pH 7.4)

Another cell culture (20 ml) induced with the addition of IPTG to a final concentration of 1 mM was used to test the solubility of GST fused protein expressed by pET-41a+SiaN3 construct. The cell pellet harvested by centrifugation was then resuspended in Tris-HCl buffer (25 mM Tris-HCl, 0.5 M NaCl, 5% glycerol, pH 7.4). After addition of lysozyme (1 mg/ml), the cell suspension was sonicated for 2 minutes with 50% duty cycle and centrifuged at 10 000 ×g. Then the pellet and the supernatant were analysed by SDS-PAGE.

4.6 Purification of proteins

4.6.1 Purification of soluble His-tagged recombinant protein

The cell pellet of a 500 ml induced culture of Top 10 *E. coli* containing pQE-80L with *nanH1* (pQE-80L+*nanH1*) was resuspended in 30 ml of lysis buffer (25 mM Tris-HCl, 0.5 M NaCl, 5% glycerol, pH 7.4) and sonicated for 5 × 2 minute pulses with 1 minute rest on ice in between. Cell debris was pelleted by centrifuging at 15000 × g for 20 minutes and the supernatant (cleared lysate) was added to a Gravitrap Nickel column (GE Healthcare UK Limited, Buckinghamshire, England) that was pre-equilibrated with 15 column volumes (CV) lysis buffer. The flow-through was reserved and the His-tagged protein bound column was washed twice with wash buffer (25 mM

Tris-HCl, 0.5 M NaCl, 50 mM imidazole, 5% glycerol, pH 7.4). The protein was eluted using elution buffer (25 mM Tris-HCl, 0.5 M NaCl, 500 mM imidazole, 5% glycerol pH 7.4) and eluate fractions of 1.5 ml were collected into 1.5 ml microcentrifuge tubes. Fractions were examined by SDS-PAGE, including a pre-stained protein ladder (Fisher Bioreagents, Fisher Scientific, Ottawa, ON). After pooling the fractions containing purified protein, the protein concentration was determined by Nanodrop spectrophotometer. The eluate was dialysed and concentrated using a protein concentrator (Amicon Ultra-15 Centrifugal Filter unit 30 kDa, Burlington, Massachusetts, U.S.A). Concentrated protein solutions were transferred into micro-centrifuge tubes in 50 μ l aliquots, and flash frozen in liquid nitrogen before being transferred to -80 °C freezer for storage.

4.6.2 Purification of insoluble protein

Harvested cells (*E. coli* BL21 containing pET41a +SiaN3) were resuspended in 20 ml of TNE (Tris-EDTA, 150 mM NaCl), 1% TX-100 (Triton X – 100)). Lysozyme was added (final concentration 1 mg/ml) to the cell suspension and kept on ice for 15 mins. The insoluble portion of the lysed cell suspension was pelleted by centrifuging at 4000 \times g for 40 minutes. The pellet was then resuspended in a solution that contained Complete Mini protease inhibitor cocktail (Roche Molecular biochemicals, Manheim, Germany) and 1% TX-100 in 0.1M PBS. Then phenylmethylsulfonyl fluoride (PMSF) was added to a final concentration of 100 μ M, sonicated for 120 seconds and centrifuged. The pellet was resuspended in Tris-EDTA buffer (10 mM Tris (pH 8), 1mM EDTA) and 1.5% N-Lauroylsarcosine solution and centrifuged at 7000 \times g for 10 minutes. The supernatant was saved. Glutathione Sepharose Beads (Glutathion Sepharose 4B, GE Healthcare Bio-Sciences AB, Uppsala, Sweden) were prepared for binding by mixing with 0.1 M PBS (pH 7) and spinning 1 ml of the solution at 500 \times g for 3 minutes. The supernatant was removed, the pellet (glutathione beads) was mixed with the solubilized protein and incubated at 4 °C for an hour. The glutathione beads were washed once with 0.1 M PBS, 1% TX-100 and twice with PBS by centrifuging at 1400 \times g for 4 minutes. To elute the bound protein, the beads were mixed with a freshly prepared reduced glutathione solution (10 mM glutathione, 50 mM Tris pH 8.0), vortexed gently and centrifuged at 1400 \times g for 30 seconds. The elution step was repeated 3 times to elute all the protein bound to the beads. At each step, aliquots were saved and loaded on to a 12% SDS-poly-acrylamide gel along with the final product.

Since the pET-41a+*nanH3* sialidase domain construct had 6x Histidine tags fused to the C-terminus and N-terminus, protein purification using a Ni-NTA column was also attempted. As the protein was insoluble, buffers containing anionic denaturing detergents were used (Schlager et al, 2012). The cell pellet obtained from a 30 ml induced culture was resuspended in lysis buffer (8 mM Na₂HPO₄, 286 mM NaCl, 1.4 mM KH₂PO₄, 2.6 mM KCl, 1% SDS, pH 7.4) with the addition of 1 mM dithiothreitol (DTT) and sonicated for 2 minutes. Cell debris was incubated in ice/water bath for 30 minutes and centrifuged at 4000 × g for 40 minutes. The supernatant (cleared lysate) was extracted and added into an equilibrated Ni-affinity spin column. The flow-through was discarded and the His-tagged protein bound column was washed twice with wash buffer (8 mM Na₂HPO₄, 286 mM NaCl, 1.4mM KH₂PO₄, 2.6 KCl, 0.1% Sarkosyl, pH 7.4). The protein was eluted using elution buffer (8 mM Na₂HPO₄, 286 mM NaCl, 1.4 mM KH₂PO₄, 2.6 KCl, 0.1% Sarkosyl, 500 mM imidazole, pH 7.4). Eluted protein concentrations were quantified by NanoDrop. The products of the purification process were analysed by SDS-PAGE.

4.7 Sialidase assay

Sialidase activity of *Gardnerella* isolates and purified or crude extracts of recombinant proteins was assessed using a qualitative filter paper spot test (Briselden et al., 1992; Pleckaityte et al., 2012), or a quantitative assay. Both assays used the fluorogenic substrate 2'-(4-methylumbelliferyl)- α -D-N-acetylneuraminic acid sodium salt hydrate (MUNANA) (Sigma-Aldrich Canada, Oakville, ON). Substrate was dissolved in water to a concentration of 300 μ M (0.015% w/v) and aliquots were stored at -20 °C. Prior to each assay, aliquots of substrate were thawed and diluted with 1.0 M sodium acetate (pH 5.8) to a final concentration of 243 μ M.

For the filter spot test, 10 μ l of substrate was added to Whatman qualitative filter paper Grade 1 circles (made using a regular hole punch) placed in a Petri dish, and 10 μ l bacterial culture or protein extract was added to each circle and incubated for 30 minutes in the dark at 37 °C. Sialidase activity was detected by visualizing and photographing filter paper circles under UV light.

For the quantitative sialidase assay, 90 μ l of substrate was combined with 10 μ l of culture or protein extract in duplicate wells of opaque 96-well plates, prior to measuring relative fluorescence units (RFU) over time in a Varioskan LUX Multimode microplate reader (Thermo Scientific). Readings of fluorescence (excitation at 350 nm, emission at 450 nm) were taken at 30 °C over a 60-minute

period at 2-minute intervals with pulsed shaking between readings. The increase in RFU (final fluorescence - initial fluorescence) over time was reported.

4.8 Amplification of homopolymer region of *nanH3*

To determine the length of the homopolymer region in *nanH3* of *Gardnerella piovii*, four strains (W11, VN014, VN015, NR032) were cultured on Columbia Agar plates with 5% Sheep Blood. Primers JH0780 and JH0781 (Table 4.2) were designed to PCR amplify a region including base pairs 2-310 of the open reading frame.

The DNA sequence (308 bp) was amplified by PCR in a Mastercycler Pro 6321 (Eppendorf AG, Hamburg, Germany). The components of the PCR reaction mix (50 µl per reaction) were added to achieve final concentrations of 1×High Fidelity PCR buffer (60 mM Tris-SO₄ (pH 8.9), 18 mM (NH₄)₂SO₄), 0.2 mM dNTP mix, 2 mM MgSO₄, 1U Platinum High Fidelity (Hi-Fi) proof reading Taq polymerase (Invitrogen, Carlsbad, CA, USA). Two colonies of each *G. piovii* strain were randomly picked and added to separate PCR reactions using sterile toothpicks. Thermocycling conditions included 35 cycles of denaturation at 94 °C for 15 seconds, annealing at 55 °C for 30 seconds, extension at 68 °C for 1 minute, and final extension at 68 °C for 5 minutes. PCR products were visualized on a 1% agarose gel along with GeneRuler 100bp Plus ladder (Thermo Scientific, Lithuania). PCR products were purified using QiaQuick PCR purification kit (Qiagen, Hilden, Germany) and purified PCR products obtained from *Gardnerella* strain W11 were sequenced by Sanger sequencing (Macrogen, Korea) using amplification primers (JH0780, JH0781).

To clarify the exact length of the homopolymer region, amplicons generated from the two colonies of each of the four *G. piovii* strains were A-tailed in 10 µl reactions containing 1× Platinum PCR buffer (20 mM Tris HCl (pH 8.4), 50 mM KCl), 2 mM MgCl₂, 0.5 mM dATP, 5 U Platinum Taq polymerase (Invitrogen, Carlsbad, CA, USA) and less than 500 ng of the purified PCR product. The reaction mixture was incubated at 72 °C for 20 minutes in Mastercycler Pro 6321 (Eppendorf AG, Hamburg, Germany). End-modified PCR products were ligated into pGEM-T vector (Promega, Madison, WI, USA). The vector-insert construct was used transform chemically competent DH5α cells or OneShot Top 10 *E. coli* (Invitrogen, Carlsbad, CA, USA) and plated on LB/AMP/X-gal agar media. Ten white colonies were randomly selected from the transformants of each *Gardnerella* strain and transferred into LB + ampicillin broth. Cultures were grown overnight

at 37 °C. Plasmid DNA was isolated using QiaPrep Spin Miniprep kit (Qiagen, Hilden, Germany). Plasmids were sequenced by Sanger sequencing (Macrogen, Korea) with vector primers T7 and SP6.

4.9 Production of polyclonal antibodies against NanH1

The experiments were designed and performed in accordance with the Canadian Council for Animal Care and all work was approved by the University of Saskatchewan Animal Research Ethics Board (protocol 20090050).

Purified recombinant NanH1 protein was used to immunize two rabbits (rabbit A and rabbit B) in order to raise polyclonal antibodies against the protein. Rabbits were immunized with NanH1 protein mixed with TiterMax gold (TiterMax, Norcross, GA, U.S.A) adjuvant. After initial immunization (0.44 mg), two booster doses (0.22 mg, on day 27 and day 42) were injected. Rabbits were exsanguinated on day-57. All the immunization procedures (Figure 4.3) were carried out in the Animal Care Unit (ACU), WCVU.

Blood collected from the pre-bleed (day-0), test bleed (day-41) and final exsanguination was incubated overnight at 4 °C (to allow coagulation) and then centrifuged at 2000 × g for 10 minutes to separate sera. The specificity and the affinity of polyclonal serum was tested using western blot.

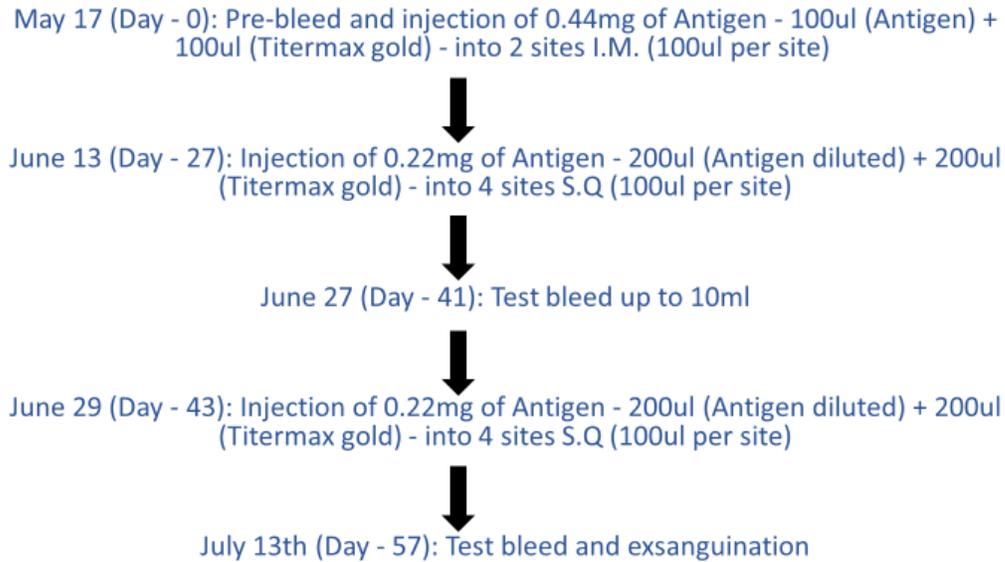


Figure 4.3: Rabbit immunization protocol for NanH1 protein.

4.10 Western blotting

NanH1 (10 µg/ml) protein was mixed with sample buffer (62.5 mM Tris, 4.65% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.25% bromophenol blue) and 10 µl was loaded into two wells of a 12% polyacrylamide gel along with molecular weight marker. Protein was separated by electrophoresis (protocol mentioned previously) using the Amersham miniV system. First, the transfer membrane (Amersham Hybond low Fluorescence 0.2 µM PVDF membrane, GE Healthcare Life Science, Germany) was rinsed in methanol for 20 seconds and in ultrapure water for 20 seconds. The gel, transfer membrane and blotting paper (blotting paper 9 × 10.5 cm, GE Healthcare Bio-Sciences, Piscataway, NJ, U.S.A) were soaked in transfer buffer (25 mM Tris, 192 mM glycine, 20% (v/v%) methanol) for 15 minutes. The blotting sandwich was assembled in the wet-transfer cassette of the Amersham MiniV electro transfer unit by placing sponge pads, blotting paper, the gel and the PVDF membrane as instructed in the user manual. The module was closed and about 300-350 ml of transfer buffer was poured to the cassette module. The water tank was filled with deionized water and ice to avoid overheating of the cassette. Wet transfer was conducted at 400 mA for 90 minutes. After the transfer, the PVDF membrane was soaked in 0.1 M phosphate

buffered saline (PBS) and then blocked with 5% bovine serum albumin (0.1M PBS pH 7.4, 0.1% Tween-20 (PBS-T), 5% BSA). The membrane was separated into two after blocking was completed and rinsed and washed twice with PBS-T. The membranes were incubated in 1:500 (serum: PBS-T) dilution of serum of rabbit A and B separately overnight on a rocking platform. The membranes were rinsed twice and washed with PBS-T for 2× 5 minutes. A dilution of 1:10000 of ECL Plex goat-anti-rabbit Cy5 (Amersham GE Healthcare Biosciences, Pittsburgh, PA, U.S.A) antibodies was used to develop the fluorescent signal. The membranes were rinsed three times and washed with PBS-T 4× 5 minutes. After drying, the membrane was scanned with the Typhoon Trio Variable Mode Imager System (Amersham Biosciences, Sweden) to visualize the results.

To detect the affinity and the specificity of the antibodies towards the native protein, a western blot analysis was carried out using *G. piovii* W11 cell lysate. W11 was cultured in NYC III broth for 48 hours at 37 °C and then the cells were harvested by centrifugation. The cell pellet was then mixed with Tris-HCl (pH 7.4) and treated with lysozyme (1 mg/ml). Cell suspension was sonicated, and the supernatant was clarified by centrifugation. The supernatant was mixed with sample buffer (62.5 mM Tris, 4.65% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.25% bromophenol blue) in 1:1 ratio and loaded on to a 12% polyacrylamide gel along with molecular weight markers. Protein was separated by electrophoresis and the gel was then transferred onto a PVDF membrane following the same protocol mentioned above. The membrane was analysed by Typhoon Trio Variable Mode Imager System (Amersham Biosciences, Sweden).

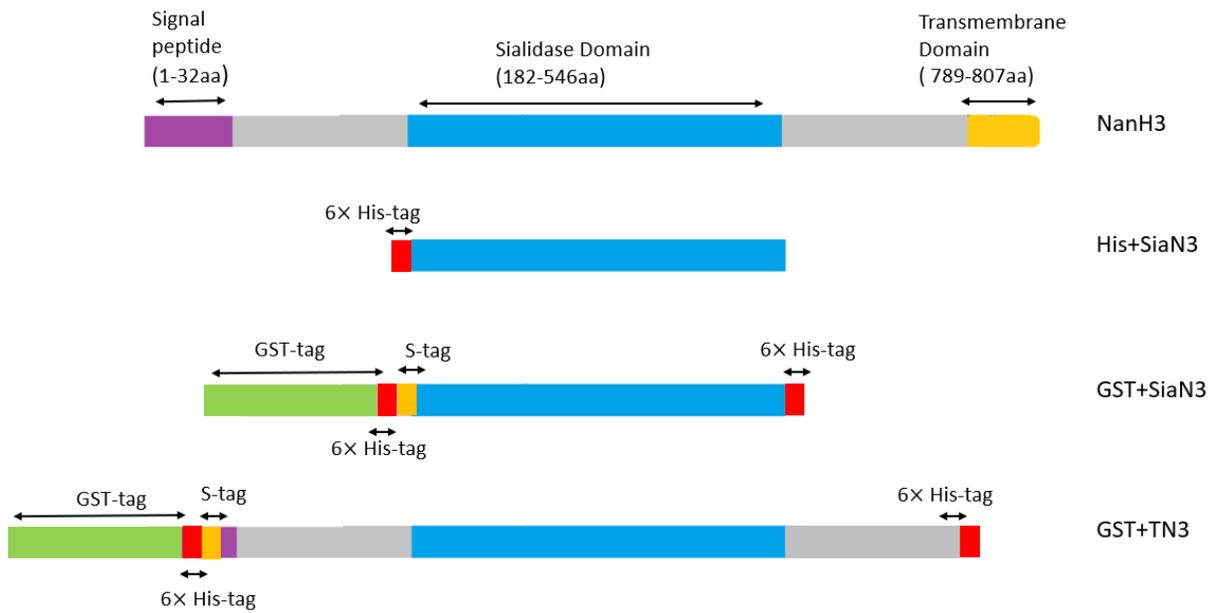


Figure 4.4: Recombinant protein constructs built in this study.

5.0 Results

5.1 Comparison of sialidases

5.1.1 Comparison of NanH3 of *Gardnerella* sp. strain W11 with exo- α -sialidase of *Bifidobacterium bifidum*

Bifidobacterium bifidum belongs to the same family as *Gardnerella* spp. and *B. bifidum* (ATCC 15696) encodes of a well-characterized extra cellular sialidase encoded by siabb2 gene (Genbank accession LC228603). Nishiyama et al. demonstrated that extracellular sialidase in *B. bifidum* enables adhesion of the organism to mucosal surfaces and cleaves sialyl-oligosaccharides and mucin glycans to produce oligosaccharides that support growth of *B. bifidum* (Nishiyama et al., 2017). As this research article was published around the same time we started our investigation, we decided to analyze the sialidase of *B. bifidum* that could potentially be an ortholog of NanH3. We compared the predicted domain structure (Figure 5.1) and the amino acid sequence (Figure 5.2) of the sialidase encoded by *nanH3* in *Gardnerella* spp. with the *B. bifidum* extracellular sialidase homolog. Both proteins are predicted to include an N-terminal signal peptide, a catalytic domain and a C-terminal transmembrane domain. Additionally, the extra cellular sialidase of *B. bifidum* has an alpha-galactosidase domain between the sialidase domain and the transmembrane domain. The two amino acid sequences showed 46% overall sequence identity and 59% similarity in the sialidase domains.

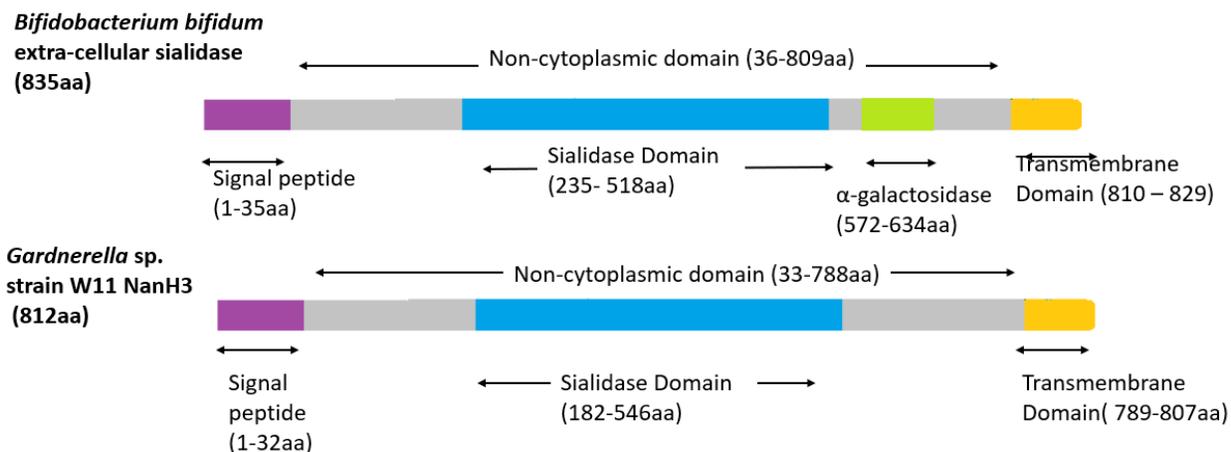


Figure 5.1: Predicted domain structure of sialidases of *Bifidobacterium bifidum* and *Gardnerella* spp..

Amino acid sequence of exo- α -sialidase of *Bifidobacterium bifidum* was recognized to have a sialidase domain, alpha-galactosidase domain with an N-terminal signal peptide and C-terminal transmembrane domain. Amino acid sequence of NanH3 of *Gardnerella* sp. strain W11 was predicted to have a sialidase domain with an N-terminal signal peptide and C-terminal transmembrane domain.

<i>B. bifidum</i>	VQRISKDGGKSWTP	IQYVAQ	QGNAPER	YGYS	SDPSYVVD	KETGE	IFLFFV	HSYKKG	FADSD	Q	60																																													
W11	VQRISKDGGKTFEK	IQIVA	KGKENS	NKYGF	SDPSYVVD	QETGN	IFLFFV	KSYDGG	VWV	DST	60																																													
<i>B. bifidum</i>	LGVDE	SN	NV	LH	VV	SKDNGE	TWSK	PRD	IT	ADITK	GYE	NEWKS	RFATSG	AGIQL	KY	GK	120																																							
W11	ASTNQ	SD	RH	I	L	D	A	V	T	C	SKDNGL	TWSE	PKV	I	T	KDITK	Y	PKNE	-RARFATSG	HGIQL	KY	GK	119																																	
<i>B. bifidum</i>	YKGRLI	QQYA	VGR	TTGS	NA	-----	AVSVYS	DDHGK	TWQ	AGNPV	----	TGML	MDEN	KV	168																																									
W11	YKGRLI	QQYA	IVN	SSNG	SVQR	NNEK	Q	AVSVYS	DDHG	V	TWK	AGNPV	GLGKE	QSH	MDEN	KV	179																																							
<i>B. bifidum</i>	VELSDGR	V	MLNSR	P	---	GAAGY	R	V	A	I	SE	DGG	V	NYG	T	V	K	NET	Q	LPD	P	NN	AH	I	T	R	A	F	P	N	A	225																								
W11	VELSDGR	I	MLNSR	P	HEG	G	ANN	H	R	I	I	A	F	SN	D	G	GET	Y	G	K	A	Y	K	D	E	T	L	P	D	P	G	N	A	Q	I	T	R	A	F	P	D	A	239													
<i>B. bifidum</i>	PEG	S	A	K	A	V	L	L	Y	S	P	R	A	N	E	G	R	A	N	G	V	V	R	I	S	L	D	G	T	T	W	S	S	G	K	L	Y	K	E	G	S	M	A	Y	S	V	I	T	A	L	S	282				
W11	P	I	G	S	D	A	A	K	I	L	L	Y	S	S	S	G	S	G	---	R	A	N	G	L	I	R	V	S	Y	N	D	G	E	N	W	T	E	E	K	G	F	K	N	G	A	M	A	Y	S	T	I	Q	A	L	S	294

Figure 5.2: Sequence alignment of sialidase domains of *Gardnerella* sp. strain W11 NanH3 and *B. bifidum* extracellular sialidase.

Identical amino acid residues are highlighted in blue.

5.1.2 Comparison of putative cell-wall tethered sialidases of *Gardnerella* spp.: NanH3 of strain W11, NanH2 and NanH3 of JCP8151B

Lately, Robinson et al. described three sialidases that are present in *Gardnerella* spp. and two of them were extracellular: NanH2 and NanH3 (Robinson et al., 2019). We compared the amino acid sequence of NanH3 of strain W11 with NanH2 and NanH3 of JCP8151B (Figure 5.3). According to predicted structure, NanH2 consists a signal peptide, translocase domain and a transmembrane domain in addition to the sialidase domain. NanH3 of strain W11 and NanH3 of JCP8151B showed 94% overall sequence identity. The similarity between NanH2 JCP8151B and NanH3 of W11 was 47%.

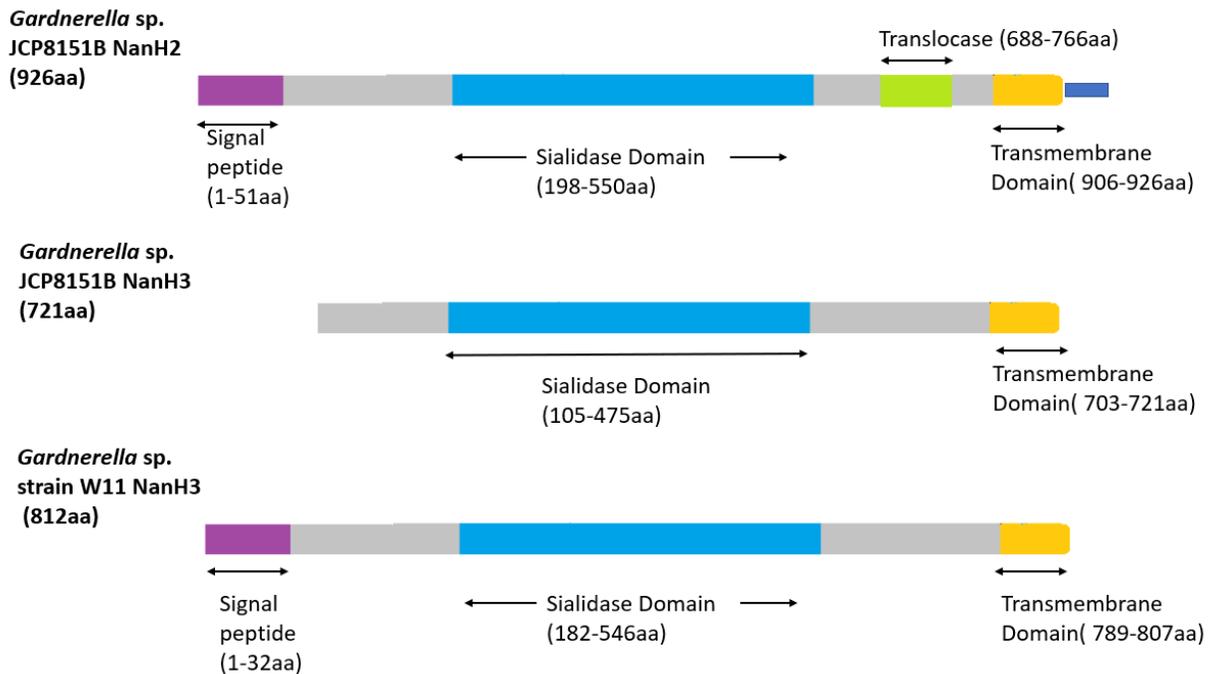


Figure 5. 3 : Predicted domain structures of NanH2, NanH3 of JCP8151B and NanH3 of W11.

5.2 Expression of NanH3 sialidase domain

5.2.1 Expression of NanH3 sialidase domain with an N-terminal 6×His tag (His-SiaN3)

5.2.1.1 Amplification and cloning of *nanH3* sialidase domain into pQE-80L vector

The segment of *nanH3* predicted to encode the sialidase domain of NanH3 of *Gardnerella* isolate W11 was amplified by PCR using proof-reading Taq polymerase (Figure 5.4A) for insertion into the pQE-80L vector as described in the Methods. After ligation and transformation, candidate clones were identified by screening ten colonies with gene-specific PCR (Figure 5.4B). Plasmid was purified from one selected clone and sequenced to confirm that the insert was in-frame with the 6×-His tag of the vector.

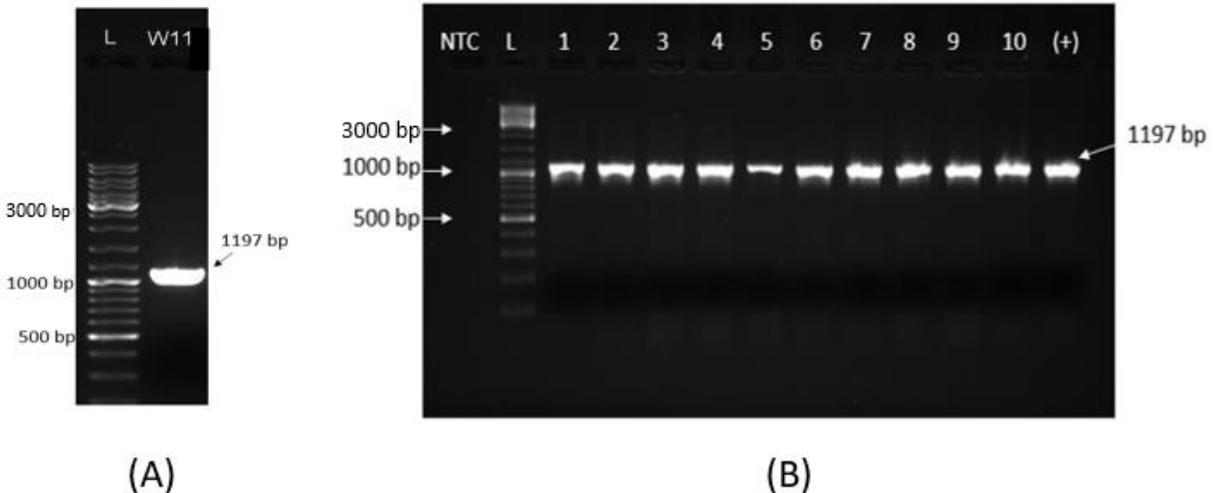


Figure 5.4: Amplification and cloning of the predicted sialidase domain encoding region of *nanH3* with an N-terminal His-tag.

(A) W11 genomic DNA was used as template DNA to amplify the predicted sialidase domain of *nanH3* using primers JH0766 and JH0767. The expected product of 1197 bp is indicated. (B) Ten randomly selected transformed colonies of BL21 *E. coli* (pQE-80L+SiaN3) were screened by PCR using primers JH0766 and JH0767 to detect the presence of the insert. The insert was detected in all ten colonies.

5.2.1.2 Expression of His-tagged *nanH3* sialidase domain

The clone that was confirmed by sequencing was cultured and gene expression was induced with the addition of IPTG (0.5 mM or 1 mM). Figure 5.5A shows the SDS-PAGE separated protein profiles of lysates of induced and uninduced cultures. There was no obvious band in the induced lysate profile at the expected size (43.1 kDa), however, this region is a “crowded” area of a bacterial protein profile and so it was possible that a low level of expression of the sialidase domain was occurring. To investigate this possibility, the His-tagged protein was purified using a Ni-NTA column to detect any expression of the protein. Figure 5.5B shows the presence of the His-tagged sialidase domain (43.1 kDa) in the eluate at low concentration.

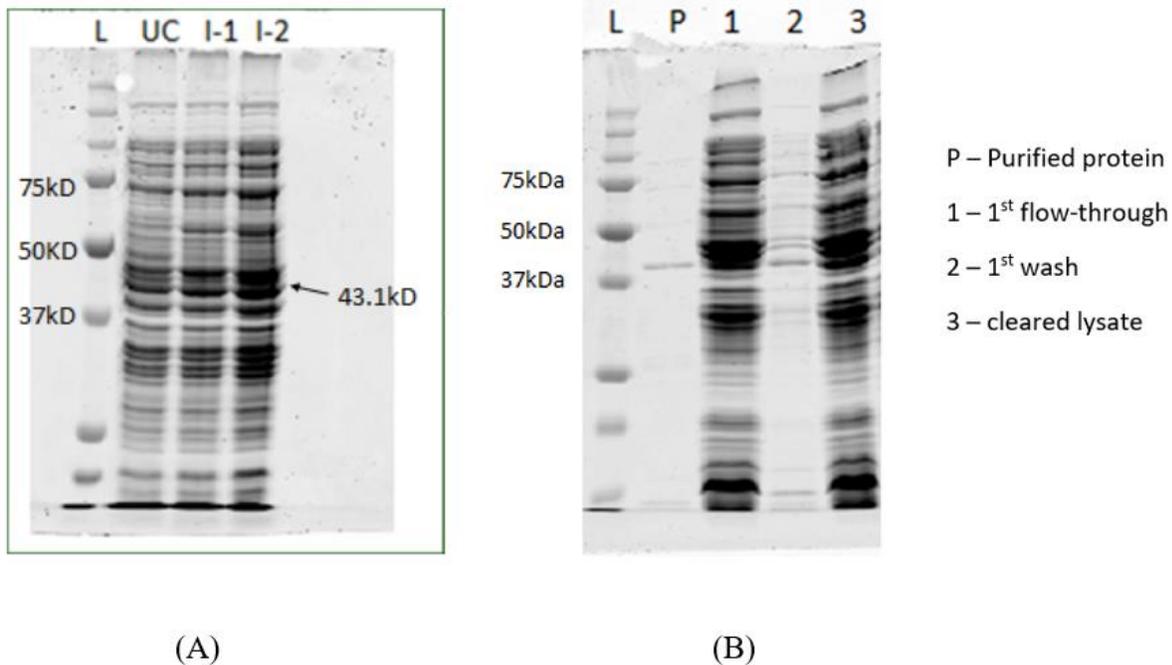


Figure 5.5: SDS PAGE separated protein profiles of *E. coli* (pQE-80L+ SiaN3) cell lysates and protein products extracted by Ni-NTA column purification.

(A) Protein profiles of uninduced (UC) and induced *E. coli* cells cultured with 0.5 mM (I-1) and 1 mM (I-2) IPTG concentrations. (B) Protein profiles of purified His-tagged Sialidase domain of NanH3 protein (43.1 kDa), cleared lysate, 1st wash and 1st flow through of Ni-NTA column purification

5.2.2 Expression of the NanH3 sialidase domain as a GST fusion protein

5.2.2.1 Amplification of *nanH3* sialidase domain and cloning into pET-41a vector

Since the His-tagged NanH3 sialidase domain did not express at an adequate level, the same region of NanH3 was then expressed as a GST fusion protein. The region of *nanH3* encoding the predicted sialidase domain was PCR amplified and ligated into expression vector pET-41a as described in Methods (Figure 5.6A). The insert-vector construct was then used to chemically transform competent *E. coli* (One shot™ Top 10). The transformants were selected in a medium containing kanamycin. Ten colonies were screened by PCR using primers JH0770 and JH0771. All the colonies screened had inserts with the correct size (Figure 5.6B). One selected clone (clone 2) was sequenced to confirm if the insert was in-frame with the N-terminal GST and 6x-His tag and the C-terminal 6x-His tag of the vector.

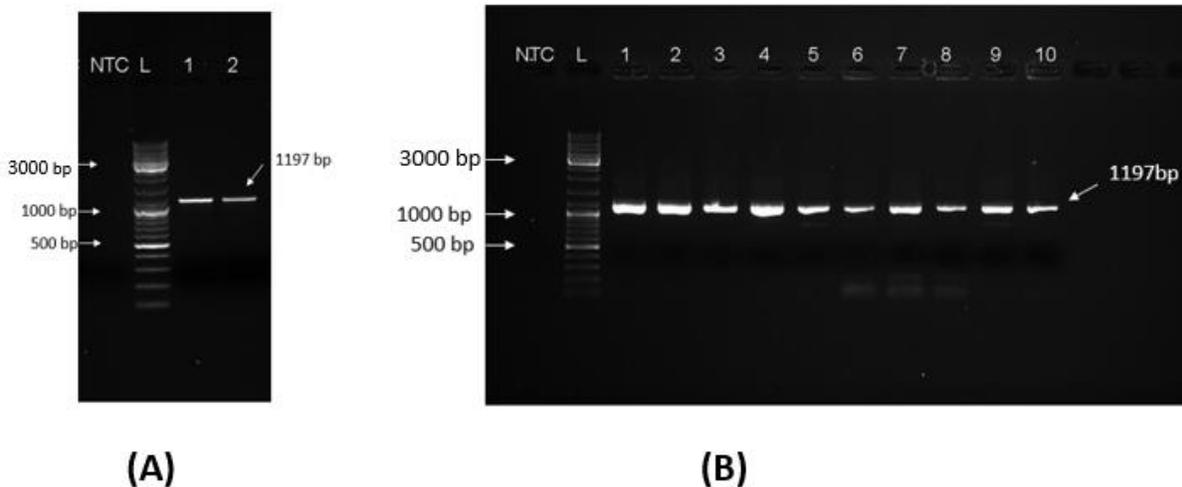


Figure 5.6: Amplification and cloning of the predicted sialidase domain encoding region of *nanH3* with an N-terminal GST tag.

(A) W11 genomic DNA was used as template DNA to amplify sialidase domain of *nanH3* using primers JH0770 and JH0771. (B) Ten randomly selected transformed colonies of BL21 *E. coli* (pQE-80L+*nanH3* sialidase domain) were screened by PCR using primers JH0770 and JH0771 to detect the presence of the insert. The insert (1197 bp) was detected in all ten colonies.

5.2.2.2 Expression of NanH3 sialidase domain using pET41 a

To express the NanH3 sialidase domain (SiaN3) as a GST fusion protein in *E. coli*, the plasmid constructs (pET41a+SiaN3) from clone 2 were extracted and used to transform chemically competent BL21(DE3) *E. coli*. A selected transformant was cultured, and protein expression was induced with the addition of isopropyl β -D-1-thiogalactopyranoside (1mM). Figure 5.7 shows the SDS-PAGE separated protein profiles of induced and uninduced cell lysates. Expression of the GST tagged NanH3 protein (GST+SiaN3) was observed (predicted mass based on the sequence was 68kDa).

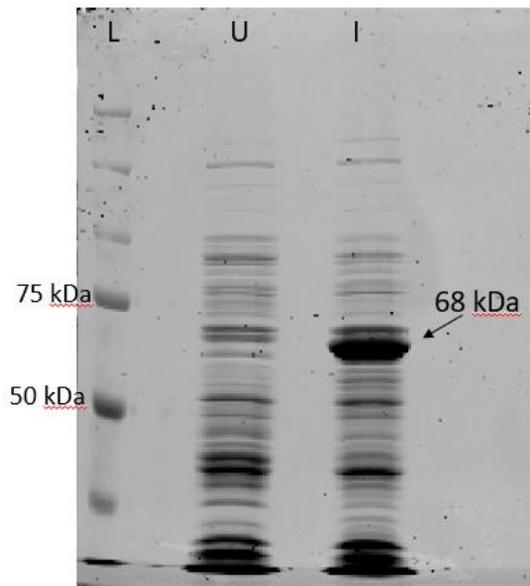


Figure 5.7: SDS PAGE separated protein profiles of BL21 *E. coli* with GST tagged NanH3 sialidase domain.

Cell lysates extracted from uninduced (U) and induced (I) *E. coli* (BL21) cultures with 1mM IPTG concentration. The predicted size of the fusion protein (GST+SiaN3) is 68 kDa.

5.2.2.3 Detection of optimal conditions for gene expression

5.2.2.3.1 Detection of optimal IPTG concentration for induction of gene expression

To determine the optimal isopropyl β -D-1-thiogalactopyranoside (IPTG) concentration for inducing expression of the sialidase domain GST fusion., *E. coli* BL21 (DE3) cells containing the *nanH3* sialidase domain fused to pET41a vector were cultured in each of five snap-cap tubes containing 5 ml of LB/KAN. The cultures were induced with 0 mM, 0.1 mM, 0.25 mM, 0.5 mM and 1 mM IPTG concentrations respectively. After induction the cultures were incubated for 4 hours at 37 °C and the cells were harvested by centrifugation. Figure 5.8A shows the SDS-PAGE separated protein profiles of induced and uninduced cell lysates. All concentrations of IPTG resulted a prominent band at 68 kDa. A concentration of 0.1 mM of IPTG was sufficient for induction of protein expression.

5.2.2.3.2 Detection of solubility of GST fused SiaN3

A portion of the cell culture which induced with 1 mM of IPTG was used to test the solubility of the GST fused protein. The cell pellet harvested by centrifugation was then resuspended in a Tris-HCl buffer (25mM Tris-HCl, 0.5 M NaCl, 5% glycerol, pH 7.4). After lysing the cells by adding lysozyme (1 mg/ml), the cell suspension was sonicated and centrifuged. Then the pellet and the supernatant were analysed by SDS-PAGE (Figure 5.8 B). The protein of interest was associated with pellet and was not soluble in the given conditions.

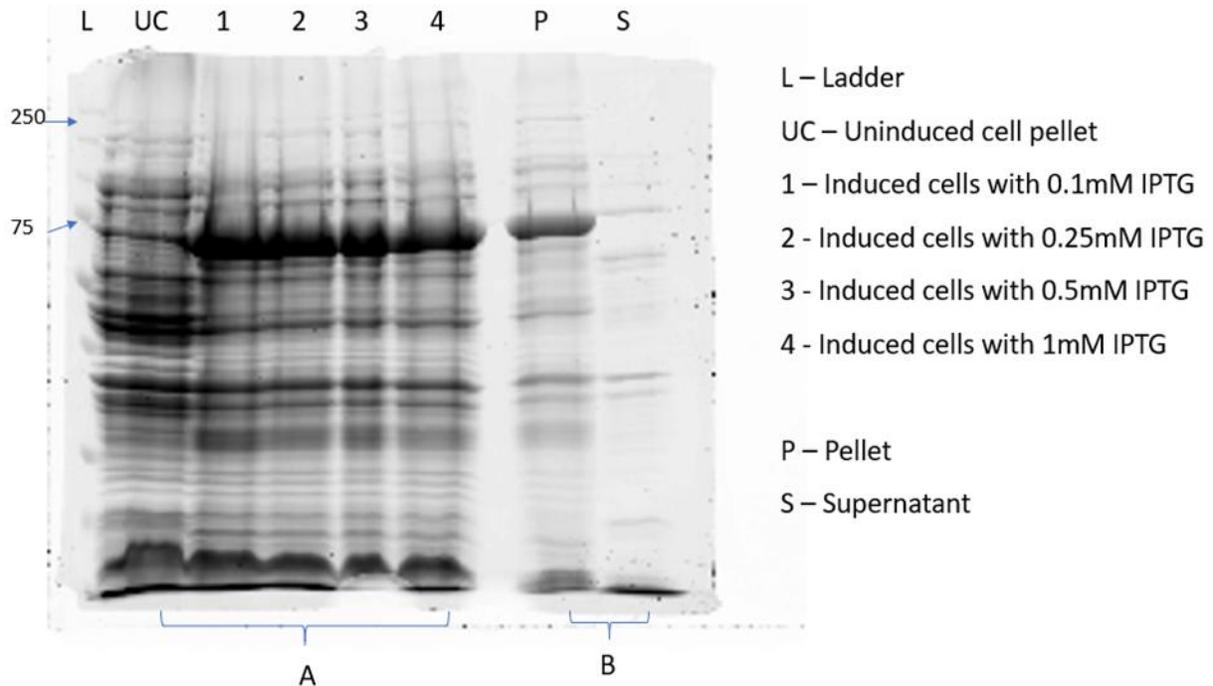


Figure 5.8: Protein profiles of cell lysates of *E. coli* BL21 (pET-41a + SiaN3).

(A) Protein profile of cell lysates of uninduced (UC) and induced (lanes 1- 4) *E. coli* BL21 culture with different IPTG concentration. The predicted size of the GST fusion protein of *nanH3* sialidase domain is 68 kDa. (B) Protein profiles of *E. coli* BL21 (pET-41a + SiaN3) cell pellet and the supernatant indicated that GST fused SiaN3 was largely insoluble.

5.2.2.3.3 Low temperature protein expression of GST fused SiaN3

A possible approach to solubilize and properly fold insoluble or partially soluble recombinant protein is to induce protein expression at a lower temperature than 37°C. Two 30 ml BL21 *E. coli* cultures carrying pET-41a + *nanH3* sialidase domain insert were given a cold shock by incubation on ice for 30 minutes before induction. After inducing with 0.1 mM IPTG, two cultures were incubated at 10 °C and 25 °C separately overnight in shaking incubators. Cells were harvested and the protein profiles (Figure 5.9) were screened by SDS-PAGE. The temperatures and the condition we used did not increase solubility of the protein.

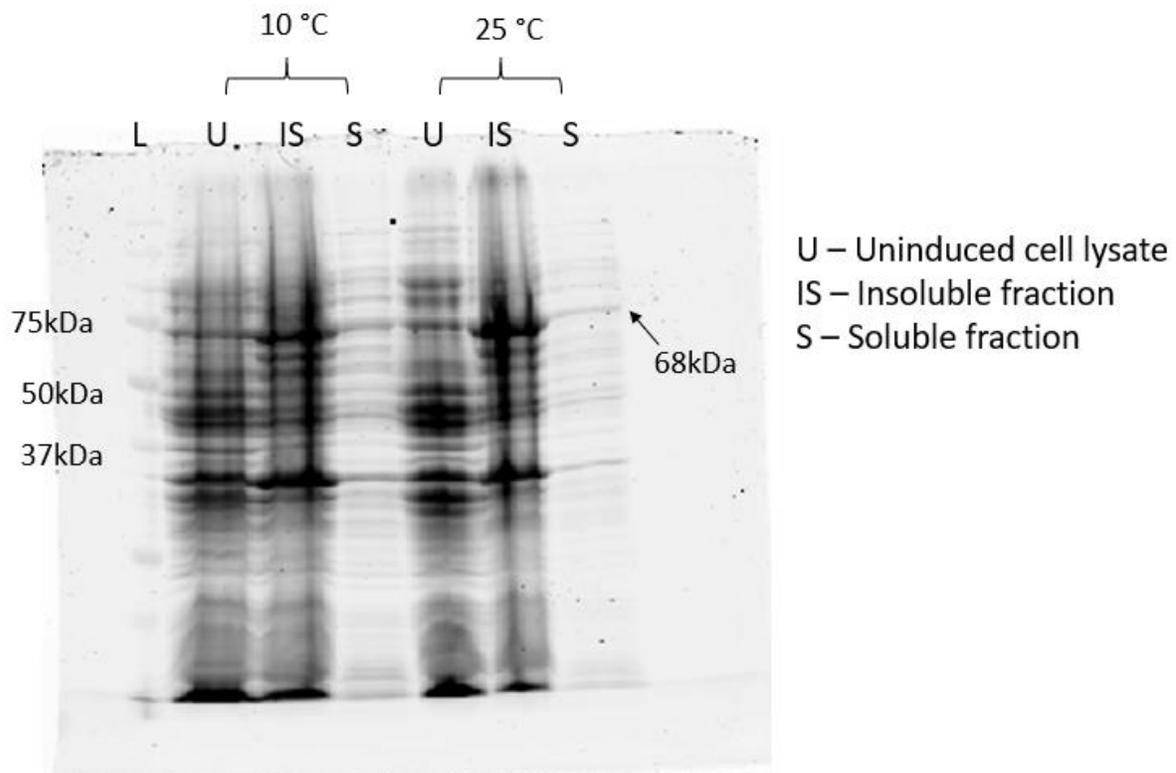


Figure 5.9: Protein profiles of *E. coli* BL21 (DE3) cultures (pET-41a + *nanH3* sialidase domain) induced at 10 °C and 25 °C.

GST fused SiaN3 (68 kDa) was associated with the pellet at both temperatures.

5.2.2.4 Purification of insoluble GST fused sialidase domain of NanH3

5.2.2.4.1 Purification of GST fused sialidase domain of NanH3 (GST+SiaN3) using glutathione beads

BL21 (*nanH3* sialidase domain + pET-41a) cells harvested from a 500 ml culture induced with 0.1 mM IPTG was used to purify the recombinant protein (GST+SiaN3) using glutathione beads. Figure 5.10 shows the SDS-PAGE separated protein profiles of the purification products. Addition of N- Lauroylsarcosine resulted in improved solubility the protein, but the GST-fused protein did not bind to glutathione beads. This could be due the folding of the protein sterically hindering the GST tag and preventing binding to the beads.

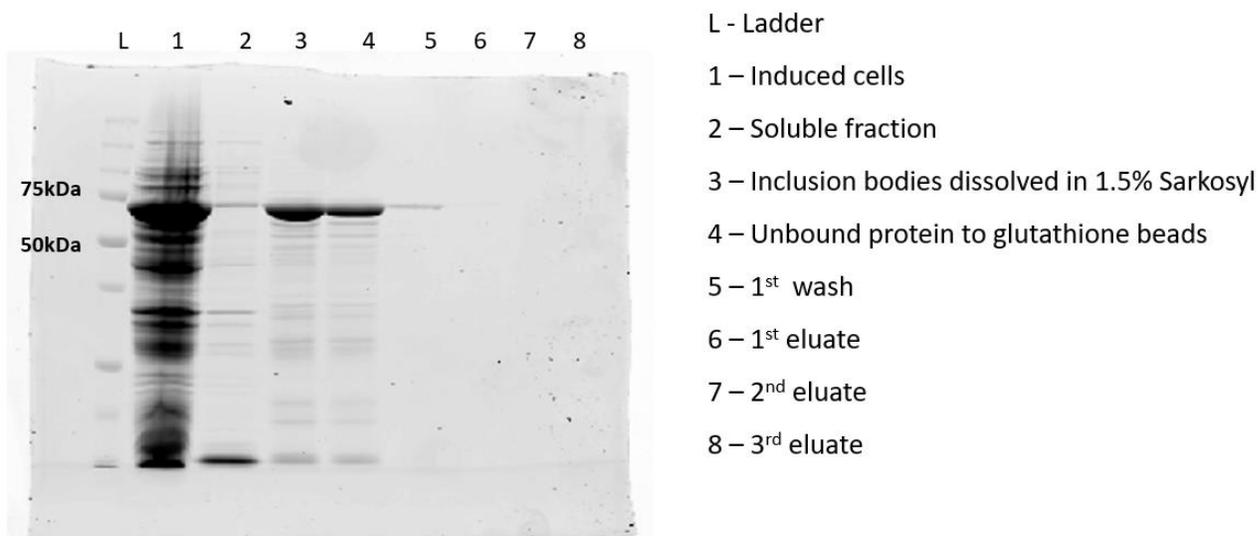


Figure 5.10: Protein profiles of purification products of glutathione bead purification.

5.2.2.4.2 Purification of sialidase domain of NanH3 (SiaN3) using Ni-NTA column

Since the vector- insert construct (pET-41a+Sia2) included a C-terminal and N-terminal 6× Histidine tags, we attempted purifying the protein using a Ni-NTA column. As the protein was insoluble, buffers containing anionic denaturing detergents were used (Schlager et al, 2012) for the purification process. The protein was eluted from the column using an elution buffer containing Sarkosyl (8 mM Na₂HPO₄, 286 mM NaCl, 1.4 mM KH₂PO₄, 2.6 mM KCl, 0.1% Sarkosyl, 500 mM imidazole, pH 7.4). The protein concentrations of the eluate fractions were quantified by NanoDrop (Table 5.1). The products of the purification process were analysed by SDS-PAGE (Figure 5.11). Most of the protein bound to the column was eluted in the second fraction (3.6 mg/ml).

Table 5.1: Concentration of protein eluted by Ni-NTA column purification.

Sample ID	Protein concentration (mg/ml)
1 st eluate	0.27
2 nd eluate	3.60
3 rd eluate	1.87

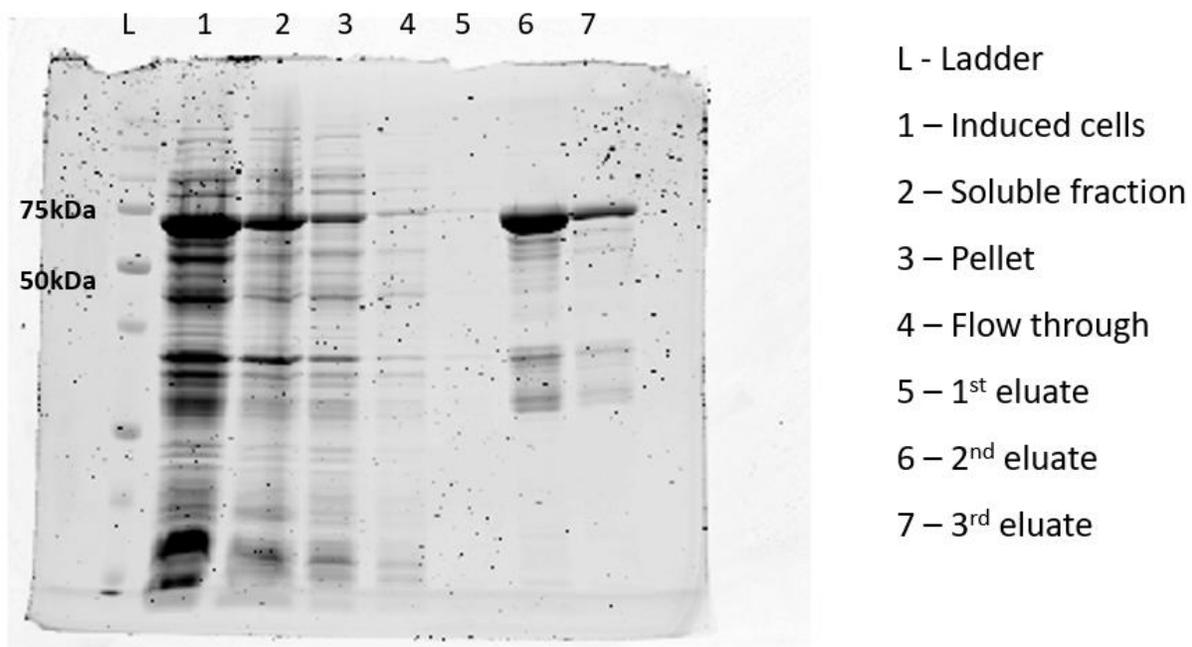


Figure 5. 11: Protein profiles of purification products of Ni-NTA column purification.

5.2.2.4.3 Sialidase activity of GST fused SiaN3

To detect if the Ni-NTA column purified sialidase domain of NanH3 had Sialidase activity, we carried out a filter spot test using 2'-(4-methylumbelliferyl)- α -D-N-acetylneuraminic acid (MUNANA). An image was taken under UV light after incubating the filter paper circles at 37 °C for 30 minutes (Figure 5.12). Recombinant NanH1 protein was used as the positive control. The 2nd eluate did not show sialidase activity likely due to the usage of denaturing detergents in the purification process.

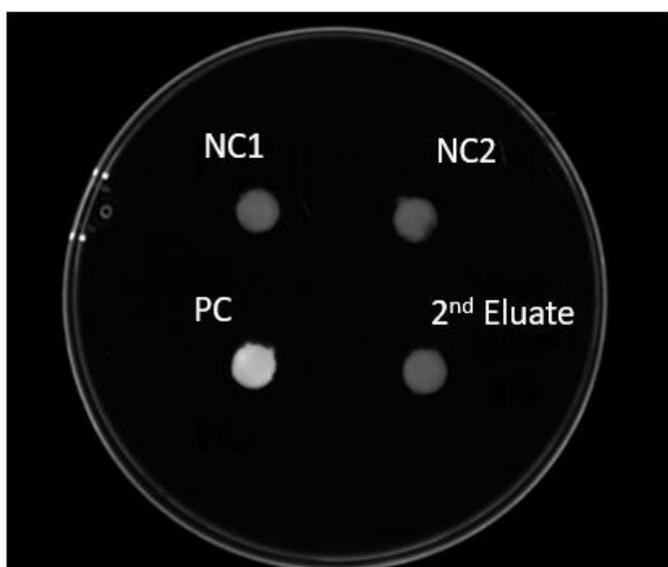


Figure 5. 12: Filter spot test to detect sialidase activity of 2nd Eluate of Ni-NTA column purification. NC1 – Substrate only, NC2 – buffer only, PC – Recombinant NanH1

To remove the detergents, the 2nd eluate (1 ml) was mixed with lysis buffer (1.5 ml) without SDS (8 mM Na₂HPO₄, 286 mM NaCl, 1.4 mM KH₂PO₄, 2.6 KCl) added to a dialysis tube. The tube was then immersed in 1 litre of PBS (pH 7.4) and kept overnight at 4 °C. A portion of the dialysed solution was added to an Amicon concentrator to concentrate the protein. Both these solutions were used for the filter spot test (Figure 5.13). Neither of the dialysed protein solutions showed sialidase activity.

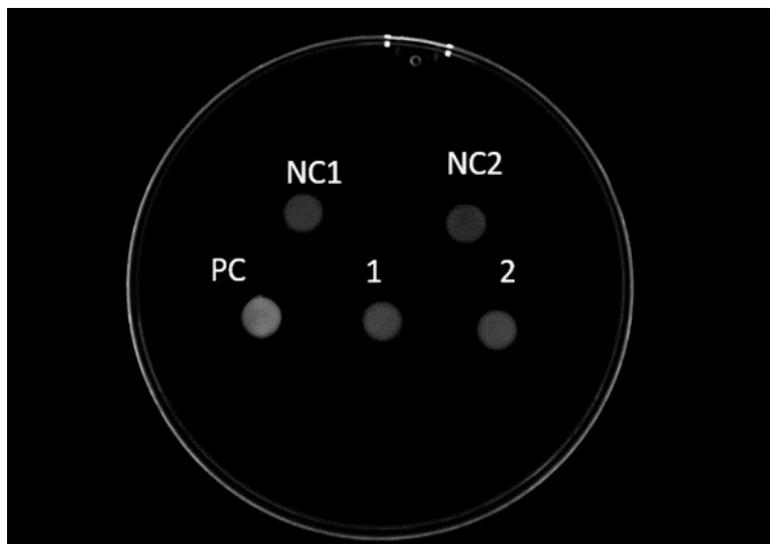


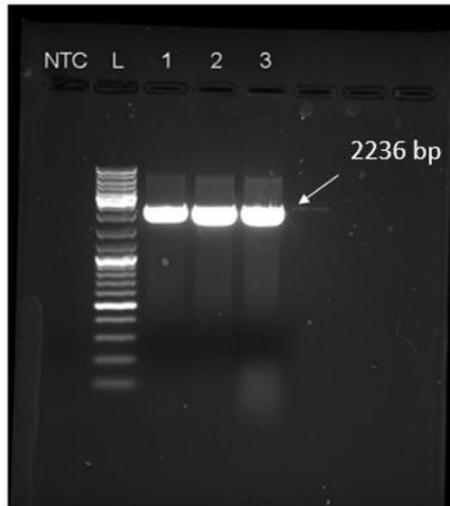
Figure 5.13: Filter spot test to detect sialidase activity of 2nd eluate of Ni-NTA column purification.

NC1 – Substrate only, NC2 – buffer only, PC – NanH1, 1 – Dialyzed eluate, 2- Dialyzed and concentrated eluate.

5. 3 Expression of Truncated *nanH3* with a GST tag

5.3.1 Amplification and cloning of Truncated *nanH3* into pET-41a

Since the recombinant sialidase domain of NanH3 did not show activity we decided to express a truncated version of *nanH3* (121-2340 bp of *nanH3*) in an attempt to produce a properly folded protein. A truncated version of *nanH3* (Truncated *nanH3*/TN3) of *Gardnerella* strain W11 was amplified by PCR excluding signal peptide and the trans-membrane encoding regions (Figure 5.14). TN3 was inserted into pET-41a as described in the Methods. After ligation and transformation, candidate clones were identified by screening ten colonies with gene-specific PCR (Figure 5.15). Plasmids were purified from one selected clone and sequenced to confirm that the insert was in-frame with the GST tag of the vector.



NTC – No template control
 L – Ladder
 Lanes 1, 2, 3, - Genomic DNA of W11

Figure 5.14: PCR amplification of the truncated version of *nanH3* (TN3).

W11 genomic DNA was used as the template DNA to amplify the predicted sialidase domain of *nanH3* using primers JH0799 and JH0800. The expected product of 2236 bp is indicated.

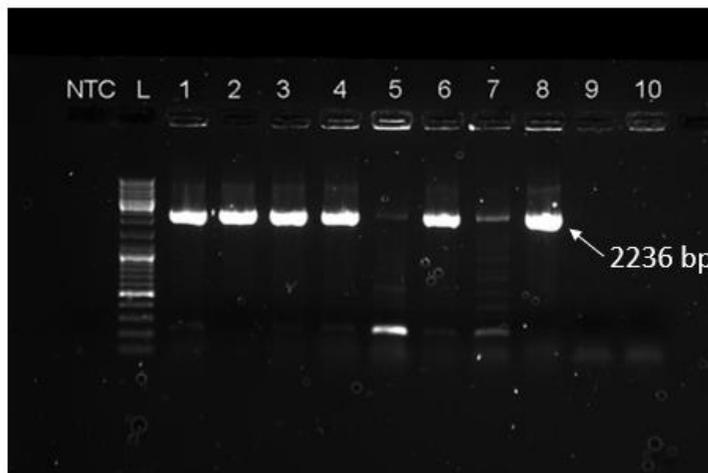


Figure 5.15: PCR screening of transformants after cloning of TN3 with an N-terminal GST tag.

Ten randomly selected transformed colonies of BL21 *E. coli* (pET-41a+TN3) were directly screened by PCR using primers JH0799 and JH0800. The insert was detected in clone 1, 2, 3, 4, 6 and 8.

5.3.2 Expression of GST fused TN3

To express the GST fused TN3 (GST+TN3), the plasmid construct from clone 2 was extracted and used to transform chemically competent *E. coli* BL21 (DE3). A selected clone was cultured, and protein expression was induced with the addition of isopropyl β -D-1-thiogalactopyranoside (0.5 mM). Figure 5.16 shows the SDS -PAGE separated protein profiles of induced and uninduced cell lysates. Expression of GST+TN3 was observed (predicted mass based on the sequence was 114 kDa).

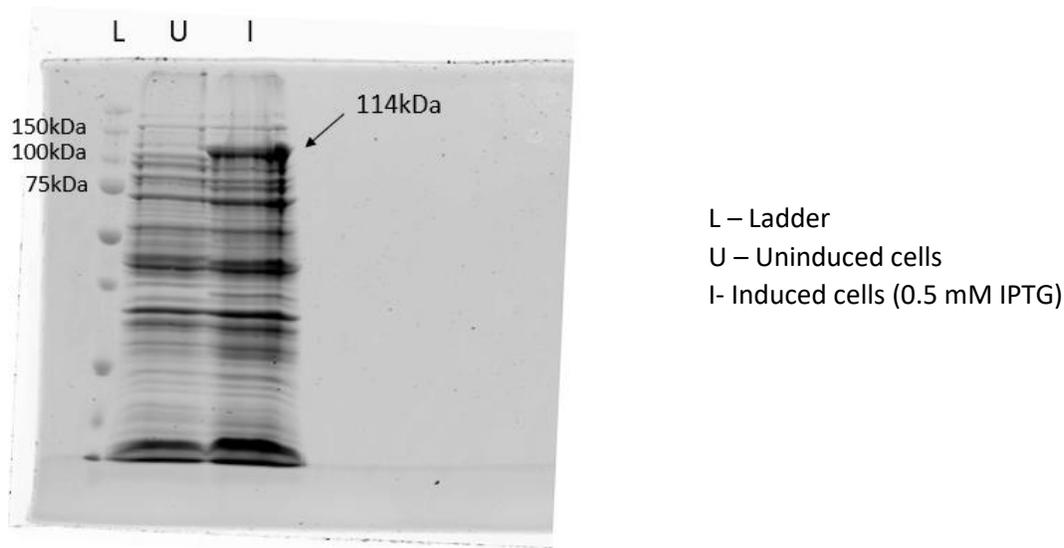


Figure 5.16: SDS PAGE separated protein profiles of *E. coli* (pET-41a+TN3) cell lysates and purified protein products extracted Ni-NTA column purification.

(A) Protein profiles of uninduced (UC) and induced *E. coli* cells cultured with 0.5 mM (I) IPTG concentration.

5.3.3 Determining the solubility of GST+TG2

Another cell culture was induced with the addition of IPTG to a final concentration of 1 mM was used to test the solubility of the GST fused protein. The cell pellet harvested by centrifugation was then resuspended in a Tris-HCl buffer (25 mM Tris-HCl, 0.5 M NaCl, 5% glycerol, pH 7.4). After lysing the cells by adding lysozyme (1 mg/ml), the cell suspension was sonicated and centrifuged. Then the pellet and the supernatant were analysed by SDS-PAGE (Figure 5.17). The protein of interest was associated with pellet and was not soluble in the given conditions.

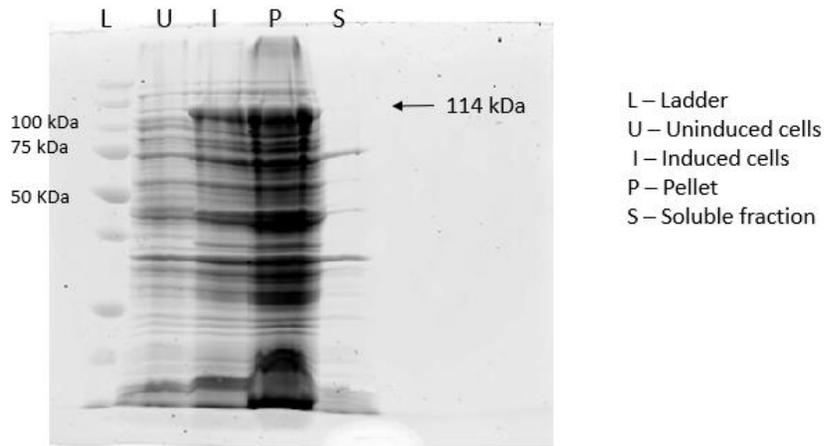


Figure 5.17: Solubility of GST+TN3. Protein profiles of *E. coli* BL21 (pET-41a + TN3) cell pellet and the supernatant indicated that GST fused TN3 was largely insoluble.

5.3.4 Low temperature expression of GST tagged TN3

A possible approach to solubilize and properly fold insoluble or partially soluble recombinant protein is to induce protein expression at a lower temperature than 37°C. A 30 ml BL21 *E. coli* culture carrying pET-41a+TN3 insert was given a cold shock by incubation on ice for 30 minutes before induction. After inducing with 0.1 mM IPTG, the culture was incubated at 20 °C overnight in a shaking incubator. Cells were harvested and the protein profiles (Figure 5.18) were screened by SDS-PAGE. The low temperature and the condition we used did not increase solubility of the protein.

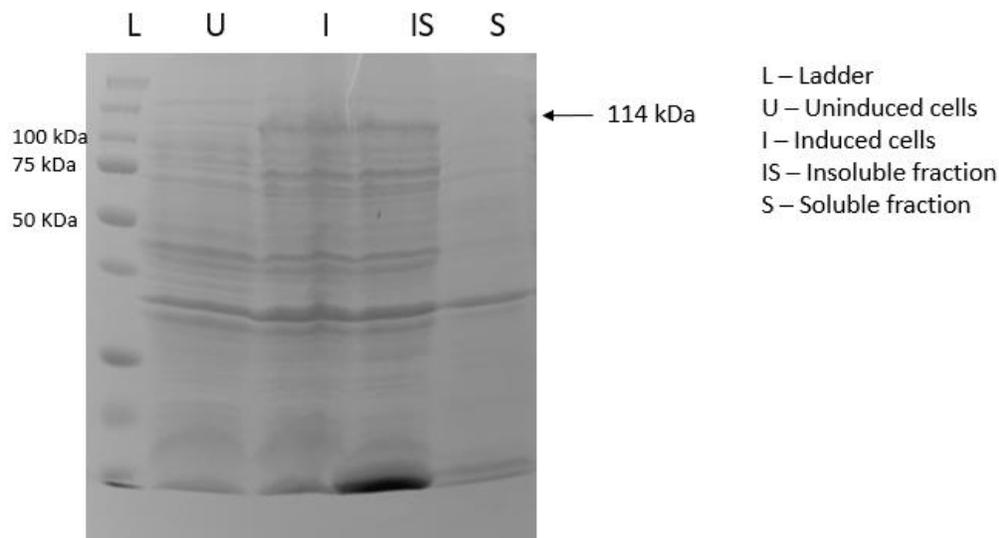


Figure 5.18: Protein profiles of *E. coli* BL21 (DE3) culture (pET-41a+TN3) induced at 20 °C. GST fused TN3 (114 kDa) was associated with the pellet.

5.4 Evidence of slipped strand mispairing in *nanH3*

Previous attempts to express the *nanH3* gene from *G. piovii* isolate W11 using the pQE-80L vector system were not successful, and it was discovered that this was due to the cloned sequences containing premature stop codons, even though the sequences were amplified with proof-reading Taq polymerase (Patterson, 2018). When the sequence of *nanH3* was analyzed, a homopolymer of repeated Cytosines was identified in the open reading frame (73-84 bases from the start site) that could possibly undergo slipped strand mispairing, resulting a frame shift that could lead to an early termination of translation. Here we set out to determine if the homopolymer region of *nanH3* varies in length within and between strains of *Gardnerella piovii* (cpn60 subgroup B).

5.4.1 Amplification and screening of homopolymer region of *nanH3* in *Gardnerella piovii* strain W11

To visualize the homopolymer region in *nanH3* of *G. piovii* strain W11, primers JH0780 and JH0781 were designed to amplify a portion of the gene (2-310 bp) containing the homopolymer. High fidelity, proof-reading Taq polymerase was used for direct PCR from W11 colonies, to minimize opportunities for polymerase slippage. We sequenced a 308 bp amplicon from each of the ten selected colonies of W11 grown on sheep blood agar. Sanger sequencing was performed with both amplification primers. Electropherograms (Figure 5.19) obtained from Sanger sequencing showed at least two overlapping sequences after the poly-C region in the forward strand. This could be due to slippage of polymerase enzyme or presence of two/more homopolymer lengths within the PCR product pool. Similar results were observed for the other seven colonies in the forward and reverse primer sequence electropherograms.

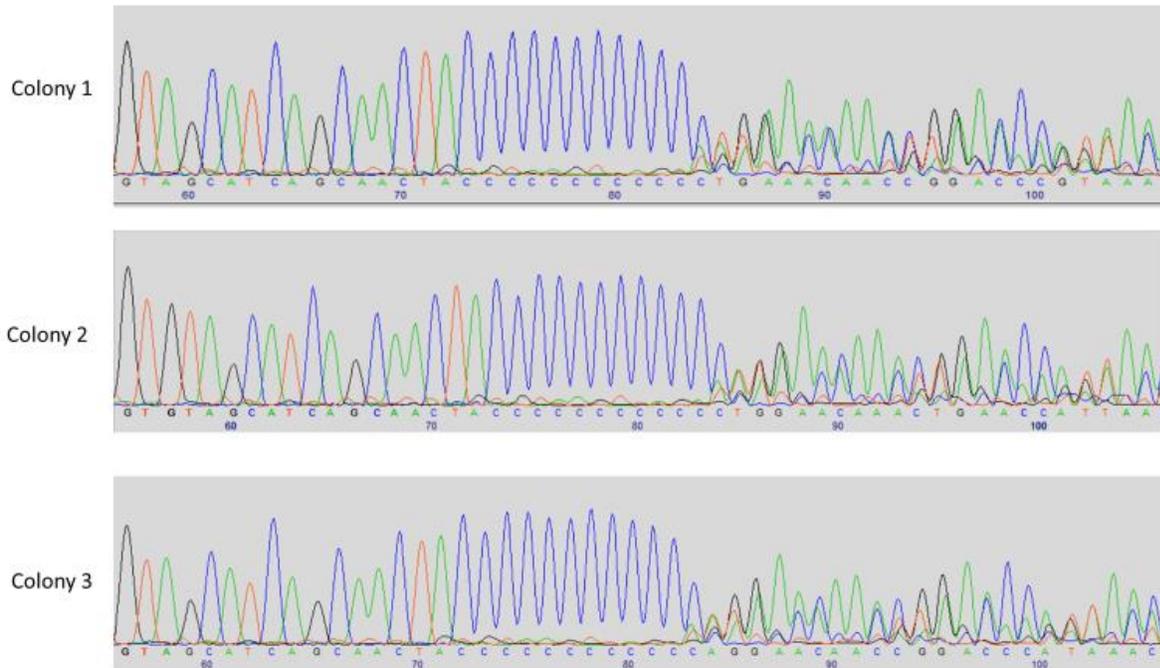


Figure 5.19: Electropherograms showing the poly-C region of *nanH3* in three colonies of W11.

In each electropherogram (sequenced with forward primer), after eleven peaks of cytosine, the sequencing signal shows multiple peaks suggesting the existence of at least two overlapping sequences. This indicates that each colony had individuals with different lengths in the homopolymer.

5.4.2 Cloning PCR products by ligating into pGEM-T vector

To clarify the exact length of the homopolymer region, amplicons generated from two isolated colonies (colony 6, colony 8) were A-tailed and ligated into pGEM-T vector. The vector-insert construct was used to transform chemically competent DH5 α cells and plated in LB/AMP/X-gal media. Ten white colonies were randomly selected from the transformants of colony 6 and 8 and plasmids were extracted. Extracted plasmids were screened for the insert by PCR using JH0780 and JH0781. All 20 colonies showed the correct sized insert (Figure 5.20)

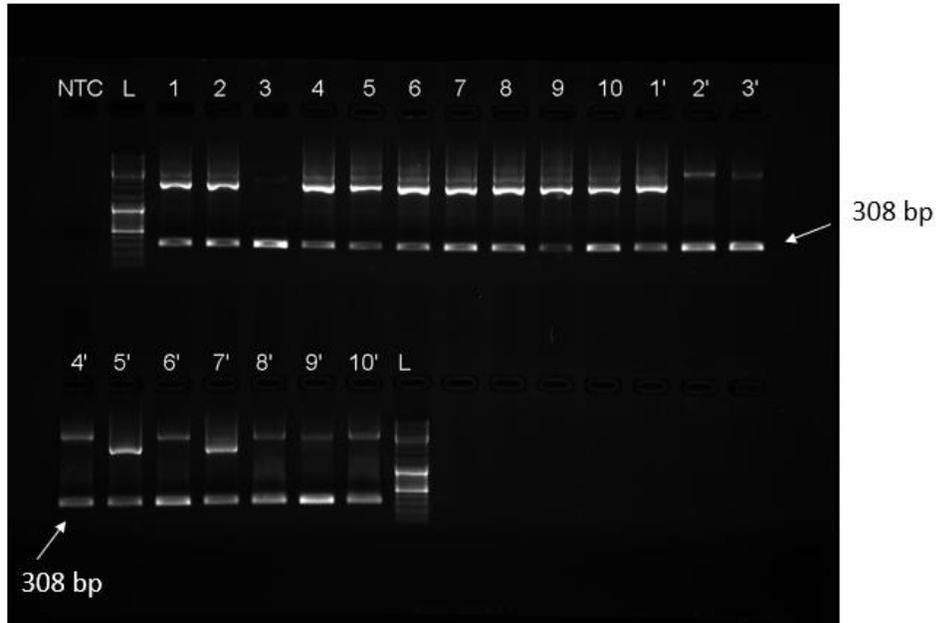


Figure 5.20: PCR amplification of plasmid (pGEM-T+ homopolymer) extractions.

All ten transformants of colony 6 (1-10) and all ten transformants of colony 8 (1'-10') contained the 308 bp insert.

5.4.3 Length of the homopolymer of *nanH3* in *Gardnerella piotii* strain W11

Plasmids (pGEM-T +homopolymer region) were sequenced by Sanger sequencing with T7 and SP6 primers. The sequencing revealed that the cloned PCR products had identical sequences apart from variation in the number of cytosine residues in the homopolymeric tract. Out of the ten colonies cloned using PCR amplicons of W11 colony 6 (Figure 5.21A) and 8 (Figure 5.21B), eight and seven colonies respectively gave clear electropherograms. In the poly-C region, 12 C's was the most common and the only length that makes the coding region in-frame. When there are 11, 13 or 14 C's, a premature stop codon is generated when translated in silico.

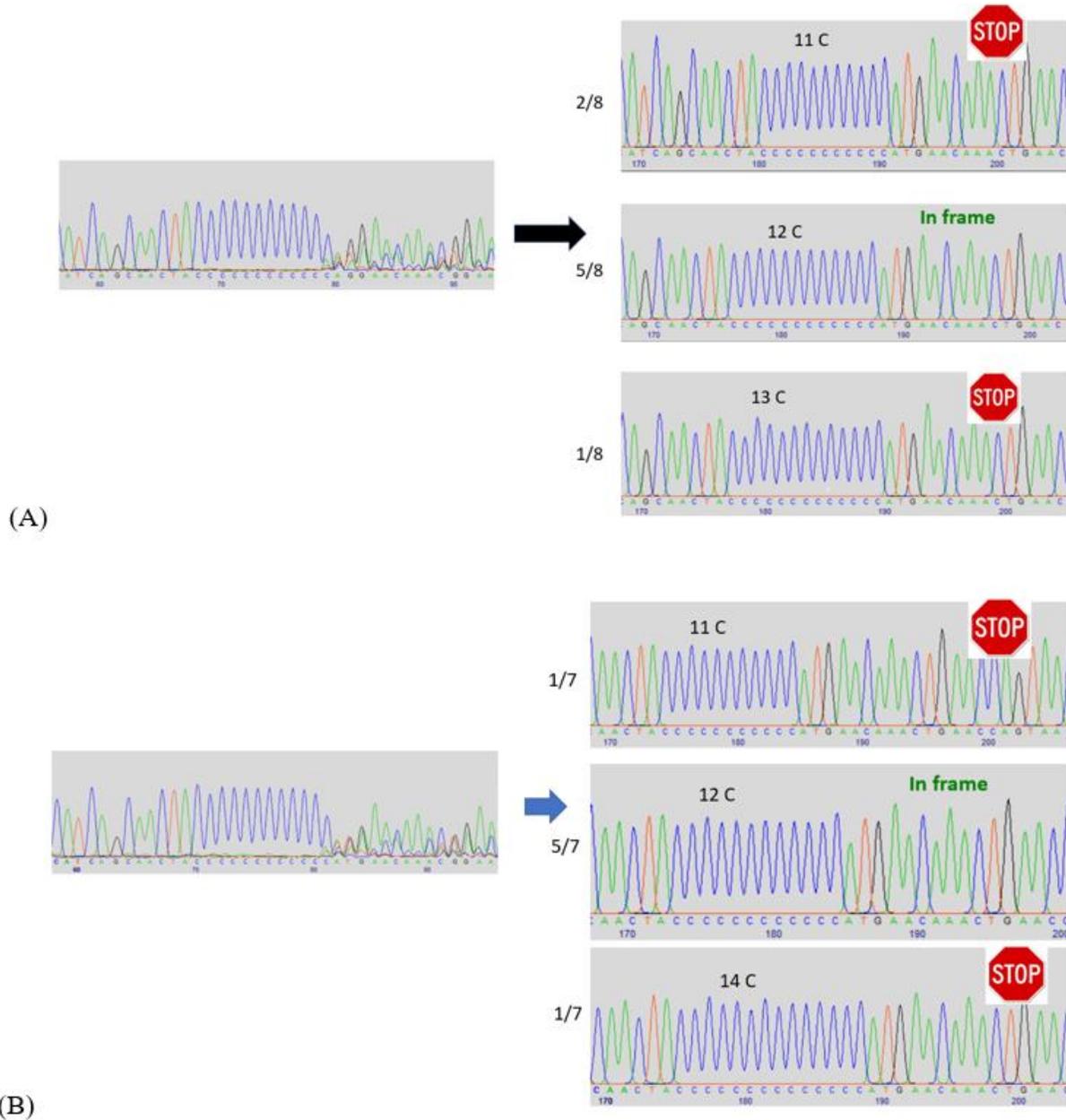


Figure 5.21: Electropherograms of the homopolymer regions in two W11 colonies.

(A) Colony 6 had individuals with 11, 12 and 13 C's in the homopolymeric region. Five out of eight colonies had 12C's in the homopolymer region. (B) Colony 8 had individuals with 11, 12 and 14 C's in the homopolymeric region. Five out of eight colonies had 12 C's in the homopolymer region. Having 12 C's makes the coding region in-frame. When there is 11,13 or 14 C's, a premature STOP codon is generated possibly making the protein truncated.

5.4.4 Length of the homopolymer of *nanH3* in other *Gardnerella piovii* strains

To determine if slipped strand mispairing occur in other isolates of *Gardnerella piovii*, we followed the same study design (Figure 5.22) for biological replicates of three more isolates. We sequenced 20 plasmids (pGEM-T+ homopolymer) per isolate. Out of 20 we could recover 14, 18 and 16 clear and informative sequences from VN014, VN015 and NR032 isolates respectively. Table 5.2 shows the different homopolymer variants of the four isolates and the size of the protein when translated in silico. The length of the homopolymer varied from 8 to 14 in these isolates. Within an isolate the sequence that flanks the homopolymer was identical. This putative sialidase gene is translated into the full-length protein (812 aa) when there are nine or twelve cytosines in the homopolymer region.

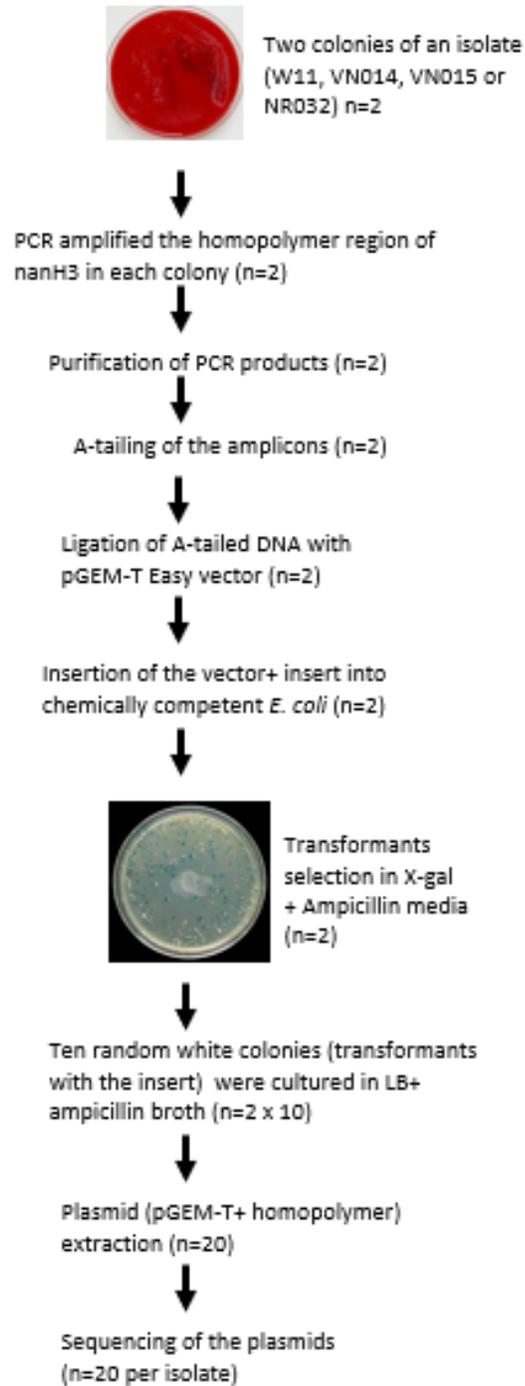


Figure 5.22: Study design of detecting the length of the homopolymer region of *nanH3*. Four isolates of *Gardnerella piotti* (cpn60 subgroup B) were used; W11, VN014, VN015, NR032. According to study design we ended up with 20 sequences of the homopolymer per isolate.

VN014 had three different homopolymer variants of 9, 10 or 11 cytosines. Of these, nine cytosines in the homopolymer produced the full length NanH3 that consists signal peptide, sialidase domain and trans-membrane domain. When there was 10 or 11 cytosines, a premature stop codon was generated after 129 bases producing a truncated protein of 43 aa in-silico. Having ten cytosines in the homopolymer region was the most common among the sequences of VN014 (Figure 5.23) and eleven was the least frequent.

VN015 isolate had the highest number of homopolymer variants. The length of the homopolymer varied from nine to thirteen producing the full-length protein when there were nine or twelve cytosines. Other sequences generated a stop codon after 129 bases generating a truncated peptide of 43 aa. Although ten cytosine residues in the homopolymer region produced a truncated protein, it was the most frequently observed variant in VN015 (Figure 5.23).

Two homopolymer variants were found in clones from NR032 isolates. The most frequent (15/16) length of the homopolymer was nine and it was predicted to translate the whole gene (Figure 5.23).

During examination of the predicted peptide sequences generated with various homopolymer lengths, we identified a second open reading frame within the *nanH3* open reading frame, starting at nucleotide 1191 and corresponding to amino acids 397-812 of the NanH3 protein. These 415 aa polypeptides encompass part of the predicted sialidase domain, and the transmembrane domain but lacks a signal peptide.

Table 5. 2: Homopolymer variants in *Gardnerella piotii* (cpn60 subgroup B) isolates

Isolate	Homopolymer sequence	Predicted protein size
W11	CAACTA- C ₁₁ -ATGAACAAA	Truncated protein 43 AA, 4.43 kDa
	CAACTA- C ₁₂ -ATGAACAAA	Full length protein 812 AA, 89.31 kDa
	CAACTA- C ₁₃ -ATGAACAAA	Truncated protein 38 AA, 3.8 kDa
	CAACTA- C ₁₄ -ATGAACAAA	Truncated protein 44 AA, 4.52 kDa
VN014	CAACTA- C ₉ -TCGAACAAA	Full length protein 812 AA, 89.31 kDa
	CAACTA- C ₁₀ -TCGAACAAA	Truncated protein 37 AA, 3.7 kDa
	CAACTA- C ₁₁ -TCGAACAAA	Truncated protein 43 AA, 4.43 kDa
VN015	CAACTA- C ₉ -TCGAACAAA	Full length protein 812 AA, 89.31 kDa
	CAACTA- C ₁₀ -TCGAACAAA	Truncated protein 37 AA, 3.7 kDa
	CAACTA- C ₁₁ -TCGAACAAA	Truncated protein 43 AA, 4.43 kDa
	CAACTA- C ₁₂ -TCGAACAAA	Full length protein 812 AA, 89.31 kDa
	CAACTA- C ₁₃ - -GAACAAA	Truncated protein 43 AA, 4.43 kDa
NR032	CAACTA- C ₈ -ATGAACAAA	Truncated protein 42 AA, 4.33 kDa
	CAACTA- C ₉ -ATGAACAAA	Full length protein 812 AA, 89.31 kDa

*Frequencies of occurrence of each sequence are shown in Figure 5.23

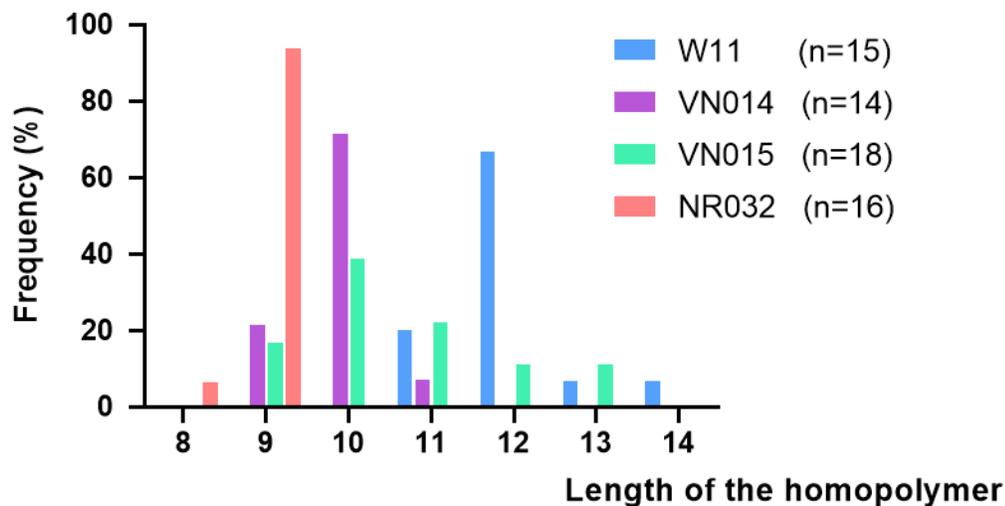


Figure 5. 23: Frequencies of the homopolymer variants in four *Gardnerella piotii* isolates. The length of the homopolymer varies from eight to fourteen. Frequencies of different homopolymer variants of each isolate (W11, VN014, VN015 or NR032) is indicated by colour.

5.5 Localization of sialidases in *Gardnerella* strains

Two putative sialidase genes (*nanH1* and *nanH3*) have been recognized among the collection of whole genome sequences of 36 *Gardnerella* spp. isolates (Patterson, 2018). Robinson *et al.* proposed the presence of three different sialidases and predicted the localization of them in *Gardnerella* spp. cells. They suggested that two of the sialidases (NanH2 and NanH3) of *Gardnerella* strain JCP8151B are likely to be extracellular based on the predicted domain structure of the amino acid sequences (Robinson *et al.*, 2019), however, there is no experimental evidence for the localization of sialidases in *Gardnerella* spp..

Our goal was to examine where NanH1 and NanH3 reside by immunogold staining and transmission electron microscopy (TEM). In order to visualize the localization of NanH1 by TEM, antibodies against NanH1 were raised in rabbits against the recombinant protein as described in the Methods. The specificity and the affinity of the antibodies were tested towards NanH1 (recombinant protein which was used to immunize the rabbits) by western blotting using fluorescence tagged anti-rabbit secondary antibodies (Figure 5.24). Both membranes showed a

prominent band at 101 kDa (the predicted mass of NanH1 based on its sequence) implying that polyclonal antibodies raised, bind to the epitopes of recombinant NanH1 protein.

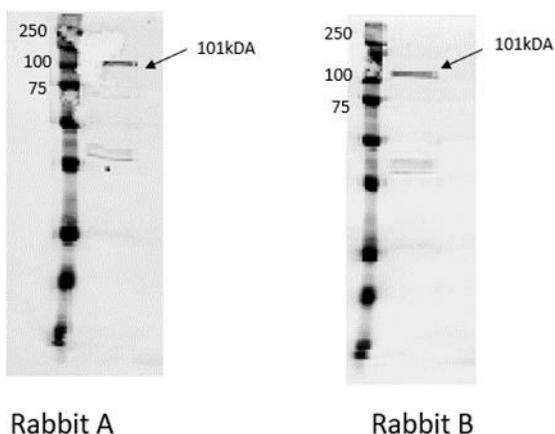


Figure 5.24: Western blot analysis of polyclonal antibodies raised in rabbit A and B. Size of the recombinant NanH1 protein is 101 kDa. The PVDF membranes blotted with recombinant NanH1 protein (10 $\mu\text{g/ml}$), showed a band at 101 kDa when treated with rabbit serum.

To determine the affinity and the specificity of the antibodies towards the native NanH1 protein, a western blot analysis was carried out using *Gardnerella piovii* strain W11 (cpn60 subgroup B) cell lysate (Figure 5.25). PVDF membrane treated with post-immune serum showed a prominent band at 75kDa, compared to the pre-immune serum (Figure 5.25). If W11 cells expressed *nanH1*, the size of the native protein (NanH1) was predicted to be 99.8 kDa. There was no band observed at ~100 kDa region in the PVDF membrane treated with post-immune serum. Although the polyclonal antibodies bind to the recombinant protein, they do not bind to a native protein of the predicted size but rather a different protein with an apparent molecular weight of ~75 kDa. This protein may represent a processed or modified form of *nanH1* product, or perhaps a different protein with related epitopes.

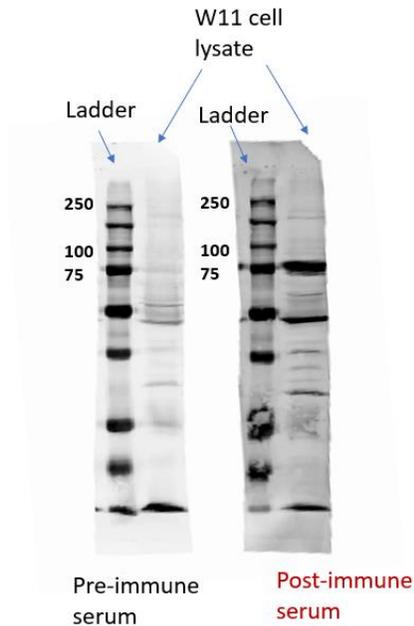


Figure 5.25: Western blot analysis of polyclonal antibodies raised in rabbit.

W11 cell lysate blotted with pre-immune and post-immune serum. In the PVDF membrane treated with post immune rabbit serum, a prominent band was not observed at the predicted size of NanH1 (99.8 kDa) but at 75 kDa.

To identify the protein that binds with the NanH1 polyclonal antibodies, the protein band was cut out from the gel and sent to CAGEF facility for protein identification. The identification of the peptide/s from acquired spectra was performed using a database search approach. Out of the many hits obtained by NCBI and UniProt, both databases provided transketolase (found in *Gardnerella* spp.) as the best match for the peptide sequence (NCBI Accession RDW95899.1, UniProt accession A0A3D8TB29).

Lack of availability of soluble NanH3 prevented antibody production against this protein.

5.6 Localization of sialidase activity in *Gardnerella* spp..

Presence of a predicted transmembrane domain and a signal peptide in NanH3 suggests that it is extracellular and anchored in the cell wall. In that case, sialidase activity should be associated with the cells, not the culture supernatant with a possibility of releasing the enzyme when cells are disrupted by sonication. If *Gardnerella* spp. have a sialidase that is secreted but not anchored to

the cell wall, the cell-free culture supernatant should be sialidase positive. To determine the cellular localization of sialidase enzymes in different *Gardnerella* spp., eight strains were selected. Four subgroup B, sialidase activity positive, *nanH3* positive isolates (W11, VN015, VN014 and NR032) and four subgroup C, sialidase activity negative, *nanH3* negative, *nanHI* positive isolates (VN001, VN007, NR038 and NR042) were included in this study (Table 4.1). These isolates were cultured on Columbia Sheep Blood-agar (BD, Germany) for 48 hours at 37 °C under anaerobic conditions. A few colonies of each isolate were used to inoculate 20 ml of NYC III broth and incubated for 48 hours at 37 °C under anaerobic conditions. After exactly 48 hours, the cultures were centrifuged at 4000 × g for 20 minutes to separate whole cells from the culture media. The supernatant was then removed, and the pellet was resuspended in PBS (pH 7.4) with 1 mg/ml lysozyme. The solution was incubated on ice for 15 minutes and sonicated for 3 minutes with 50% duty cycle. The suspension was clarified by centrifugation at 4000×g for 20 minutes. Whole cell pellets, culture supernatants, disrupted cell pellets and supernatants after sonication were tested in duplicates for sialidase activity using the quantitative assay.

As expected, no sialidase activity was detected in culture supernatants or cell pellets of subgroup C isolates NR042, NR038, VN001 and VN007 (Figure 5.26). In subgroup B isolates W11, VN015, VN014 and NR032 sialidase activity was detected in cell pellets (Figure 5.26). Although sialidase activity was detected in the clarified culture supernatant of VN015 isolate, no activity was seen in supernatant of other subgroup B isolates suggesting there are no secreted sialidases in isolates W11, VN014 and NR032.

When the cells were lysed by lysozyme and sonication higher levels of sialidase activity were detected in the pelleted cell debris in all four subgroup B isolates (Figure 5.27). In isolates VN015 and NR032, sialidase activity was detected in the supernatant after sonication, but at much lower levels than the corresponding pellets. The sialidase activity was very low in the supernatants of W11 and VN014 after sonication, suggesting that disrupting the cell wall of these isolates did not release sialidase enzyme to the medium. Even though subgroup C isolates possess an intracellular sialidase encoding gene (*nanHI*), they did not show sialidase activity in their supernatants after cell sonication.

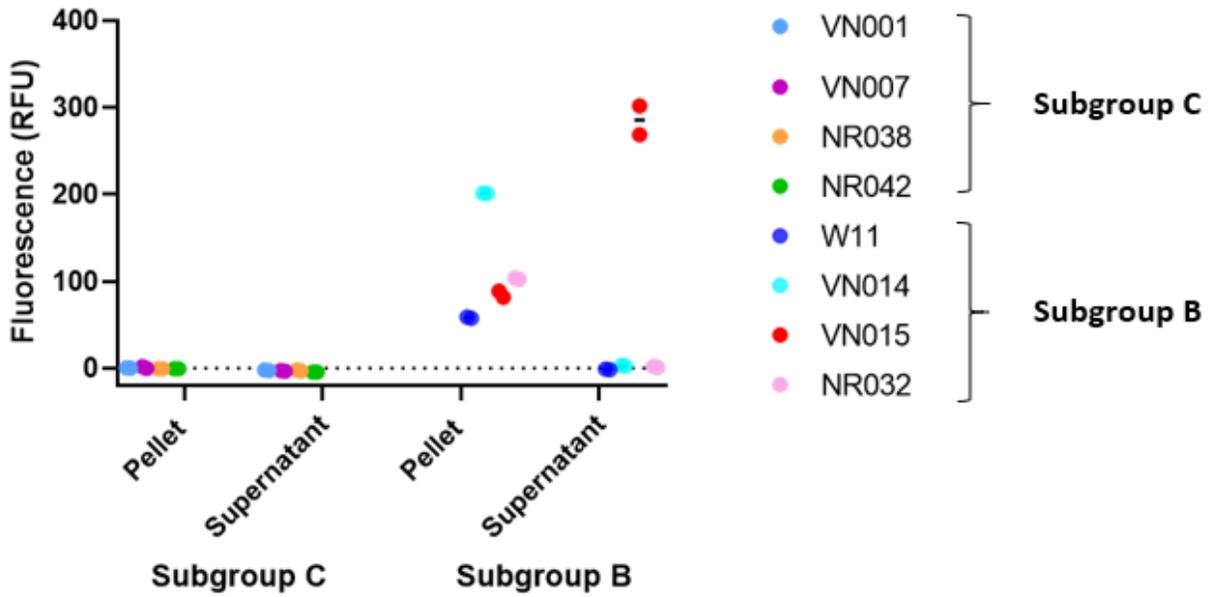


Figure 5.26: Localization sialidase activity of *Gardnerella* spp. isolates.

Quantitative assay was performed on cell pellets and culture supernatants of cpn60 subgroup C and B isolates. Y axis represents the difference of fluorescence (final fluorescence – initial fluorescence) after 1 hour of incubation. Each color represents an isolate and duplicate assay are indicated by colored circles.

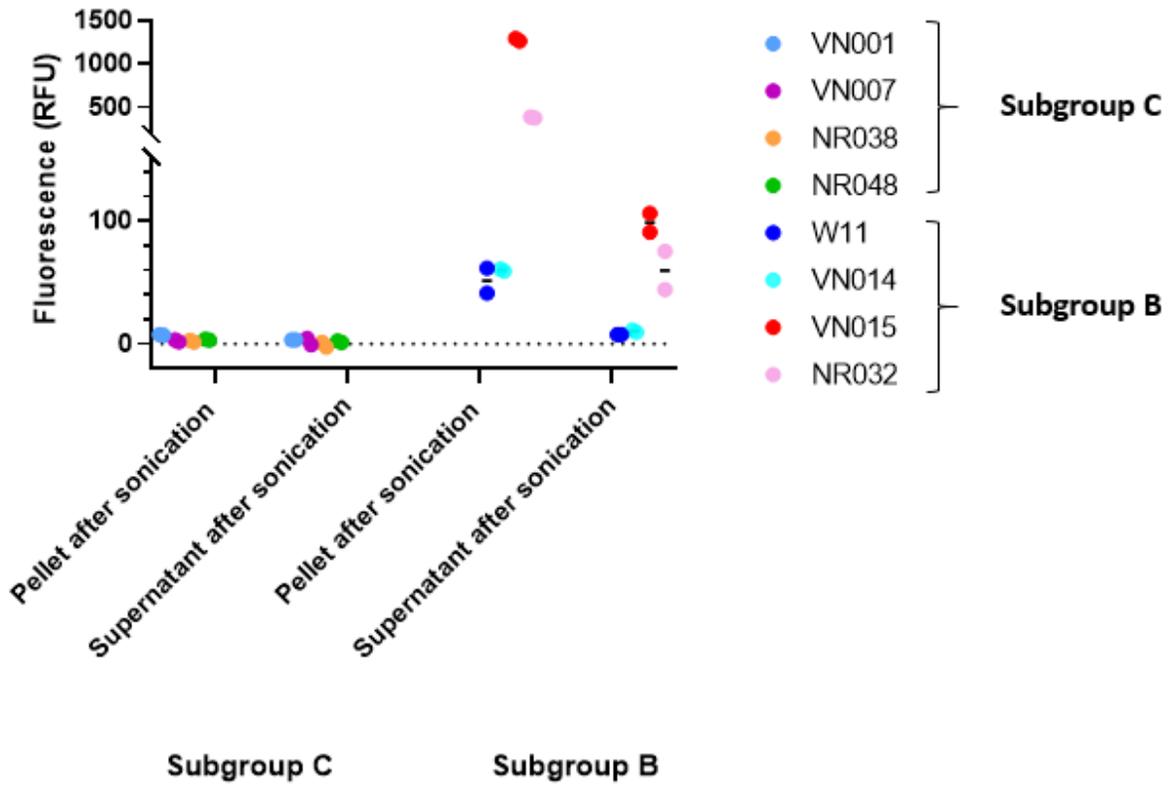


Figure 5.27: Localization sialidase activity of *Gardnerella* spp. isolates.

Quantitative assay was performed on pellets and supernatants after sonicating the cells of cpn60 subgroup C and B isolates. Y axis represents the difference of fluorescence (final fluorescence – initial fluorescence) after 1 hour of incubation. Each color represents an isolate and duplicate assay are indicated by colored circles.

6.0 Discussion

The genus *Gardnerella* is known as the hallmark of BV and abundantly observed in the vaginal fluid of women with BV (Swidsinski et al., 2005). Sialidase activity is an important virulence factor only found in *G. piotii* and *Gardnerella* genome species 3. Sialidase enzymes secreted by *Gardnerella* spp. can provide a carbon source by releasing sialic acid moieties from vaginal sialylated mucins (W. G. Lewis et al., 2013). Removal of sialic acid residues may reveal cryptic receptors and expose the underlying molecules for other enzymes and possible toxins (Galen et al., 1991). Thus, bacterial sialidases can contribute to degradation of mucus and alter its physical properties. Presence of sialidase could be beneficial to *Gardnerella* spp. in accessing nutrients or facilitating attachment to vaginal epithelial cells, initiating the formation of biofilm characteristic of BV.

6.1 Putative sialidase gene, *nanH3* was successfully expressed in *E. coli*, but the fusion proteins lacked sialidase activity.

By whole genome analysis we have identified two genes that potentially encode sialidases in *Gardnerella* spp; *nanH1* and *nanH3*. Previously, our lab group demonstrated the sialidase activity of the *nanH1* product (Patterson, 2018). Based on the amino acid sequence and the predicted domain structure of the sialidase encoded by *nanH1*, we assume it is an intra-cellular enzyme (Figure 2.4). In contrast, the sialidase (NanH3) that is predicted to be encoded by *nanH3*, has a signal peptide, sialidase domain and a transmembrane domain. When we compared the amino acid sequence of NanH3 of *Gardnerella* strain W11 to the exo- α -sialidase of *Bifidobacterium bifidum*, we observed 59% similarity in the sialidase domains. The degree of sequence similarity and the orientation of the predicted domains of the two sialidases is not surprising as *Gardnerella* spp. and *B. bifidum* belong to the same bacterial family, Bifidobacteriaceae. It is probable that NanH3 of *Gardnerella* sp. strain W11 has a similar role in adhesion to the mucosal surface as the extracellular sialidase of *B. bifidum* (Nishiyama et al., 2017).

Initial attempts to express NanH3 in *E. coli* had failed possibly due the presence of a homopolymer in the open reading frame and the difficulty to find a clone that does not generate a premature stop codon (Patterson, 2018). The rationale behind expression of the sialidase domain encoding region of *nanH3* was to avoid the homopolymer and the transmembrane encoding region. Sialidase

domain of *nanH3* was expressed in *E. coli* Top10 with a 6×His tag in the N-terminus. Although the recombinant protein was expressed (Figure 5.5) the amount of protein was insufficient for downstream assays.

When we expressed sialidase domain of NanH3 as a GST fusion protein (GST+SiaN3), the expression level was high (Figure 5.7), but the protein was insoluble (Figure 5.8B). We suspected that GST+SiaN3 was insoluble because of the inclusion bodies formed due to over expression of the protein. Even though the aggregates dissolved well in 1.5% Sarkosyl solution, GST+SiaN3 did not bind glutathione beads in the purification process. We suspect that this could be due to folding of the protein which could possibly sterically hinder the N-terminal GST from binding to the glutathione beads. The recombinant protein (GST+SiaN3) contained a C-terminal 6×His tag, and readily bound to a Ni-NTA affinity column when the buffers contained SDS and/or Sarkosyl. However, the purified protein did not show sialidase activity (Figure 5.10, 5.12). We suspect the lack of catalytic activity of GST+SiaN3 is because of the elution buffer contained ionic detergents (SDS and Sarkosyl). These ionic detergents affect the overall charge of the protein and make it soluble yet interfere with the folding, potentially disrupting the binding sites for the substrate. Removing the detergents by dialysis did not appear to help in terms of sialidase activity (Figure 5.13). Another possible reason for lack of activity of the sialidase domain could be that the gene region we cloned (sialidase domain of *nanH3*, 550 – 1554bp) does not translate to a protein with proper folding that shows sialidase activity.

To address this possibility, we decided to clone a larger part of *nanH3* lacking only the homopolymer region and the transmembrane domain (Truncated *nanH3*). Robinson et al. reported that a recombinant protein (a sialidase) could be insoluble when it carries a transmembrane domain (Robinson et al., 2019). Although truncated *nanH3* was expressed as a GST fusion protein (GST+TG2) the recombinant protein was still insoluble (Figure 5.17) possibly due to massive aggregation between the protein molecules.

The purpose of incorporating affinity tags like Histidine or GST tags into the DNA sequences is to facilitate the purification of the recombinant protein, but these tags can interfere with the solubility, conformation and other physical properties of the protein. Considering the size of the 6×His tag (0.8kDa), it is not believed to interact with the protein of interest (Uhl, Forsberg, Moks, Hartmanis, & Nilsson, 1992). His-tags can affect the 3-dimensional structure and cause significant reduction

in the enzymatic activity of enzymes (Mohanty & Wiener, 2004). Glutathione-S-transferase (GST) tag on the other hand is known as a protein stabilizer that prevents the target protein being degraded and helping to bring the protein to the soluble fraction (Kaplan, Husler, Klump, Sluis-cremer, & Dirr, 1997). However, GST fusion can often result in the target protein accumulating into inclusion bodies (Dümmler, Lawrence, & Marco, 2005). GST being a large tag (26 kDa) can affect catalytic sites of the recombinant enzyme. Removal of the GST tag from the recombinant protein by cleavage would be a potential solution to the insolubility of GST+SiaN3 and GST+TN3. The isoelectric point (PI) of a recombinant protein is significant in purification because it represents the pH where solubility is typically minimal. The pH of the solutions we used to purify GST+TN3 was close to its PI of 7.44, so that might have been a factor for poor solubility.

Although the sequence of NanH3 of *Gardnerella* strain JCP8151B investigated by Robinson et al. was very close to NanH3 of strain W11 (NanH3 of JCP8151B lacks signal peptide), we could not reproduce their expression results. This could be due to the usage of different expression vectors and *E. coli* strains in that study (Robinson et al., 2019).

6.2 The homopolymer region of *nanH3* varied in length within and between four isolates of *Gardnerella piovii* (cpn-60 subgroup B)

Sequence analysis of *nanH3* of *Gardnerella piovii* strain W11, revealed the presence of a homopolymer in the open reading frame (Patterson, 2018). This homopolymer consists about 12 cytosine residues and resides at 73-84 bp from the start site. Genomic regions that contain homogenous or heterogenous repeats are more prone to change in length of the repeat at each replication due to slipped strand mispairing (SSM). SSM is one of the mechanisms involved in phase variation which can influence the expression of a gene.

This is the first report to show evidence of slipped strand mispairing in *Gardnerella* spp.. Although *nanH3* of *Gardnerella* strain JCP8151B (Robinson et al., 2019) described by Robinson et al. had a similar sequence as *nanH3* of strain W11, it lacked the signal peptide encoding sequence where homopolymer is located.

To approach the research question, we first sequenced the PCR amplified homopolymer region from different colonies of W11. It was interesting to see the electropherograms (Figure 5.19) showing a mixture of signals after eleven peaks of cytosine (sense strand) suggesting each colony

had individuals with different lengths in the homopolymer region. Each colony consists of millions of individuals and at each replication, there is a chance of losing or adding one/more cytosine residue/s due to polymerase enzyme slippage. One explanation for the mixture of signals (Figure 5.19) could be due to polymerase enzyme slippage during PCR. Since we used proof-reading high fidelity Taq polymerase for PCR amplifications, PCR artifacts were minimized.

In this study, we demonstrated the presence of homopolymer variants not only in W11, but also in other *G. piotii* isolates: VN014, VN015 and NR032. According to our results (Figure 5.23) some lengths of the homopolymer region were more frequent than others. In isolates W11 and NR032, we found high abundance of homopolymer variants that potentially produce the full-length protein. In contrast, homopolymer variants that give rise to a truncated protein of isolates VN014 and VN015 were more abundant. As *nanH3* is predicted encode a sialidase that is tethered to the cell wall, producing the full-length protein would be beneficial to the organism for attachment or breaking down nutrients. *Gardnerella* spp. reside in an environment where there is constant interaction between the vaginal microbial community and the host mucosal immunity (Fichorova et al., 2013). Switching off an extracellular antigen could favour the bacteria to evade the host immune system. But we do not know the factors that triggers this switch in our culture conditions. More colonies from each isolate and more isolates should be assessed to conclude the actual prevalence of homopolymer lengths of the isolates.

Our data strongly suggest the existence of slipped strand mispairing at a genetic level in the four isolates of *Gardnerella piotii*. However, more experiments should be done to investigate if the genetic modifications affect the protein level. One approach could be using a sialylated protein (i.e: porcine mucin) as the sole carbon source in the cultures and assess the expression of *nanH3*.

6.3 Sialidase activity in *Gardnerella piotti* isolates is associated with the cell wall

For many years, *nanH1* (annotated as sialidase A in *G. vaginalis* strain ATCC 1419) was assumed to encode the enzyme that is responsible for the sialidase activity seen in the cultures (W. G. Lewis et al., 2013; Lopes Dos Santos Santiago et al., 2011). Although *nanH1* is found in all sialidase positive strains the intact open reading frame is also present in many sialidase negative strains (Schellenberg et al., 2016). Our lab group discovered another putative sialidase gene, *nanH3* by analyzing whole genome sequences of *Gardnerella* spp. (Patterson, 2018). When we began this study we hypothesized that *nanH1* encodes an intracellular sialidase and *nanH3* encodes an

extracellular sialidase that contributes to the sialidase activity observed in *Gardnerella* spp.. In 2019 Robinson et al. suggested that there are three sialidase genes (*nanH1*, *nanH2* and *nanH3*) and demonstrated only NanH2 and NanH3 (of strain JCP8151B) sialidases showed significant enzymatic activity. But the authors do not provide evidence of the localization of the three sialidases.

Our initial objective was to detect where the sialidase encoded by *nanH1* and *nanH3* reside by immuno-gold electron microscopy. Even though the antibodies we raised specifically bound recombinant NanH1 (101 kDa) with a high affinity (Figure 5.24), they did not bind the native protein of the same size. Instead, the antibodies reacted with a smaller sized protein (Figure 5.25) which was identified as a transketolase. Since we expressed *nanH1* in *E. coli*, it is possible to get recombinant protein contaminated with the host protein during Ni-NTA column purification process (Robichon, Luo, Causey, Benner, & Samuelson, 2011). When the rabbits were injected with the protein, antibodies might have raised against *E. coli* protein and possibly bound to a similar protein of *Gardnerella* species isolate W11. One way to remove host protein contaminants after Ni-NTA column purification is to expose the target protein to chitin beads (Robichon et al., 2011).

Since we were unable to pursue the localization of sialidases with the immune labeling method, we focused on if activity was associated with intact cells or culture supernatants. We selected four sialidase negative *Gardnerella vaginalis* (cpn60 subgroup C) isolates (VN001, VN007, NR038, NR048) and four sialidase positive *Gardnerella piovii* (cpn60 subgroup B) isolates (W11, VN014, VN015, NR032). The sialidase activity of these isolates were confirmed by filter spot test and the quantitative assay prior to the experiment. Our findings suggested in each isolate sialidase activity is more associated with cellular fraction than the supernatant, consistent with the responsible protein being tethered to the cell (Figure 5.26). The sialidase activity of the culture supernatant of VN015 could be due to the presence of an unbound sialidase (intracellular or secreted), but we suspect that it could be due to the presence of dead cells as we did not filter the supernatant. We tried to reduce the variabilities among the broth cultures of different isolates by growing them exactly 48 hours and adjusting the turbidity. Yet the number of cell numbers could have varied depending on the size of the cells of each isolate and a consequent lack of consistency in the relationship between OD and cell number. Sonication of the intact cells appeared to partially liberate sialidase/s to the soluble fraction in VN015 and NR032 but the sialidase activity was more

associated with the pellets than the supernatants of the respective isolates even after sonication. Von Nicolai and colleagues were able to liberate sialidase enzyme by ultrasonic treatment (Von Nicolai et al., 1984), but we only observed that in isolates VN015 and NR032. In all isolates, the sialidase activity was concentrated to the cellular fraction suggesting that there could be one or more sialidase/s that is/are tethered to cell wall of *Gardnerella* spp. Interestingly, supernatants after sonication of all four subgroup C isolates did not show sialidase activity even though they contain an intracellular sialidase encoding gene, *nanH1*. This could be due to regulation of expression of *nanH1* or the sialidase activity of NanH1 was not detectable under the given conditions.

6.4 Future prospects

In this study we faced some challenges in expressing *nanH3* of *Gardnerella* strain W11, purifying the recombinant protein and confirming its catalytic activity. As we mentioned in the discussion one possible reason for insolubility of GST+TN3 could be that the PI is closer to the buffers we used for purification. It is ideal to use a buffer with pH that makes the net charge of the protein negative or positive. To determine at which pH the protein become soluble, a range of buffers could be prepared with a pH gradient. The optimal pH which gives the highest solubility could be used for purification of GST+TN3. After purification GST tag would be cleaved off and the sialidase activity of the protein will be tested.

An approach to confirm localization of NanH3 would be extracting native extracellular cell bound proteins by cell shaving method (Marín, Haesaert, Padilla, Adán, & Hernández, 2018) and identifying the “shaved” protein using mass spectrometry.

In our study we demonstrated the existence of slipped strand mispairing in *nanH3* at genomic level. Future work in this regard would be focused on how this process alters protein expression in *Gardnerella piovii*.

References

- Ada, G. L., & French, E. L. (1959). Purification of Bacterial Neuraminidase (Receptor-destroying Enzyme). *Nature*, *1*, 3–4.
- Ahmed, A., Earl, J., Retchless, A., Hillier, S. L., Rabe, L. K., Cherpes, T. L., ... Ehrlich, G. D. (2012). Comparative genomic analyses of 17 clinical isolates of *Gardnerella vaginalis* provide evidence of multiple genetically isolated clades consistent with subspeciation into genovars. *Journal of Bacteriology*, *194*(15), 3922–3937.
- Andrewes, F. W. (1922). Studies in group -agglutination. The salmonella group and its antigenic structure. *The Journal of Pathology and Bacteriology*, *25*(4), 8–10.
- Benito, R., Vazquez, J. A., Berron, S., Fenoll, A., & Saez-Nieto, A. (1986). A modified scheme for biotyping *Gardnerella vaginalis*. *Journal of Medical Microbiology*, *21*, 357–359.
- Boskey, E. R., Cone, R. A., Whaley, K. J., & Moench, T. R. (2001). Origins of vaginal acidity: high D/L lactate ratio is consistent with bacteria being the primary source. *Hum Reprod*, *16*(9), 1809–1813.
- Briselden, A. M., Moncla, B. J., Stevens, C. E., & Hillier, S. L. (1992). Sialidases (Neuraminidases) in Bacterial Vaginosis and Bacterial Vaginosis-Associated Microflora. *Journal of Clinical Microbiology*, *30*(3), 663–666.
- Burnet, F. M., McCrea, J. F., & Stone, J. D. (1946). Modification of human red cells by virus action; the receptor gradient for virus action in human red cells. *British Journal of Experimental Pathology*, *27*(4), 228–22836.
- Camara, M., Boulnois, G. J., Andrew, P. W., & Mitchell, T. J. (1994). A neuraminidase from *Streptococcus pneumoniae* has the features of a surface protein. *Infection and Immunity*, *62*(9), 3688–3695.
- Cauci, S., Monte, R., Ropele, M., ero, C., Not, T., Quadrifoglio, F., & Menestrina, G. (1993). Pore-forming and haemolytic properties of the *Gardnerella vaginalis* cytoiysin. *Molecular Microbiology*, *9*(6), 1143–1155.
- Chaban, B., Links, M. G., Jayaprakash, T., Wagner, E. C., Bourque, D. K., Lohn, Z., ... Money,

- D. M. (2014). Characterization of the vaginal microbiota of healthy Canadian women through the menstrual cycle. *Microbiome*, 2(1), 23.
- Claire, M., Steven, J., & George, M. (1998). Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science*, 281, 375–378.
- Corfield, A. P., Wagner, S. A., Clamp, J. R., Kriaris, M. S., & Hoskin, L. C. (1992). Mucin degradation in the human colon: Production of sialidase, sialate O- acetyltransferase, N-acetylneuraminidase, arylesterase, and glycosulfatase activities by strains of fecal bacteria. *Infection and Immunity*, 60(10), 3971–3978.
- Corfield, T. (1992). Bacterial sialidases — roles in pathogenicity and nutrition. *Glycobiology*, 2(6), 509–521.
- Criswell, B. S., Ladwig, C. L., Gardner, H. L., & Dukes, C. D. (1969). *Haemophilus vaginalis*: Vaginitis by Inoculation from Culture. *Obstetrics and Gynecology*, 33(2), 195–199.
- Criswell, B. S., Marston, J. H., Stenback, W. A., Black, S. H., & Gardner, H. L. (1971). *Haemophilus vaginalis* 594, a Gram-negative organism? *Journal of Microbiology*, 17, 865–869.
- Dechering, K. J., Cuelenaere, K., Konings, R. N. H., & Leunissen, J. A. M. (1998). Distinct frequency-distributions of homopolymeric DNA tracts in different genomes. *Nucleic Acids Research*, 26(17), 4056–4062.
- Deitsch, K. W., Lukehart, S. A., & Stringer, J. R. (2009). Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. *Nature Reviews Microbiology*, 7(7), 493–503.
- Donders, G. G. G., Vereecken, A., Dekeersmaecker, A., & Bulck, B. Van. (2000). Wet mount microscopy reflects functional vaginal lactobacillary flora better than Gram stain. *Journal of Clinical Pathology*, 53, 308–313.
- Dümmler, A., Lawrence, A., & Marco, A. De. (2005). Simplified screening for the detection of soluble fusion construct expressed in *E. coli* using a modular set of vectors. *Microbial Cell Factories*, 10, 1–10.

- Egan, M., Motherway, O. C., Ventura, M., & Sinderen, V. (2014). Metabolism of Sialic Acid by *Bifidobacterium breve* UCC2003. *Journal of Applied and Environmental Microbiology*, 80(14), 4414–4426.
- Fichorova, R. N., Buck, O. R., Yamamoto, H. S., Fashemi, T., Dawood, H. Y., Fashemi, B., ... Onderdonk, A. B. (2013). The Villain Team-Up or how *Trichomonas vaginalis* and bacterial vaginosis alter innate immunity in concert. *Sexually Transmitted Infections*, 89, 460–466.
- Galen, J. E., Ketley, J. M., Fasano, A., Richardson, S. H., Wasserman, S. S., & Kaper, J. B. (1991). Role of *Vibrio cholerae* neuraminidase in the function of cholera toxin. *Infection and Immunity*, 60(2), 406–415.
- Gelber, S. E., Aguilar, J. L., Lewis, K. L. T., & Ratner, A. J. (2008). Functional and phylogenetic characterization of vaginolysin, the human-specific cytolysin from *Gardnerella vaginalis*. *Journal of Bacteriology*, 190(11), 3896–3903.
- Giacani, L., Lukehart, S., & Centurion-lara, A. (2007). Length of guanosine homopolymeric repeats modulates promoter activity of subfamily II tpr genes of *Treponema pallidum* ssp . pallidum. *FEMS Immunology and Medical Microbiology*, 51, 289–301.
- Gottschalk, A. (1957). Neuraminidase: the specific enzyme of influenza virus and *Vibrio cholerae*. *Journal of Chemical Society*, 23, 645–646.
- Greenwood, J. R., & Pickett, M. J. (1980). Transfer of *Haemophilus vaginalis* Gardner and Dukes to a New Genus, *Gardnerella*; *G.vaginalis* (Gardner and Dukes) comb.nov. *International Journal of Systematic Bacteriology*, 30(1), 170–178.
- Gutman, R. E., Peipert, J. F., Weitzen, S., & Blume, J. (2005). Evaluation of Clinical Methods for Diagnosing Bacterial Vaginosis. *American Journal of Obstetrics and Gynecology*, 105(3), 551–556.
- Hagblom, P., Segalt, E., & Billyardt, E. (1985). Intragenic recombination leads to pilus antigenic variation in *Neisseria gonorrhoeae*. *Nature*, 315, 2–4.
- Hardy, L., Jaspers, V., Bulck, M. Van Den, Buyze, J., Mwambarangwe, L., Musengamana, V., ... Crucitti, T. (2017). The presence of the putative *Gardnerella vaginalis* sialidase A gene in vaginal specimens is associated with bacterial vaginosis biofilm. *PLOS ONE*, 1–11.

- Heimer, R., & Meyer, K. (1956). Studies on Sialic Acid of Submaxillary Mucoïd. *Proceedings of the National Academy of Sciences of the United States of America*, 42(10), 728–734.
- Hickey, R. J., Zhou, X. I. A., Pierson, J. D., Ravel, J., & Forney, L. J. (2012). Understanding vaginal microbiome complexity from an ecological perspective. *Translational Research*, 160(4), 267–282.
- Hill, J. E., & Albert, A. Y. (2019). Resolution and Cooccurrence Patterns of *Gardnerella leopoldii*, *G. swidsinskii*, *G. piotii*, and *G. vaginalis* within the Vaginal Microbiome. *Infection and Immunity*, 87(12), 1–12.
- Hill, J. E., Goh, H., Money, D. M., Doyle, M., Li, A., Crosby, W. L., ... Hemmingsen, S. M. (2005). Characterization of vaginal microflora of healthy, nonpregnant women by chaperonin-60 sequence-based methods. *American Journal of Obstetrics and Gynecology*, 193, 682–692.
- Hillier, S. L., Nugent, R. P., Eschenbach, D. A., Krohn, M. A., Gibbs, R., Martin, D. H., ... Klebanoff, M. A. (1996). “Association between Bacterial Vaginosis and Preterm Delivery of a Low-birth-weight Infant.” *The New England Journal of Medicine*, 333(26), 1737–1742.
- Hirst, G. K. (1941). The Agglutination of Red Cells by Allantoic Fluid of Chick Embryos Infected with Influenza Virus. *Science*, 94(2427), 22–23.
- Ingianni, A., Petruzzelli, S., Morandotti, G., & Pompei, R. (1997). Genotypic differentiation of *Gardnerella vaginalis* by amplified ribosomal DNA restriction analysis (ARDRA). *FEMS Immunology and Medical Microbiology*, 18(1), 61–66.
- Jayaprakash, T. P., Schellenberg, J. J., & Hill, J. E. (2012). Resolution and Characterization of Distinct cpn60-Based Subgroups of *Gardnerella vaginalis* in the Vaginal Microbiota. *PLOS ONE*, 7(8), 1–11.
- Joesoef, M. R., Hillier, S. L., Josodiwondo, S., & Linnan, M. (1991). Reproducibility of a scoring system for gram stain diagnosis of bacterial vaginosis. *Journal of Clinical Microbiology*, 29(8), 1730–1731.
- Kaplan, W., Husler, P., Klump, H., Sluis-cremer, N., & Dirr, H. (1997). Conformational stability of pGEX-expressed *Schistosoma japonicum* glutathione S-transferase: A detoxification

- enzyme and fusion-protein affinity tag. *Protein Science*, 6, 399–406.
- Kenyon, C., Colebunders, R., & Crucitti, T. (2013). The global epidemiology of bacterial vaginosis : a systematic review “ B. *The American Journal of Obstetrics & Gynecology*, 209(6), 505–523.
- Kiyohara, M., Tanigawa, K., Chaiwangsri, T., Katayama, T., Ashida, H., & Yamamoto, K. (2011). An exo - Sialidase from bifidobacteria involved in the degradation of sialyloligosaccharides in human milk and intestinal glycoconjugates. *Glycobiology*, 21(4), 437–447.
- Koumans, E. H., Sternberg, M., Bruce, C., McQuillan, G., Kendrick, J., Sutton, M., & Markowitz, L. E. (2007). The prevalence of bacterial vaginosis in the United States, 2001-2004; associations with symptoms, sexual behaviors, and reproductive health. *Sexually Transmitted Diseases*, 34(11), 864–869. <https://doi.org/10.1097/OLQ.0b013e318074e565>
- Lewis, A. L., & Lewis, W. G. (2012). Microreview Host sialoglycans and bacterial sialidases : a mucosal perspective. *Clinical Microbiology*, 14(8), 1174–1182.
- Lewis, W. G., Robinson, L. S., Gilbert, N. M., Perry, J. C., & Lewis, A. L. (2013). Degradation , Foraging , and Depletion of Mucus Sialoglycans by the Vagina-adapted Actinobacterium *Gardnerella*. *The Journal of Biological Chemistry*, 288(17), 12067–12079.
- Lopes Dos Santos Santiago, G., Deschaght, P., El Aila, N., Kiama, T. N., Verstraelen, H., Jefferson, K. K., ... Vanechoutte, M. (2011). *Gardnerella vaginalis* comprises three distinct genotypes of which only two produce sialidase. *American Journal of Obstetrics and Gynecology*, 204(5), 450.e1-450.e7.
- Lu, H., Jiang, P., Zhang, X., Hou, W., Wei, Z., Lu, J., ... Han, Y. (2015). Characteristics of bacterial vaginosis infection in cervical lesions with high risk human papillomavirus infection. *Int J Clin Exp Med*, 8(11), 21080–21088.
- Ma, Z. S., Li, L., & Gotelli, N. J. (2019). Diversity-disease relationships and shared species analyses for human microbiome-associated diseases. *The ISME Journal*, 13, 1911–1919.
- Marín, E., Haesaert, A., Padilla, L., Adán, J., & Hernáez, M. L. (2018). Unraveling *Gardnerella vaginalis* Surface Proteins Using Cell Shaving Proteomics. *Frontiers in Microbiology*,

9(May), 1–15.

- Martin-Platero, A. M., Valdivia, E., Maqueda, M., & Martinez-Bueno, M. (2007). Fast , convenient , and economical method for isolating genomic DNA from lactic acid bacteria using a modification of the protein ““ salting-out ”” procedure. *Analytical Biochemistry*, *366*, 102–104.
- Mattarelli, M. M. B. B. P. (2006). Occurrence of the Family Bifidobacteriaceae in Human Dental Caries and Plaque. *Carries Research; Basel*, *40*(3), 271–276.
- May, M., & Brown, D. R. (2009). Secreted sialidase activity of canine mycoplasmas. *Veterinary Microbiology*, *137*, 380–383.
- Mcgregor, J., French, J., Jones, W., Milligan, K., Mckinney, P., Patterson, E., & Parker, R. (1994). Bacterial vaginosis is associated with prematurity and vaginal fluid mucinase and sialidase : results of a controlled trial of topical clindamycin cream . *American Journal of Obstetrics and Gynecology*, *170*(4), 1048–1059.
- Mohanty, A. K., & Wiener, M. C. (2004). Membrane protein expression and production : effects of polyhistidine tag length and position. *Protein Expression and Purification*, *33*, 311–325.
- Myziuk, L., Romanowski, B., & Johnson, S. C. (2003). BVBlue Test for Diagnosis of Bacterial Vaginosis. *Journal of Clinical Microbiology*, *41*(5), 1925–1928.
- Newton, E. R. (1997). Wet Smear Compared With Gram Stain Diagnosis of Bacterial Vaginosis in Asymptomatic Pregnant Women. *Obstetrics and Gynecology*, *78*44(00), 504–506.
- Nishiyama, K., Yamamoto, Y., Sugiyama, M., Takaki, T., Urashima, T., Fukiya, S., ... Mukai, T. (2017). *Bifidobacterium bifidum* Extracellular Sialidase Enhances Adhesion to the Mucosal Surface and Supports Carbohydrate Assimilation. *American Society for Microbiology*, *8*(5), 1–15.
- Nugent, R. P., Krohn, M. A., & Hillier, S. L. (1991). Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *Journal of Clinical Microbiology*, *29*(2), 297–301.
- Olmsted, S. S., Meyn, L. A., Rohan, L. C., & Hillier, S. L. (2003). Glycosidase and Proteinase

- Activity of Anaerobic Gram-Negative Bacteria Isolated From Women With Bacterial Vaginosis. *Sexually Transmitted Diseases*, 30(3), 257–261.
- Patterson, M. (2018). *Characterization of putative sialidase genes in Gardnerella vaginalis*. University of Saskatchewan. Retrieved from <https://harvest.usask.ca/handle/10388/9587>
- Piot, P., Dyck, E. V. A. N., Peeters, M., Hale, J., Totten, P. A., & Holmes, K. (1984). Biotypes of *Gardnerella vaginalis*. *Journal of Clinical Microbiology*, 20(4), 677–679.
- Pleckaityte, M., Janulaitiene, M., Lasickiene, R., & Zvirbliene, A. (2012). Genetic and biochemical diversity of *Gardnerella vaginalis* strains isolated from women with bacterial vaginosis. *FEMS Immunology and Medical Microbiology*, 65, 69–77.
- Ravel, J., Gajer, P., Abdo, Z., Schneider, G. M., Koenig, S. S. K., McCulle, S. L., ... Forney, L. J. (2011). Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences*, 108(Supplement_1), 4680–4687.
- Reyn, A., & Birch-Anderson, A. (1966). An electron microscope study of thin sections of *Haemophilus vaginalis* (gardner and dukes) and some possibly related species. *Canadian Journal of Microbiology*, 12.
- Robichon, C., Luo, J., Causey, T. B., Benner, J. S., & Samuelson, J. C. (2011). Engineering *Escherichia coli* BL21 (DE3) Derivative Strains to Minimize *E. coli* Protein Contamination after Purification by Immobilized Metal Affinity Chromatography. *Applied and Environmental Microbiology*, 77(13), 4634–4646.
- Robinson, L. S., Schwebke, J., Lewis, W. G., & Lewis, A. L. (2019). Identification and characterization of NanH2 and NanH3, enzymes responsible for sialidase activity in the vaginal bacterium *Gardnerella vaginalis*. *JBC*, 294, 5230–5245.
- Roggentin, P., Kleineidam, R. G., & Schauer, R. (1995). Diversity in the Properties of Two Sialidase Isoenzymes Produced by *Clostridium perfringens* SPP. *Biological Chemistry Hoppe-Seyler*, 376(9), 569–576.
- Rup, L. (2012). The Human Microbiome Project. *Indian Journal of Microbiology*, 52(3), 315.
- Sadhu, K., Domingue, P. A. G., Chowt, A. W., Nelligan, J., Chengt, N., & Costerton, J. W.

- (1989). *Gardnerella vaginalis* has a gram-positive cell-wall ultrastructure and lacks classical cell-wall lipopolysaccharide. *Journal of Medical Microbiology*, 29, 229–235.
- Schellenberg, J. J., Jayaprakash, T. P., Gamage, N. W., Patterson, M. H., Vaneechoutte, M., & Hill, J. E. (2016). *Gardnerella vaginalis* Subgroups Defined by cpn 60 Sequencing and Sialidase Activity in Isolates from Canada, Belgium and Kenya. *PLOS ONE*, 1–12.
- Seifert, H. S., Ajioka, R. S., Marchal, C., Sparling, P. F., & So, M. (1988). DNA transformation leads to pilin antigenic variation in *Neisseria gonorrhoeae*. *Nature*, 336, 4–7.
- Sela, A. D. A., Chapman, J., Adeuya, A., Kim, J. H., Chen, F., Whitehead, T. R., ... Mills, D. A. (2008). The Genome Sequence of *Bifidobacterium longum* subsp. *infantis* Reveals Adaptations for milk utilization within the infant microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, 105(48).
- Sewankambo, N., Gray, R. H., Wawer, M. J., Paxton, L., McNairn, D., Wabwire-Mangen, F., ... Konde-Lule, J. (1997). HIV-1 infection associated with abnormal vaginal flora morphology and bacterial vaginosis. *Lancet*, 350(9077), 546–550.
- Smayevsky, J., Canigia, L. F., & Lanza, A. (2001). Vaginal microflora associated with bacterial vaginosis in nonpregnant women : reliability of sialidase detection. *Infectious Diseases in Obstetrics and Gynecology*, 17–22.
- Stern, A., Brown, M., Nickel, P., & Meyer, T. F. (1983). Opacity Genes in *Neisseria gonorrhoeae* : Control of Phase and Antigenic Variation. *Cell*, 47, 61–71.
- Stern, A., Nickel, P., & Meyer, T. F. (1984). Opacity Determinants of *Neisseria gonorrhoeae* : Gene Expression and Chromosomal Linkage to the Gonococcal Pilus Gene. *Cell*, 37(June 1964), 447–456.
- Swidsinski, A., Mendling, W., Loening-Baucke, V., Ladhoff, A., Swidsinski, S., Hale, L. P., & Lochs, H. (2005). Adherent Biofilms in Bacterial Vaginosis. *Obstet Gynecol*, 106(5), 1013–1023.
- Turroni, F., Bottacini, F., Foroni, E., Mulder, I., Kim, J.-H., Zomer, A., ... Klaenhammer, T. R. (2010). Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-derived glycan foraging. *National Academy of Sciences*, 107(45), 19514–19519.

- Uhl, M., Forsberg, G., Moks, T., Hartmanis, M., & Nilsson, B. (1992). Fusion proteins in biotechnology. *Current Opinion in Biotechnology*, 3, 363–369.
- Vaneechoutte, M., Guschin, A., Simaey, L. Van, Gansemans, Y., Nieuwerburgh, F. Van, & Cools, P. (2019). Emended description of *Gardnerella vaginalis* and description of *Gardnerella leopoldii* sp. nov., *Gardnerella plotii* sp. nov. and *Gardnerella swidsinskii* sp. nov., with delineation of 13 genomic species within the genus *Gardnerella*. *Systematic and Evolutionary Microbiology*, 69, 679–687.
- Von Nicolai, H., Hammann, R., Salehnia, S., & Zilliken, F. (1984). A newly discovered sialidase from *Gardnerella vaginalis*. *Zentralbl Bakteriell Mikrobiol Hyg A*, 258(1), 1984.
- Woude, M. W. Van Der, & Baumber, A. J. (2004). Phase and Antigenic Variation in Bacteria. *Clinical Microbiology Reviews*, 17(3), 581–611.
- Zieg, J., Silverman, M., Hilmen, M., & Simon, M. (1977). Recombinational Switch for Gene Expression. *Science*, 196(4286), 170–172.