The effect of Bacterial Vaginosis on HIV infection

by

Bianca Abrahams

Thesis presented for the degree of

DOCTOR OF PHILOSOPHY

in the Department of Integrative Biomedical Sciences

Faculty of Health Sciences

With Associate Professor Zenda Woodman

and Dr Brian Kullin

December 2023

Table of Contents

Declaration

I, Bianca Abrahams, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I have used the referencing style from the journal Nature. Each significant contribution to, and quotation in, this dissertation from the work(s) of other people has been attributed, cited and referenced.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature:

Acknowledgements

I would like to thank the following people and institutions:

The National Research Foundation, Medical Research Council, Poliomyelitis Research Foundation and the University of Cape Town for funding my studies and my research.

Hannah Livingstone for the isolation of the clinical *Gardnerella* strains.

Sylva Schwager and Vinasha Ramasamy for assisting with the speedvac.

Viantha Naidoo, Dirk Lang and Mike Reiche for assistance with the confocal microscopes.

Maré Vlok for Mass spec analysis and the team at CAF for sequencing.

Melissa-Rose Abrahams for always being willing to help with equipment and reagents.

Everyone on level 6 in Falmouth but especially Shaheeda, Ebbie, Graham, Abdu, Mara, Joyce and Ricki for ensuring that everything ran smoothly in the lab and for going out of their way when I really needed help.

To all my fellow students, current and past, especially Raabs, Shabs, Gabriella, Caryn Rajh and Arad for the support, encouragement and laughs.

To my fellow PhD students, Welcome, Hendrina, Nonku, Jamie, Sinead, Claire and Deeds. Not many people will ever understand this journey, thank you for the support and help with experiments and reagents. The biggest thank you to Humaira who has been by my side almost every step and every day of my postgraduate experience and has supported me through the toughest of times. For always being willing to help me with experiments, reagents and equipment and being a great listener when I needed to complain, every day.

To my co-supervisor, Brian, thank you for your willingness to teach and help whenever needed. Your quiet demeanour and kindness did not go unnoticed and I am extremely thankful for all the guidance you have provided.

To my family, my mom and dad for making my education at UCT possible and always encouraging me to further my education. And for absolutely everything that simply can't be put into words. To my sister, for inspiring me to pursue a PhD by showing me it is possible rather than telling me.

To Daiyaan, for sticking by my side through 13 years of studying and never once questioning why it was taking so long. Your patience, your encouragement and your belief in me is what prevented me from giving up in my darkest of times and I am forever grateful to you.

Lastly, and most importantly, to my supervisor, Zenda. I can't thank you enough for the past 8 years. It's been the best and worst of my life, but you have always encouraged and supported me, unconditionally. I am forever grateful for all the knowledge you have imparted on me which is so invaluable and unique because you are a one-of-a-kind human. Your guidance and support and care is certainly unmatched by any other supervisor. Words will never be enough to thank you for everything you have done for me. I hope we will have the opportunity to work together in the future because I really don't know what you'll do without me.

Abstract

Human Immunodeficiency Virus-1 (HIV) and Bacterial Vaginosis (BV) are both among the most common diseases affecting young women in Sub-Saharan Africa. BV is characterised by dysbiosis in the female reproductive tract (FRT) when optimal *Lactobacillus* spp. such as *L. crispatus,* are displaced by anaerobes such as *Gardnerella* spp., consistently isolated from the FRT of BV-positive women. *Gardnerella* spp. produce a number of important virulence factors such as vaginolysin (VLY) and sialidase and are known to initiate the formation of biofilms and hence, contribute to the pathology associated with BV. As many studies have suggested that BV increases susceptibility to HIV infection, it stands to reason that *Gardnerella* spp. might either indirectly, (initiate the onset of BV and subsequent immune responses) or directly (via its virulence factors) play a role in enhancing HIV acquisition. This study investigated whether VLY, sialidase and biofilm formation played a role in HIV infection. Recently, it was discovered that *Gardnerella* spp. comprises *G. leopoldii, G. vaginalis, G. piotii* and *G. swidsinskii* as well as nine other genome species which vary in virulence potential. When we compared twenty strains isolated from BV-positive women belonging to *G. vaginalis* (n = 16), *G. piotii* (n = 2) and *G. swidsinskii* (n = 2), we found differences in the presence, expression, and activity of VLY and sialidase as well as biofilm-forming capacity between the strains, suggesting a wide range in virulence. However, there was no overt association between HIV infection and *Gardnerella* virulence factors. *Gardnerella* sialidase consists of three isoforms: NanH1, NanH2 and NanH3 and as the latter is responsible for the sialidase activity in the FRT, *nanH3* was cloned, expressed in *E.coli* and purified by His-Tag affinity chromatography. We found that purified, recombinant NanH3 increased HIV infection *in vitro,* most likely by removing the sialic acid moieties on the surface of host cells, reducing the negative repulsive force between the viral Envelope and cell membrane. This may then facilitate virus accumulation at the cell surface, favouring attachment to CD4 and/or its co-receptors and thereby enhance HIV infection. Interestingly, only the two *G. piotii* strains expressed NanH3, suggesting that perhaps only some species of *Gardnerella* may play a role in enhancing HIV infection. This has important implications for diagnosis and treatment as BV-positive women shown to carry *G. piotii* strains, might benefit from preexposure prophylaxis with antiretroviral drugs.

List of abbreviations

µg	Microgram
µl	Microlitre
µM	Micromolar
µm	Micron
AIDS	Acquired immunodeficiency syndrome
ARDRA	Amplified ribosomal DNA restriction analysis
ATCC	American type culture collection
BAP	Biofilm associated protein
BCA	Bicinchoninic acid
bp	Base pair
BPROM	Predictor of bacterial promoter
BV	Bacterial vaginosis
BVAB	Bacterial vaginosis associated bacteria
℃	Degrees celius
CAPRISA	Centre for the AIDS Program in South Africa
CD59	Complement regulatory molecule 59
CDC	Cholesterol-dependent cytolysin

cDNA	Complementary deoxyribonucleic acid
CE	Crude extract
Cfu	Colony forming units
cm	Centimeter
CO_2	Carbon dioxide
Cpn60	Chaperonin protein 60
CVM	Cervico vaginal mucus
DANA	2,3-dehydro-2-deoxy-N-acetylneuraminic acid
DD	Double digested
DEPC	Diethyl pyrocarbonate
DMEM	Dulbecco's modified eagle's medium
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
eGFP	Enhanced green fluorescent protein
Env	Envelope
EtOH	Ethanol
FBS	Fetal bovine serum
FGT	Female genital tract

FRT	Female reproductive tract
FT	Flow through
g	g-force
gDNA	Genomic deoxyribonucleic acid
GFP	Green fluorescent protein
H_2O	Hydrogen oxide
H_2O_2	Hydrogen peroxide
HEK293T	Human embryonic kidney
HIV-1	Human Immuno Deficiency Virus-1
H-NS	Histone-like nucleiod structuring protein
hOAT	Human organic anion transporter
HSV-2	Herpes Simplex Virus-2
IgA	Immunoglobulin A
IL-10	Interleukin-10
IL-8	Interleukin-8
IMC	Infectious molecular clone
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	Killodalton

KOH	Potassium hydroxide
LB	Luria broth
LTR	Long terminal repeat
MFI	Mean fluorescent intensity
mg	Miligrams
$MgCl_2$	Magnesium chloride
ml	mililitre
mM	Milimolar
mm	Milimeter
MSM	Men who have sex with men
MTT	3- [4,5-dimethylthiazol-2-yl]-2,5 diphenyl tetrazolium bromide
MUN	4-methylumbelliferyl-α-d-N-acetylneuraminic acid
MW	Molecular weight
ng	nanogram
ng	Nanograms
nm	Nanometers
NTC	No template control
NYC III	New york city three

OD_{600}	Optical density 600
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEI	Polyethylenimine
Pgi	Phosphoglucose isomerase/phosphoglucoisomerase
PMSF	Phenylmethylsulfonyl fluoride
PrEP	Pre-exposure Prophylaxis
PVDF	Polyvinyldene difluoride
qPCR	Quantitative polymerase chain reaction
RC-101	Retrocyclin 101
RFU	Relative fluorescent units
RLU	Relative light units
RNA	Ribonucleic acid
rpm	Rotations per minute
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
sIgA	Secreted immunoglobulin A
STD	Sexually transmitted disease
STI	Sexually transmitted infection

TCEP	Triscarboxyethyl phosphine
TCID50	Tissue culture infectious dose
U	Units
UD	Undigested
UTI	Urinary tract infection
V	Volts
VLP	Virus-like particle
VLY	Vaginolysin
w/v	Weight/volume
WCBS	Western Cape blood service

Chapter 1: Literature review

1.1 Introduction

Human immunodeficiency virus-1 (HIV-1) is a global epidemic and one of the leading causes of death worldwide [11]. Sub-Saharan Africa has the highest incidence of HIV infection which is most likely due to multiple risk factors, such as gender-based violence, economic disempowerment, lack of condom usage, drug abuse, number of sexual partners, infection by other sexually transmitted diseases (STDs) and more recently, bacterial vaginosis (BV) [12-15]. Commensal microbiota of the female reproductive tract (FRT), such as *Lactobacillus* spp., provide protection against HIV infection, and hence the displacement of these species by anaerobic bacteria, such as *Gardnerella* spp. and *Prevotella bivia* might increase the chance of viral infection [15]. It is also possible that *Gardnerella* spp. might directly enhance HIV infection through the action of its virulence determinants. Comparison of the global distribution of HIV and BV supports the relationship between HIV infection and BV although the mechanism remains unknown (Figure 1.1).

1.2 Global distribution of HIV infection

HIV was first described over 30 years ago [16] and now has a large, but unequal, global distribution [17,18]. In 2021, UNAIDS reported that globally, 4000 adults and children become infected daily, 60 % of which occur in sub-Saharan Africa [4]. The affected cohort is further skewed towards young women of child-bearing age [19] and despite early diagnosis, treatment and the availability of pre-exposure prophylaxis (PrEP), there were still 39 million people living with HIV, globally, in 2022 [20]. The HIV pandemic is thus not yet resolved highlighting the need for research into factors increasing susceptibility to infection.

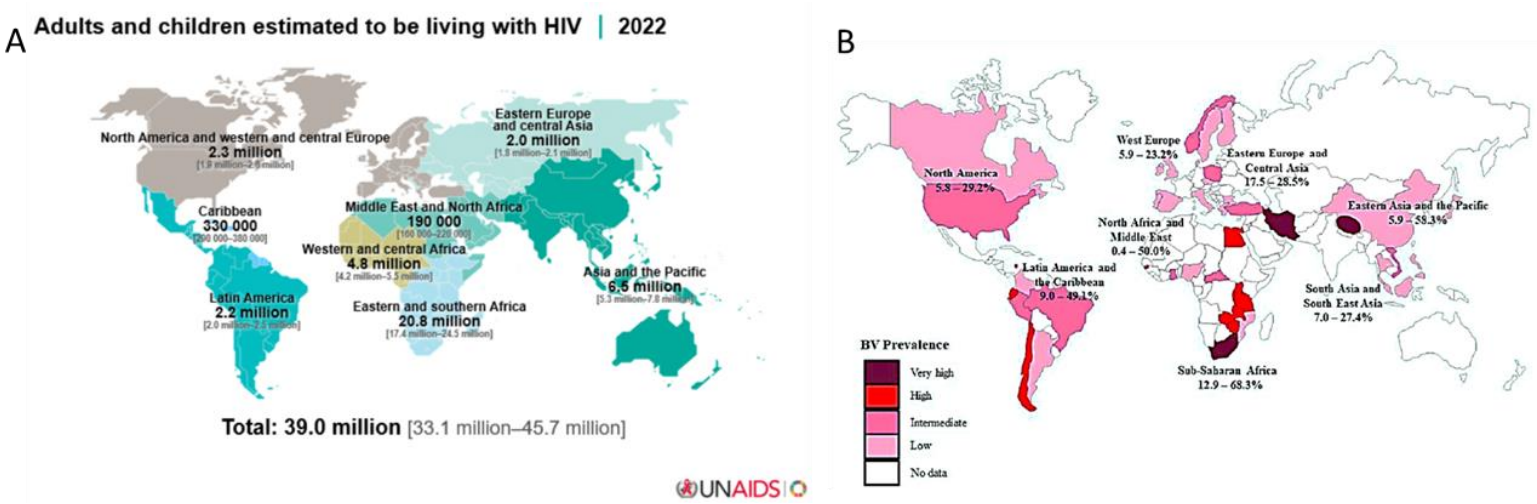


Figure 1.1: Global distribution of HIV and BV. A) Indicates the percentage of persons aged 15-46 infected with HIV in 2019 [4] B) Prevalence of BV in 2013 [9]

1.3 HIV structure

HIV is an RNA virus with a genome encoding 15 proteins, including the matrix, capsid, nucleocapsid, protease, reverse transcriptase, and envelope (Env) [21]. HIV replication begins with the attachment of Env to its target cells, $CD4^+$ T cells. Upon fusion of the viral and target cell membranes, the viral RNA and viral core are released into the target cell. Reverse transcriptase then initiates the formation of double-stranded DNA from the viral RNA which is transported into the nucleus and integrated into the host genome. The long terminal repeat (LTR) is activated, resulting in the expression of the viral transcriptome and proteome. The capsid is assembled, enclosing the viral RNA, and the mature viral particle, with Env at its surface, buds off from the plasma membrane [22]. The mature viral particles are able to infect HIV permissive cells and undergo further rounds of replications [23].

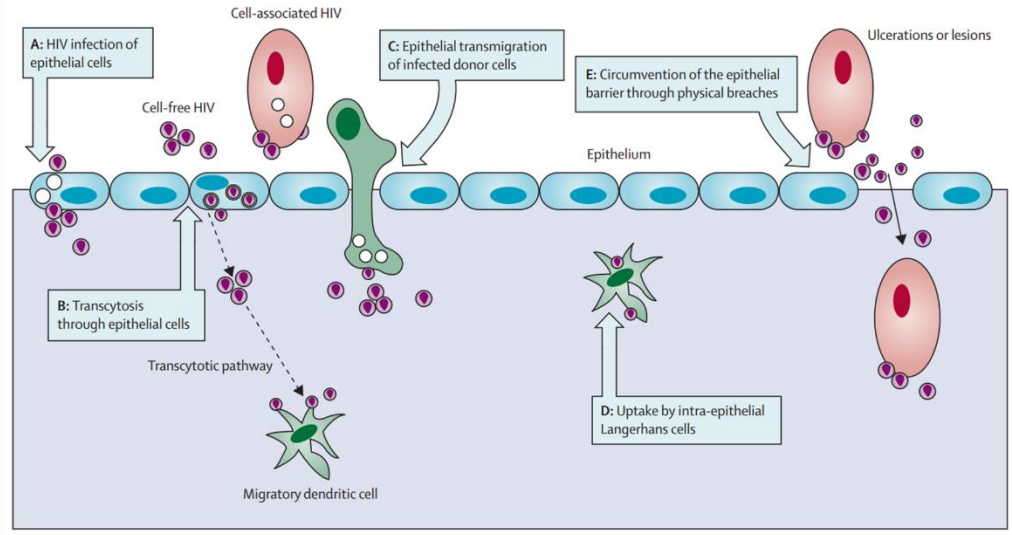
HIV infection is dependent on the interaction between Env and the CD4 and CCR5 (or CXCR4) receptors on the surface of T cells [24]. Env, which consists of the glycoproteins gp120 and gp41, binds to the CD4 receptor and undergoes a conformational change which enables binding to either the CCR5 or CXCR4 co-receptor. The shifts in structure facilitate the fusion of the viral and host cell membranes, mediated by gp41 [25]. Viruses differ in tropism as some use CCR5 to enter

cells whereas gp120 of other viruses binds to CXCR4 [26-28]. Most transmitted founders are R5 tropic, suggesting a selective advantage during transmission [29]. Gp120 is highly glycosylated comprising of high mannose and complex N-glycans [30]. Complex N-glycans have terminal sialic acids with either α-2,3, α-2,6 or α-2,8-linked bonds. The identity of the bond seems to be cell dependent as human $CD4^+$ T cells produced virions with Env carrying N-glycans with terminal sialic acid moieties linked via α-2,6-bonds, whereas pseudovirus produced in HEK293T cells had sialylated N-glycans that were only α-2,3-linked [31].

1.4 Mechanisms of HIV infection

Both cell-free and cell-associated virus have been identified in genital secretions [32] and the latter was suggested to be more efficient at HIV transmission [33]. Although HIV infection of $CD4^+$ T cells is dependent on the attachment of Env to CD4 and CCR5 (or CXCR4) [29], alternative mechanisms of entry include endocytosis of epithelial cells, macropinocytosis of endothelial cells, transcytosis through the epithelium layer, trans-infection of dendritic and Langerhans cells, migration via abrasions of the epithelium barrier (Figure 1.2) [34] and syncytium formation [35]. Although it remains controversial whether endocytosis results in productive HIV infection [36,37], it has been suggested that Env binding to the CD4 receptor and co-receptor could lead to endocytic internalisation of HIV, endosomal membrane fusion and release of HIV genome into the cytoplasm. Alternatively, a receptor-independent method could lead to recycling or degradation of HIV. Recycling of the virus could provide a reservoir of infectious HIV, able to infect $CD4^+$ T cells [38]. These pathways are most likely not mutually exclusive but preference of one over the other could be determined by a number of factors, including the FRT microbiome.

A



B

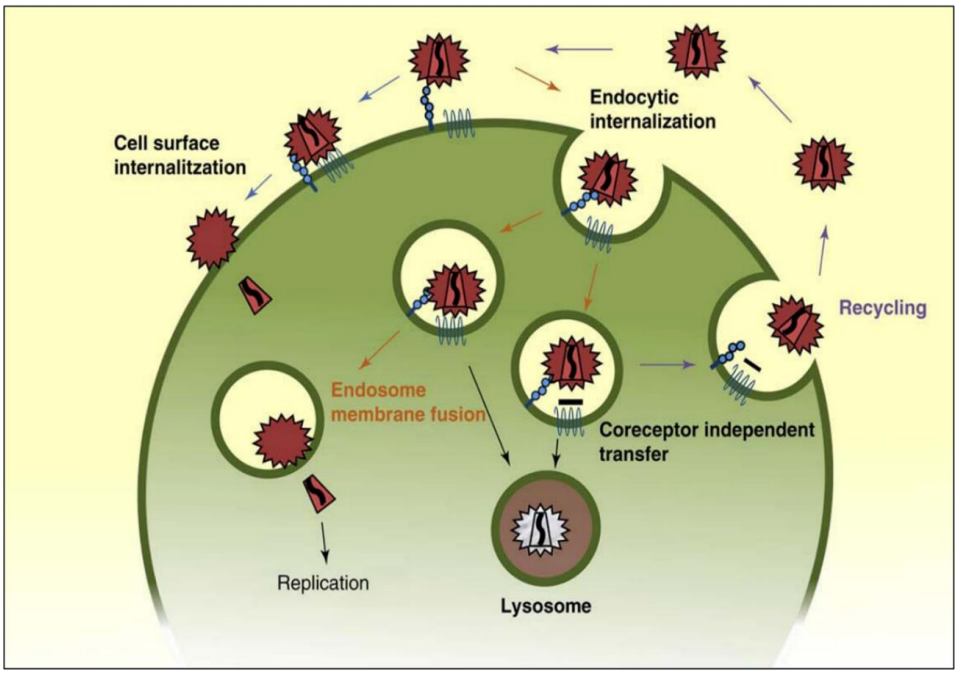


Figure 1.2: Mechanisms of HIV infection. A) Both cell-free and cell-associated HIV can cross a disrupted epithelial barrier by transmigration. Once in the sub-mucosa, infection of Langerhans cells can occur and result in transinfection of $CD4^+$ T cells [5]. B) Viral entry into $CD4^+$ T cells can occur via the CD4 receptor and CCR5 or CXCR4 co-receptor. Fusion of viral and host cell membranes can result in HIV infection or maturation of endosomes can result in viral degradation and when the necessary co-receptor is not available, the virus may be recycled and released to go on and cause productive infection of target cells [9].

1.5 Barriers to HIV infection

The FRT provides protection against HIV infection, reducing the chance of transmission per coital act [39]. Importantly, the chance of HIV infection is 14 times higher when men have sex with men (MSM) than during male to female transmission[39], emphasising the protective nature of the FRT mucosal barrier.

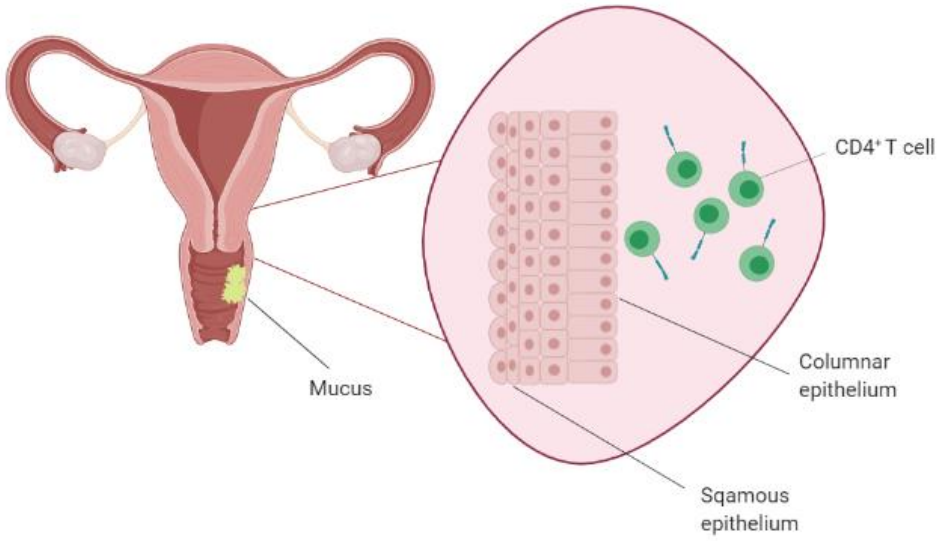


Figure 1.3: Diagram depicting the structure of the female genital tract. Barriers to infection by pathogens include cervicovaginal mucus (CVM) and epithelium which protect the underlying CD4+ T cells. (Adapted from [2]).

The first barrier to HIV transmission is the vaginal epithelium of the endocervix and ectocervix. The endocervix comprises of columnar epithelial cells, while the ectocervix comprises of squamous epithelial cells (Figure 1.3.) [39]. Another layer, called the stratum corneum, is made up of flattened differentiated epithelial cells. The endocervix is responsible for the production of cervico-vaginal mucus (CVM) which comprises proteins, such as mucins, immune cells, optimal microbiota, such as some *Lactobacillus* spp. and antimicrobial agents, which provide anti-inflammatory protection during epithelial damage [39]. Optimal FRT microflora comprises of a number of different microorganisms, each with their own innate adaptative abilities and high

spatial and temporal diversity [40]. The complexity of the FRT environment has made efforts to characterise the microbiome very challenging but it is known that an optimal microbiome prevents yeast infections, urinary tract infections (UTIs) and sexually transmitted infections (STIs), including HIV [41] (Figure 1.4). Although the precise mechanisms by which the microbiome wields its protective abilities is not yet fully understood, it is widely hypothesised that certain *Lactobacillus* spp. play a vital role in protection [40].

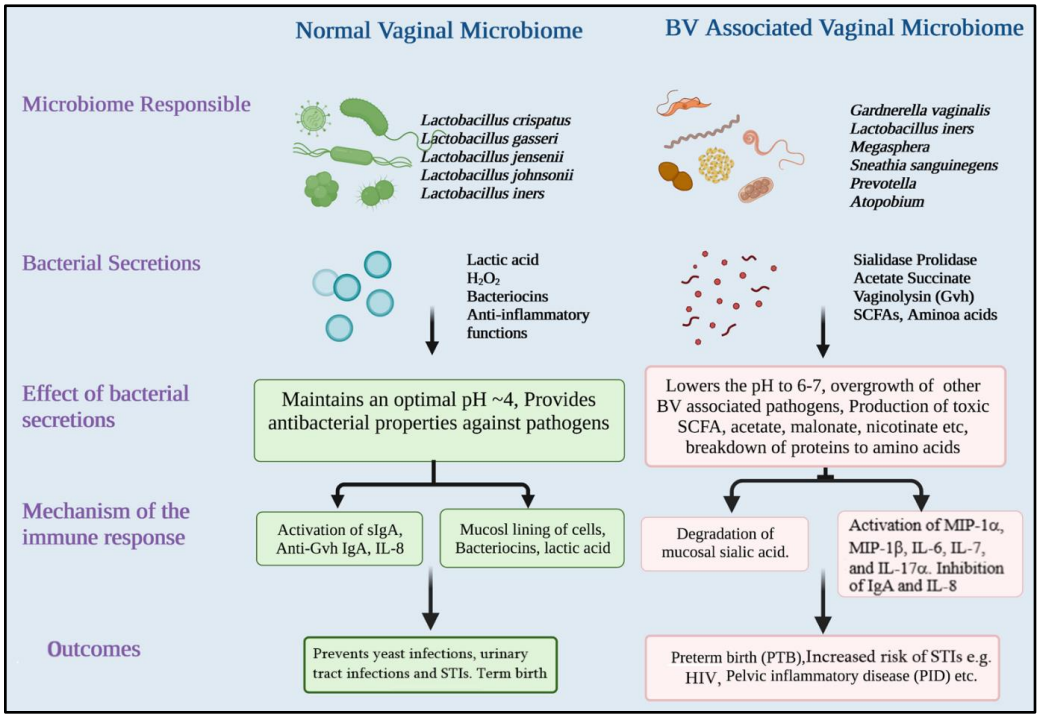


Figure 1.4. Impact of the microbiome on the environment of the female genital tract. Dysbiosis is the displacement of optimal microbiota with bacterial vaginosis associated bacteria (BVAB) such as *G. vaginalis*. Colonisation of the FRT by BVAB disrupts the protective mechanisms of the FRT, resulting in adverse effects. *Lactobacillus* spp. plays a protective role in the FRT, while *Gardnerella* does not. Adapted from Durand*, et al.* [6]. Non-optimal microbiome refers to bacteria associated with adverse health outcomes such as BVAB whereas optimal microbiome is characterised by microbiota associated with a 'healthy' FRT such as *L. crispatus* [7].

1.5.1 *Lactobacillus* spp.

Lactobacillus spp. are known producers of lactic acid and hydrogen peroxide (H_2O_2) which function to lower the pH in the FRT [40], creating a hostile environment for opportunistic pathogens (Figure 1.4). Although, *Lactobacillus* is generally considered to be associated with an optimal FRT microbiome, comparison between the global north, where most studies have been conducted, and sub-Saharan Africa have shown species variation [42]. Studies have suggested that certain *Lactobacillus* spp. are more protective than others because they have evolved to adjust to different environmental conditions within the FRT [43]. Therefore, some species are more effective at modulating pH and H_2O_2 levels while others produce bacteriocins [44,45]. Furthermore, *Lactobacillus* spp. differentially elicit immune responses characterised by specific cytokines, suggesting that some species might be proinflammatory [46]. Verstraelen*, et al.* [47] suggested that *L. iners* increased the likelihood of dysbiosis while *L. crispatus* provided greater protection [47]. Furthermore, *L. crispatus* was the only species able to reduce cytotoxicity associated with *G. vaginalis.* [48]. However, Petrova*, et al.* [49] reported that *L. iners* was present in approximately 50 % of BV-negative women which suggested that the effect of *L. iners* on dysbiosis may be due to quantity and not mere presence [49].

The FRT of North American and European women without any adverse health outcomes are dominated by *L. crispatus* [50], whereas women of African descent, tend to have microbiota dominated by *L. iners* [51]. This suggests that perhaps the composition of an optimal microbiome could vary across populations.

1.6 Bacterial vaginosis

Vaginal dysbiosis is defined by the displacement of optimal microbiota by anaerobes that tend to create an environment favourable for the survival of pathogens and predisposition to BV [52,53]. There are two methods commonly used to diagnose clinical BV, namely, (Amsel-BV) and Nugent scores (Nugent-BV). Four factors make up the Amsel criteria: vaginal discharge, pH above 4.5,

fishy odour upon addition of potassium hydroxide and 20 % clue cells. Clue cells, squamous cells coated in *Gardnerella* spp., have been observed in vaginal biopsies, urine and bladders of BV-infected women (Figure 1.5) [54]. However, Amsel criteria are mainly based on symptoms but BV may also be asymptomatic. Nugent scores are based on the ratio of the number of *Lactobacilli* to anaerobes present in a Gram smear of vaginal samples: a score < 3 is considered BV negative, a score 4-6 indicates an intermediate stage and the range 7-10 is indicative of BV [55]. Pyrosequencing revealed that the FRT microbiome of most healthy women were completely dominated by *Lactobacillus* spp., whereas a microbiome was considered in an intermediate phase when *L. gasseri, L. crispatus, L. iners, Atopobium vaginae, Gardnerella vaginalis, Peptoniphilus* spp. and *Actinomyces neuii* were present. Molecular-BV was defined by the presence of *A. vaginae, Prevotella bivia*, *Mobiluncus curtisii*, *Bacteriodies ureolyticus*, *Varibaculum cambriense*, *Arococcus christensenii*, *Dialister* spp., *L. iners* and *G. vaginalis* [49], as well as confirmation of three out of four Amsel criteria or when the Nugent score was above 7.

BV is the most common vaginal condition affecting young women. It is associated with a host of health issues, such as premature birth [56], pelvic inflammatory disease [57] and postoperative infections [58] (Figure 1.4). The most troubling effect of BV is its association with increased susceptibility to infection by several pathogens. Among these are *Chlamydia trachomatis, Neisseria gonorrhoeae* [59,60], *Trichomonas vaginalis* [59] HSV-2 [61] and HIV [62,63]. The precise aetiology of BV is yet to be fully understood, likely owing to its polymicrobial nature. However, there are two aetiological theories: the single pathogen hypothesis and the polymicrobial hypothesis [64]. The single pathogen hypothesis suggests that BV is caused by the outgrowth of a single pathogenic species, while the polymicrobial hypothesis suggests that BV is caused by a number of anaerobic, pathogenic bacteria [64].

Many species have been associated with BV and the composition of the FRT microflora is highly diverse between individuals. Nonetheless, *Gardnerella* is the genus that is most consistently present, generally considered to be a hallmark of BV and is the most virulent compared to other BVAB [65-67]. Initially, phylogenetic analysis classified *G. vaginalis* as the only species of the genus *Gardnerella* in the family Bifidobacteriales [68]. However, more recently, Vaneechoutte*, et al.* [69]

suggested that the genus be divided into *G. vaginalis, G. leopoldii, G. piotii and G. swidsinskii* and nine genome species.

1.7 *Gardnerella* diversity

A recent review discussed evidence obtained through inoculation of humans and animals with *Gardnerella*-positive samples and suggested that the bacteria "can under some conditions, recreate some of the features and complications that have been associated with BV" [70]. However, the authors do caution that much is unknown about how *Gardnerella* could influence growth of other microbial populations in the FRT, especially noting that strains were isolated from the FRT of both BV-positive and BV-negative women [71,72]. They also suggested that the presence of *Gardnerella* in the absence of BV could be due to some isolates being capable of facilitating dysbiosis while others might be non-pathogenic. Attempts have been made to classify *Gardnerella* based on different criteria, using different methods [73-75] but these studies did not reach consensus (Table 1.1). An early study by Piot*, et al.* [75] divided 359 strains into eight biotypes [75] but subsequently, Ingianni*, et al.* [76] differentiated *Gardnerella* into either three or four genotypes, depending on the restriction enzyme used in amplified ribosomal DNA restriction analysis (ARDRA) [76]. Unfortunately, they were not able to find an association between a particular genotype and clinical presentation of BV [76]. Although Santiago*, et al.* [77] confirmed this finding, they were able to find an association between ARDRA genotype and production of sialidase [77]. However, the clinical relevance of these results was not clear as they were not able to establish a link between genotype and sialidase production with BV status [77].

Amplification of the gene for the 60 kDa universal chaperonin protein (cpn60) is commonly used to distinguish between sub-species and it clustered *G. vaginalis* into four distinct subgroups (A, B, C and D) [78,79], with Paramel Jayaprakash*, et al.* [78] suggesting that there was sufficient variability to group the subspecies into four different species. This was supported by a similar study by Schellenberg*, et al.* [80]. They were able to correlate each ARDRA genotype with three cpn60 subgroups – genotype 1 with subgroups A, B and D, genotype 2 with subgroups B, C and D while genotype 3 was not detected in their study [80]. To elucidate the clinical significance of these

subgroups/species, Schellenberg*, et al.* [80] determined whether there were any phenotypic differences between *Gardnerella* isolates from different continents [80]. They found that the *sialidase A* (*sldA*) gene was present in subgroups B, C and D and absent in subgroup A isolates [80]. However, the presence of *sldA* was not indicative of active enzyme because only subgroup B consistently had sialidase activity [80]. These findings were supported by a similar study [81] which suggested that genotype 1 isolates produced little to no sialidase activity while all genotype 2 isolates were positive for sialidase activity [81]. On the contrary, there was no correlation between genotype/biotype and VLY [81].

Finally, Vaneechoutte*, et al.* [69] used whole genome sequencing to divide the *Gardnerella* genus into four novel species, namely *G. vaginalis, G. leopoldii, G. piotii and G. swidsinskii* [69]. They were able to determine that all *G. piotii* strains had sialidase activity, compared to only one out of four *G. vaginalis* strains and, all *G. leopoldii* and *G. swidsinskii* strains were negative for sialidase activity [69]. Overall, *Gardnerella* species are highly diverse which could have profound clinical implications, emphasising the need to identify whether some species are more pathogenic than others [78,79]. Moreover, many women present with several different *Gardnerella* isolates which potentially fall into different subgroups, genotypes, biotypes or even species and therefore likely differ in virulence [79]. The differential impact of species- or strain-specific virulence on HIV is of particular interest in Sub-Saharan Africa, with one of the highest prevalence of BV and HIV.

Time

Genotype [75,76]	Clade [82]	Subgroup [80]	Species [69]
2	C	1	*G. vaginalis*
1 and 2	D	3	Unnamed
1 and 2	B	2	*G. piotii*
1	A	4	*G. leopoldii* and *G. swidsinskii*

Table 1.1: *Gardnerella* classification over time. The green arrow indicates change in classification over time with *G. vaginalis* clades 1-4 corresponding with *cpn60* subgroups A-D: *cpn60* subgroup A corresponded to clade 4, subgroup B to clade 2, subgroup C to clade 1 and subgroup D to clade 3 [80]. Subsequently, the subgroups and clades were further classified into new species [69].

1.8 Factors that contribute to *Gardnerella* virulence

The ability to acquire iron is essential for the survival of nearly all organisms due to its role in enzymatic reactions [83]. *Gardnerella* expresses high-affinity iron transporters to take up iron which then facilitates the expression of iron-regulated proteins during infection that could play a role in its survival within the FRT [84]. Furthermore, it could outcompete *Lactobacilli* for an iron source, and thus influence the beneficial effect of low pH and bacteriocins in the FRT [85]. Adherence of pathogens to the vaginal epithelium is another factor that contributes to *Gardnerella* pathogenesis as it prevents removal by the movement of urine, mucus and vaginal secretions through the vaginal tract [66,86]. When compared to other BVAB, *Gardnerella* showed the greatest potential to adhere to mammalian cells *in vivo* [66]. Yeoman*, et al.* [86] showed that *Gardnerella* strains carried the genes encoding pili, which mediate epithelial adhesion. Therefore, *Gardnerella* spp. might be specially adapted to adhere to the epithelial cells of the FRT, facilitating the formation of biofilms, increasing the chances of survival of other BVAB and initiating dysbiosis [66,86].

1.8.1 Biofilms

A biofilm is a mode of growth whereby a community of bacteria adopts a sessile, adherent lifestyle (Figure 1.5) [87]. There are several advantages associated with biofilm formation in bacteria: better defence against changes in the environment, increased metabolic efficiency as a bacterial community and more effective colonisation of the FRT [87]. Lewis [88] showed that biofilm-associated bacteria were able to survive microbicides at much higher concentrations compared to plankton and were much less susceptible to phagocytosis, suggesting that biofilm-specific bacteria were more likely to persist in the FRT [88] irrespective of changes in pH, oxygen radicals and nutrient deprivation [87]. Certain bacteria have been shown to switch from biofilm to planktonic modes of growth (Figure 1.5) during nutrient deprivation to adapt to different ecological niches [87] and bacteria grown in either mode are able to sense their environment and regulate their metabolism accordingly. In biofilms regulation of gene expression may facilitate a genotypically and phenotypically heterogenous community [87] that uses a division of labour approach to combat changes in their environment and thus maximise their metabolic efficiency [87]. Treatment of BV is plagued by high recurrence rates most likely due to the versatility and persistent nature of the biofilm community.

The main species associated with biofilms in Amsel-BV were *Gardnerella* and *A. vaginae* [89,90] and it has also been proposed that the formation of biofilms by *G. vaginalis* promotes the colonisation of secondary anaerobes (Figure 1.5) [52,91]. *Gardnerella* are variable in size, measuring up to 3 µm in length and are gram-variable due to their relatively thin wall which influences Gram staining [54]: *Gardnerella* can appear Gram-positive during the exponential growth phase but can stain as Gram-negative over time because of progressive thinning of the cell wall [54]. However, after scrutinising the ultrastructure of the cell wall it was concluded that they are in fact gram-positive in their organisation [54]. Electron microscopy identified an exopolysaccharide layer characterised by strands emanating from the cell wall [54]. These strands have been suggested to play a role in biofilm formation as well as clustering of cells in broth cultures [54]. Additionally, more pili were found on the cell surface of clinical isolates compared to sub-cultured, laboratory strains, suggesting a physiological advantage within the FRT such as escaping contact with extracellular

enzymes and antibodies by binding tightly to epithelial cells [54]. *Gardnerella* strains isolated from BV-positive women had enhanced biofilm formation compared to those from BV-negative women, suggesting that some strains might be more virulent [92]. Of course, it is possible that *Gardnerella* strains that have evolved characteristics better adapted to the disruption of an optimal microbiome, such as forming biofilms to evade immune responses. Determining the association between BV and *Gardnerella* virulence factors might allow for better understanding of the causal relationship.

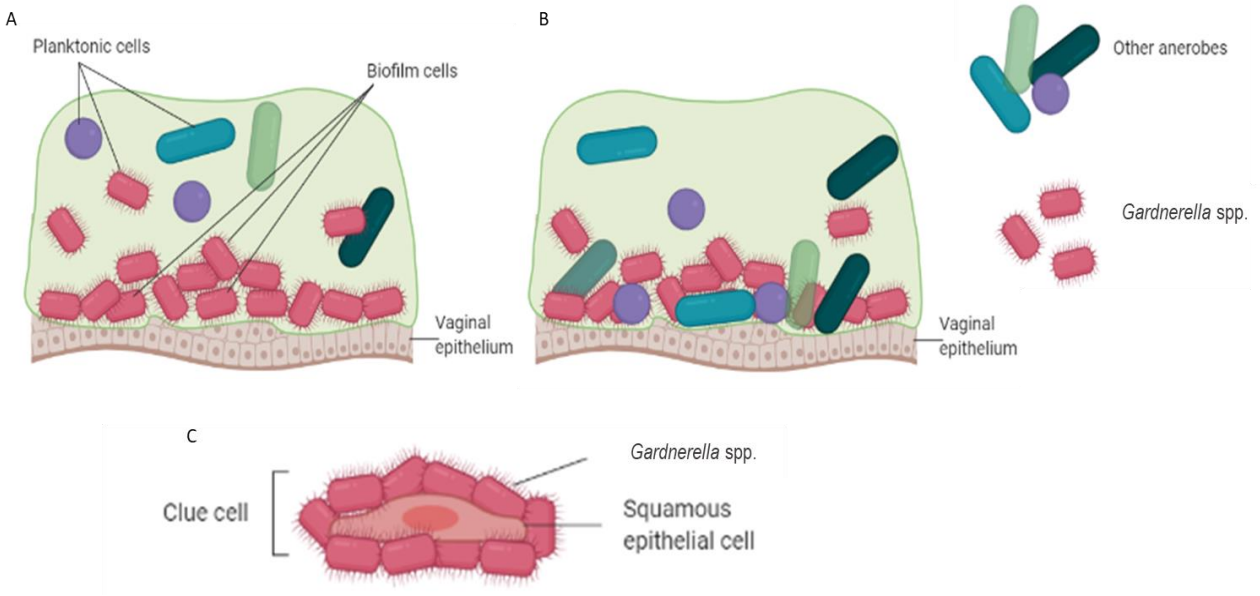


Figure 1.5: Biofilm and clue cell structure. Biofilm formation is initiated by a single colonising species, G. vaginalis (A). After a biofilm is established, other anaerobes are recruited and a polymicrobial biofilm is formed (B). Clue cells, used as a diagnostic feature of BV, are squamous epithelial cells coated with G. vaginalis (C). Adapted from Castro*, et al.* [3].

1.8.2 Sialidase

Sialidase activity has long been thought of as a diagnostic feature of Amsel-BV and Nugent-BV. Early studies demonstrated that elevated levels of sialidase activity in vaginal fluid coincided with the presence of sialidase-producing bacteria such as *Bacteriodes fragilis, Prevotella bivia* and *G. vaginalis* [93]. A more recent study by Aldunate*, et al.* [94] showed that in order to produce similar

levels of sialidase activity, *Prevotella bivia* was required in 100-fold excess compared to *Gardnerella* spp. This suggested that the presence of specific bacteria may not be as important as their expression of sialidase [95]. Furthermore, sialidase activity was associated with changes in FRT microbiota, suggesting that it could promote the outgrowth of more pathogenic strains linked to BV or conversely, loss and gain of species affected levels of sialidase [96]. Overall, sialidase activity may play an important role in the level of phenotypic diversity in the FRT microbiome.

Sialidase is an important virulence factor of *Gardnerella*. It targets the sialic acid residues found on glycoproteins, which are commonly found on mucosal surfaces, and thereby provides *Gardnerella* with a carbon source [1]. The activity of sialidases have been implicated in preterm birth and responsible for the characteristically thin vaginal discharge observed in women with BV [1]. The recent study by Robinson*, et al.* [1], demonstrated that only some strains of *Gardnerella* sialidase genes produce active sialidases. Until recently, *Gardnerella* sialidase activity was believed to be due to *sldA*, renamed NanH1. However, Robinson*, et al.* [1] found that NanH1 had little to no sialidase activity and reported that two homologues, NanH2 and NanH3, were instead responsible for the enzyme activity [1]. Although, sialidase activity has been detected in the cell free culture medium of G. *vaginalis*, the majority of activity was cell-associated [97,98]. When whole *G. vaginalis* cells were incubated with highly glycosylated secreted Immunoglobulin A (SIgA), free sialic acid was released into the extracellular environment and catabolised, [1] suggesting *G. vaginalis* sialidase was tethered to the cell surface, facing outwards [1]. Based on homology to sialidase from *Bifidobacterium*, Robinson*, et al.* [1] predicted the presence of C-terminal transmembrane α-helices which could tether NanH3 and NanH2 to the cell membrane. After attempting to purify the full-length protein, it was not detected in the soluble fraction and the authors suggested that the full-length protein formed aggregates via its hydrophobic α-helices. This was confirmed when the transmembrane domain was removed, and the protein was detected in the soluble fraction [1].

As for the activity found in culture medium, it is not yet clear whether this activity was due to active secretion of sialidase by the bacteria or whether it was merely a consequence of cell lysis or proteolytic cleavage [95]. Recently, it was shown that NanH3 could be translated in frame of a

signal peptide depending on the length of a stretch of cytosines [poly(C)] which could cause slipped strand mispairing during replication [10]. Therefore, when *Gardnerella* strains replicate, after a few generations, there would be a mixed population of isolates, some which secrete NanH3, and others that do not. It is possible that secretion, or the absence thereof, is essential for the physiological function of NanH3 in the FRT. This has ramifications on studies trying to understand differences in NanH3 between *Gardnerella* strains.

1.8.3 Vaginolysin

Gardnerella is highly cytotoxic and plays an important role in inducing apoptosis of vaginal epithelial cells [99]. The cytotoxic properties of *Gardnerella* were first observed in 1955 by Gardner and Dukes [100] and were attributed to a haemolysin characterised as a 59 kDa protein called Vaginolysin (VLY) [68]. Despite the fact that most strains identified thus far have been shown to produce VLY (Gelber et al., 2008, Soltani et al., 2007), a comprehensive study of 17 *Gardnerella* strains, conducted by Pleckaityte*, et al.* [81], indicated that expression of this toxin is highly heterogeneous [81]. VLY is a cholesterol-dependent cytolysin (CDC) which are a group of pore-forming toxins commonly utilised by Gram-positive bacteria and some Gram-negative bacteria [101].

VLY is secreted by *Gardnerella* during its exponential phase of growth and is mostly specific for human erythrocytes, neutrophils and endothelial cells [86]. The ability to form pores is dependent on the level of cholesterol in cell membranes and the complement regulatory molecule, CD59 [68]. When VLY resistant cells were transfected with human CD59 they became susceptible to VLY-induced lysis [68]. At high concentrations, VLY causes cell lysis and disruption of the mucosal barrier, while at sublytic concentrations, it initiates defence pathways to promote cell survival [101]. One such example is known as cell blebbing, whereby membrane disruption results in protrusions emanating from the cell membrane [101]. Randis*, et al.* [101] demonstrated that VLY, secreted by *G. vaginalis*, was able to induce rapid blebbing in epithelial cells [101], the first contact of pathogens in the FRT, and therefore may provide an advantage to invading pathogens such as HIV. VLY also activated the p38 mitogen-activated protein kinase pathway and induced pro-inflammatory IL-8

production by HeLa cells, a cervical epithelial cell line [68,102]. This suggests that VLY might not only be cytotoxic but could also regulate immune responses in the FRT by influencing gene expression of epithelial cells.

1.9 BV and HIV

Many studies have suggested that BV increases the chance of HIV transmission, [15,63,103,104] but have yet to reach a clear consensus on the mechanism responsible. Generally, studies on the association between STIs and susceptibility to HIV infection, suggest it is due to impaired mucosal or skin barriers, inflammatory secretions, or abnormal bleeding [104]. Increased levels of pro-inflammatory cytokines in the FRT were associated with Nugent-BV and Molecular-BV. They showed that highly diverse microbiota in the FRT had the highest association with inflammation followed by microbiomes dominated by *L. crispatus, G. vaginalis a*nd *L. iners* [105]. It has also been shown that cervicovaginal samples from BV-positive individuals activated the toll-like receptor-2 (TLR2) and induced HIV expression in infected cells *in vitro* [106]. Klebanoff and Coombs [107] were the first to suggest that BV may increase susceptibility to HIV infection due to the associated loss of *Lactobacillus* spp. A study by Gosmann*, et al.* [46] divided a cohort of South African women into four community types based on their FRT microbiome composition [46]. Their findings suggested that the group which had the least amount of "protective" *Lactobacillus* spp. and the highest diversity of anaerobes were the most susceptible to HIV infection [46].

Some *Lactobacilli* are capable of producing H_2O_2, which is toxic to many pathogens, including HIV [104]. However, this is a point of contention as it has also been shown that concentrations of H_2O_2 in the vaginal tract are not high enough to have a microbicidal effect [108]. The production of lactic acid by *Lactobacilli* maintains a low pH in the vaginal environment, which is unfavourable for the survival and growth of many microorganisms [104]. Low pH repressed $CD4^+$ T cell activation and inactivated free HIV-1 virus particles [15]. Thus, in the case of BV where pH is elevated, conditions would be more favourable for CD4 lymphocyte activation and therefore, infection by HIV [104]. It has also been shown that the FRT of women with depleted levels of *Lactobacillus* spp., had 17-fold higher concentrations of HIV target cells in the FRT, which led to higher risk of HIV acquisition

[46]. Furthermore, BV is associated with high IL-10 levels which increased susceptibility of macrophages to HIV infection [109]. It is likely that more than one of these factors indirectly increases HIV acquisition. However, it is also possible that *Gardnerella* virulence determinants directly affect the ability of the virus to establish clinical infection.

1.10 Role of *Gardnerella* virulence factors in HIV infection

Sialidase might have multiple effects on HIV infection. It is thought to contribute to the thinning of vaginal fluid by mucus degradation which increases the risk of HIV transmission by allowing a more intimate association between the virus and mucosal membranes [1]. In support of this, it was shown that when sialic acids were removed from CVM, the arrangement of mucin molecules was altered [110] [111]. Sialidases have also been implicated in inducing syncytium formation as well as enhancing the interaction of HIV with CD4 by removal of terminal sialic acids from gp120 [35,112]. Terminal sialic acids are attached to complex N-glycans via α-2,3 and α-2,6 linkages and are essential for Env function and the efficiency of HIV infection [113]. Therefore, de-sialylation of gp120 could enhance binding to CD4 and CCR5 and thus increase HIV infection [35,112].

Alternatively, it has been suggested that general loss of sialic acids from the glycoproteins on the viral and/or cell surface could reduce the negative charge repulsion, making it easier for the virus to adhere to the cells and bind to the CD4 receptor and/or co-receptors [112,114]. Whether removal of sialic acids specifically from Env and/or host cells or non-specifically from the membranes of both, could depend on the type of sialic acid linkages. Bacterial and viral sialidases vary in their specificity for sialic acid linkages with some cleaving only α-2,3 linkages [115], suggesting that sialidase substrate specificity could determine the mechanism by which it impacts HIV infection. On the contrary, HIV has been shown to utilise Siglecs as a method of infection of macrophages by interaction with sialic acids on the envelope [114]. Furthermore, N-glycosylation shields the virus from immune recognition [116]

As previously explained, VLY has been shown to cause membrane disruption of epithelial cells. This presents a possible advantage to HIV transmission in cells where blebbing occurs. Blebs are

transient membrane protrusions that occur when cell membranes separate from cytoskeletal structures and is associated with apoptosis, locomotion, and cell division. It is also linked to micropinocytosis [117], which is a form of endocytosis shown to enhance HIV infection of macrophages [118]. Furthermore, VLY activates p38 signalling pathways which has been shown to be essential for HIV replication [119]. Alternatively, VLY could enhance HIV infection by disrupting the mucosal barrier, allowing access to Langerhans cells that could mediate trans-infection of $CD4^+$ T cells [120]. There could thus be multiple ways that VLY might enhance HIV infection. Expression of VLY has been shown to be highly heterogenous across *G. vaginalis* strains [81] however, the clinical significance of this heterogeneity has yet to be determined.

1.11 Conclusion

BV is a highly complex condition that largely affects women of childbearing age from low- and middle-income countries, placing this vulnerable population at higher risk of acquiring HIV infection. Multiple mechanisms have been suggested by which BV enhances HIV infection and it is likely that more than one plays an important role. One of the challenges to understand the mechanism is the absence of a robust tissue and animal model and thus most studies are observational which has its own set of limitations. As far as we are aware, this is the first *in vitro* study to investigate the direct effect of *Gardnerella* virulence determinants on HIV infection.

1.12 Overall aim

Aim: To determine whether *Gardnerella* strains have varying impact on HIV infection because of differences in their ability to form biofilms, and variation in sialidase and VLY expression and activity.

Specific Objectives:

1. Determine whether biofilm production and vaginolysin expression and activity differ across strains.
2. Characterise the expression and activity of the sialidase homologues: NanH1, NanH2 and NanH3 present in *G. vaginalis* strains.
3. Cloning, expression, purification and enzymatic analysis of NanH3
4. Compare the effect of *Clostridium perfringens*, NanI, and NanH3 on whether HIV is able to infect the reporter cell line, TZM-bl cells.
5. Investigate how sialidase affects the mechanism HIV infection

Chapter 2: Methods and materials

2.1 Mammalian cells, *Gardnerella* strains and human erythrocytes

TZM-bl cells [Modified human cervical epithelial (HeLa)] were obtained through the NIH HIV Reagent Program [Division of AIDS, NIAID, NIH: TZM-bl Cells, ARP-8129, contributed by Dr John C. Kappes, Dr Xiaoyun Wu and Tranzyme Inc] and Human Embryonic Kidney cells (HEK293T) was a gift from Professor Carolyn Williamson, UCT. TZM-bl is a stable cell line that expresses the luciferase gene under the control of the HIV promoter, the long terminal repeat (LTR). Human erythrocytes were obtained as whole blood from the Western Cape Blood Service (WCBS). *Gardnerella vaginalis* ATCC 14018 was provided by Dr Lindi Masson (Burnet Institute) and 20 clinical *Gardnerella* isolates were obtained from Professor Joanne Passmore, UCT.

2.2 *Gardnerella* culture

Gardnerella strains were preserved in 60 % glycerol at – 80 °C. Glycerol stocks were recovered on New York City III [0.4 % HEPES, 1.5 % Proteose Peptone, 0.5 % NaCl, 0.5 % glucose, 0.25 % yeast extract, 10 % heat-inactivated horse serum (Celtic) 1.5 % agar (NYC III)] plates and incubated for 48 hours at 37 °C with anaerobic packs (Davies Diagnostics). Colony forming units (cfu) appeared after 48 hours and a single cfu was streaked on NYC III agar plates and cultured for a further 48 hours at 37 °C, anaerobically before inoculation of NYC III broth. NYC III broth 15 ml cultures were inoculated with single colonies and grown for approximately 48 hours at 37 °C, anaerobically.

2.3 Mammalian cell culture

HEK293T and TZM-bl cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM), supplemented with 10 % heat-inactivated Fetal Bovine Serum (FBS) and 1 U/ml penicillin and 1 µg/ml streptomycin, at 37 °C with 5 % CO_2 in a water-jacketed incubator. Human erythrocytes were isolated from whole blood as described by [121] and stored in phosphate buffered saline (PBS) at 4 °C for up to 50 days.

2.4 Production of infectious molecular clones

Infectious molecular clones (IMCs) were produced in HEK293T cells by transient transfection of the pNL4-3 HIV provirus cloned into the pUC18 vector [The following reagent was obtained through the NIH HIV Reagent Program, Division of AIDS, NIAID, NIH: Human Immunodeficiency Virus 1 (HIV-1), Strain NL4-3 Infectious Molecular Clone (pNL4-3), ARP-114, contributed by Dr M. Martin]. Cells were plated in 6-well plates at a density of 4 x 10^5 cells per well and allowed to adhere overnight at 37 °C and 5 % CO_2. The following day, polyethyleneimine (PEI) complexes were formed by vortexing (15 seconds) and incubating at room temperature (10 minutes) 400 µl serum-free DMEM with 1 µg of the pNL4-3 construct and 3 µg of PEI, before adding the complexes to the HEK293T cells in fresh 10 % FBS DMEM. The cells were incubated for 48 hours before the supernatant containing the pNL4-3 IMC particles were filtered through a 0.45 µM filter. The concentration of FBS was adjusted to 20 % and the IMCs were stored at – 80 °C until use.

2.5 Tissue culture infectious dose

Tissue culture infective dose that results in 50 % infected cells (TCID50) was determined by infection of the reporter cell line, TZM-bl. TZM-bl cells were plated at a density of 1 x 10^4 cells per well in a 96-well tissue culture plate and incubated at 37 °C, 5 % CO_2 overnight. The following day, the TZM-bls cells were infected with 100 µl undiluted IMC stock and ten 1:2 serial dilutions in quadruplicate in a final volume of 200 µl per well. Infected TZM-bl cells were incubated for 48 hours before 150 µl of the medium was removed from the cells and the remaining cells were lysed with 50 µl BrightGlo buffer (Promega) at room temperature for 2 minutes. After lysis, 75 µl of the lysate was transferred to an opaque 96-well plate and relative light units (RLUs) was measured in a GloMax 96 Microplate Luminometer (Promega). The TCID50 was calculated according to the Reed-Meunch method [122]. For all experiments, 200 TCID50 units was used to infect cells, unless stated otherwise.

2.6 Polymerase chain reaction

Polymerase chain reaction (PCR) using 16S universal primers (Tables 2.1 – 2.3) confirmed whether the strains belonged to the *Gardnerella* genus. PCR of *vly, nanh1, nanh2 and nanh3* genes using published primers were used to confirm the presence of the genes in the genome of all strains (Tables 2.1 – 2.3). Platinum SuperFi™ II DNA polymerase (ThermoFisher) was used for all amplification reactions according to the manufacturer's instructions and carried out using a LightCycler (Applied Biosystems).

Table 2.1: Components and final concentrations for PCR.

Component	Volume per reaction (µl)	Final concentration
5 x SuperFi™ II buffer	4	1 x
Forward primer	0.25	0.5 µM
Reverse primer	0.25	0.5 µM
10 mM dNTPs	0.4	200 µM each
Template DNA	-	5 – 100 ng gDNA
Platinum™ SuperFi™ II DNA polymerase	0.4	1 x
H_2O	Up to 20	-

Table 2.2: Cycling conditions for PCR.

Cycle	Temperature (℃)	Time	Cycles
Initial denaturation	98	30 seconds	1
Denaturation	98	10 seconds	30
Annealing	Table 2.3	10 seconds	
Extension	72	30 seconds per kb	
Final extension	72	5 minutes	1
Hold	4	∞	-

Table 2.3: Primer sequences, annealing temperatures, application, source of primer and expected product size.

Primer	Application	Sequence	Tm (°C)	Size (bp)	Source
VLY	Identification	F – 5' CTACGCCAGACAGCTTGAAG 3' R- 5' CGCCCTTAATTGCTGCTTCT 3'	55	155	[123]
NanH1 PCR	Identification	F – 5' CGAGGCAAAATGATTAACGAAGG 3' R – 5' TTGGCTCCTTTCAGTTCG 3'	54	636	[1]
NanH1 qPCR	Expression	F – 5' GATGTGTAGGCAAAGCCATTAG 3' R – 5' CTCGGCTGGCTGAATAT 3'	55	94	Designed
NanH2	Identification	F – 5' AGGAGTGCGTATGCGTAGAAG 3' R – 5' CCGCACTGCTGAGTTTCA 3'	66	348	[1]
NanH2 re-designed	Identification	F – 5' ATGCGTAGAAGTGTCATAGAAG 3' R – 5' AGGCTTTACAGAAATACCAGTC 3'	61	1788	Designed
NanH3 qPCR	Expression	F – 5' CAGTTCCAATGGAAGTGTGC 3' R – 5' AGCATCTGGGAATGCTCTTG 3'	55	322	[1]
16S universal	Identification	F – 5' AGAGTTTGATCMTGGCTCAG 3' (F27) R – 5' TACGGYTACCTTGTTACGACTT 3' (1492R)	55	2500	Gift from Dr Brian Cullin, UCT
16S qPCR	Expression	F – 5' TGAGTAATGCGTGACCAACC 3' R – 5' AGCCTAGGTGGGCCATTACC 3'	55	167	Designed
NanH1 qPCR	Expression	F – 5' GATGTGTAGGCAAAGCCATTAG 3' R – 5' CTCGGCTGGCTGAATAT 3'	55	93	Designed
NanH3 gap	Sequencing	F – 5' AATACAAATCGTAGCAAAAGG 3' R – 5' GCAGTAAGTTTATCTCAAAC 3'	48	1185	Designed
NanH3 truncated TM	Cloning	F – 5' GGATCCGGTGACTGTTTGGCCAAC 3' R – 5' ACTAGTTTAAGTTGTGTTGATTCC 3'	59	2147	Designed
M13_pUC_F	Sequencing	5' GTTTTCCCAGTCACGAC 3'	60	N/A	CAF
NanH3	Sequencing	F – 5' GTGACTGTTTGGCCAACT 3' R – 5' TTGCACACTTCCATTGGAAC 3'	60	910	Designed
pBAD R	Sequencing	5' GATTTAATCTGTATCAGG 3'	60	N/A	CAF
NanH3 middle 2 F	Sequencing	5' CCTACGAGTACTAAATATG 3'	60	N/A	Designed

2.7 RNA extractions

RNA extractions were carried out using Trizol (ThermoFisher). Briefly, *Gardnerella* cultures were grown for 48 hours under anaerobic conditions at 37 ℃, OD_{600} readings were measured with a spectrophotometer (ThermoFisher) and all cultures were diluted to an OD_{600} of 0.8 in a final volume of 12 ml NYCIII. Thereafter, cultures were centrifuged at 5 000 rpm for 10 minutes at 4 ℃ and the cell pellets were stored at – 80 ℃. For extractions, cells were thawed on ice, resuspended in 1 ml of PBS and centrifuged at 5 000 rpm for 10 minutes. Cells were resuspended in 1 ml of Trizol reagent and incubated at room temperature for 5 minutes before chloroform (200 µl) was added, mixed vigorously for 15 seconds and incubated for another 3 minutes at room temperature. The samples were then centrifuged at 4 ℃ for 15 minutes at 14 000 rpm. After centrifugation, three distinct layers were formed and the upper aqueous layer, containing total RNA, was removed. Thereafter, 500 µl of isopropanol was added and mixed vigorously. Samples were incubated at room temperature for 10 minutes before centrifugation at 14 000 rpm for 30 minutes at 4 ℃. The RNA was resuspended in 1 ml of 75 % diethyl pyrocarbonate (DEPC)-EtOH and gently vortexed for 10 seconds prior to centrifugation for 5 minutes at 8 000 rpm and at 4 ℃. The supernatant was removed, and the RNA pellets were allowed to air-dry to remove excess EtOH. The RNA was resuspended in 30 µl of DEPC-H_2O and incubated at 55 ℃ for 10 minutes. RNA concentration was determined using the Nanodrop 2000 (ThermoFisher) and its integrity was established by gel electrophoresis at 90 V using a 1 % agarose gel and then visualised by iBright FL1500 imaging system (Invitrogen). All RNA extractions were carried out of three independent bacterial cultures.

2.8 cDNA synthesis

The ImProm-II™ reverse transcription system (Promega) was used for cDNA synthesis. All samples were diluted to 90 ng in a volume of 4 µl to which 1 µl of Random primers (Promega) was added in PCR tubes (Lasec) (Table 2.4). This mixture was placed in the LightCycler (Applied Biosystems) for 5 minutes at 70 ℃ followed by 5 minutes at 4℃.

Table 2.4: cDNA synthesis master mix.

Component	Volume per reaction (µl)	Final concentration
ImProm-II™ 5 x buffer	4.0	1 x
$MgCl_2$	2.4	3.0 mM
dNTP mix	1.0	0.5 mM each dNTP
Recombinant RNasin®	0.5	20.0 units
ImProm-II™ reverse transcriptase	1.0	1.0 µl
H_2O	6.1	-
Total	15.0 µl	-

After incubation, 15 µl of a master mix (Table 2.4) was added to each sample. The tubes were returned to the Thermocycler at 25 °C for 5 minutes, 42 °C for 1 hour and 70 °C for 15 minutes.

2.9 Quantitative PCR

Quantitative PCR (qPCR) was carried out using the PowerUp SYBR Green Master Mix (Applied Biosystems) and the Quantstudio lightcycler (Applied Biosystems). Briefly, 90 ng cDNA was added to a mastermix (Table 2.5) and the reactions were carried out as shown (Tables 2.6 and 2.7). Gene expression was normalised using the delta Ct method (ΔCt), where ΔCt = Ct (16S rRNA)—Ct (target gene). A negative control lacking reverse transcriptase was included in each reaction. The expression of each gene was determined by three independent reactions, each carried out in triplicate. A melting curve (Table 2.7) determine the occurrence of non-specific amplification.

Table 2.5: qPCR master mix.

Component	Volume per reaction (µl)	Final concentration
PowerUp™ SYBR™ Master mix	10.0	-
Forward primer	0.8	From 10 µM
Reverse primer	0.8	From 10 µM
cDNA	2.0	From 90 ng
H_2O	6.4	-

Table 2.6: qPCR cycling conditions.

Step	Temperature (°C)	Time	Cycles
UDG activation	50	2 minutes	Hold
Hold Dual-Lock™	95	2 minutes	Hold
Denature	95	15 seconds	40 cycles
Anneal/extend	55	15 seconds	
	72	1 minute	1 cycle

Table 2.7: Melt curve cycling conditions.

Step	Temperature (°C)	Time (minutes)
1	95	15 seconds
2	60	1 minute
3	95	15 seconds

2.10 Preparation of *Gardnerella* abiotic culture medium and cell lysates

Gardnerella strains were cultured as described above (Section 2.2). The cultures (2 ml) were centrifuged at 5 000 rpm for 10 minutes at 4 °C and the supernatant was filter sterilised using a 0.45 µM filter and stored at – 20 °C. The bacterial pellets were resuspended in 300 µl of PBS and sonicated 5 x for 20 second intervals on ice. Initially, after sonication the samples were centrifuged at 5 000 rpm for 10 minutes at 4 °C and the supernatants representing the cell lysate fraction were collected and stored at – 20 °C. To enrich for sialidase activity, bacterial pellets were sonicated as before but were not centrifuged and the lysed cells were stored at – 20 °C. A bicinchoninic acid (BCA) assay (Pierce™) was performed on the cell lysates, according to the manufacturer's instructions and cell-free culture medium to determine the concentration of total protein. Preliminary experiments were used to determine protein concentrations that did not induce cytotoxicity (data not shown). All samples were diluted to 1 µg/µl total protein in preparation for the infection assay. All abiotic and cell lysate preparations were carried out from three independent bacterial cultures.

2.11 TZM-bl infection assay

TZM-bl cells are HeLa cells that express CD4 and CCR5 and also contain the firefly luciferase gene under the control of the HIV-1 LTR promoter [124]. Infection of TZM-bl cells is measured after activation of the LTR by *tat* from IMCs which results in luciferase expression [125]. TZM-bl cells were seeded at a density of 1 x 10^4 cells per well and allowed to adhere overnight at 37 °C, 5 % CO_2. Cells were infected with 200 TCID50 units and 10 µg total protein from either cell-free culture medium or cell lysates unless indicated otherwise. Cells were incubated at 37 °C and 5 % CO_2 for 48 hours before 50 µl of BrightGlo (Promega) was used to lyse the TZM-bl cells and measure the extent of infection by RLUs using a GloMax 96 Microplate Luminometer (Promega) as described in section 2.5.

2.12 Sialidase infection assay

The substrate 4-Methylumbelliferyl-N-acetyl-α-D-Neuraminic Acid [(MUN) Biocom] was used to determine sialidase activity in the samples. Preliminary experiments determined the presence of activity in samples as well as inhibition by the inhibitor (see section 2.18). To determine the effect of sialidase on HIV infection, TZM-bl cells were infected with pNL4 IMCs after treatment with sialidase. Cells were seeded at a density of 1 x 10^4 and allowed to adhere overnight at 37 °C and 5 % CO_2. Cells were then treated with 10 U of sialidase enzyme (NEB) purified from *Clostridium perfringens* or purified *G. piotii* NanH3 for 8 hours. Inhibition by 2,3-dehydro-2-deoxy-N-acetylneuraminic acid (DANA) (30 mg/ml) was tested by incubating the inhibitor with the sialidase for 2 hours before adding to the cells or virus for 8 hours. IMCs were also treated with sialidase [10 U per 200 TCID50 units (50 µl) of virus] for 8 hours before adding the inhibitor for 2 hours. Thereafter, untreated and treated TZM-bl cells were infected with untreated and treated virus in different combinations to determine whether the effect of sialidase on host cells and viruses affected HIV infection. The cells were then incubated for 48 hours before lysing with BrightGlo to measure infection as outlined in section 2.5.

2.12.1 Effect of sonicated *Gardnerella* cells on HIV infection

The infection assay was prepared according to section 2.11. *Gardnerella* cells were pelleted, resuspended in 100 µl PBS and sonicated as described in section 2.10. Total protein was measured by BCA assay and sonicated cells, equivalent to 10 µg of total protein, was added to the infection assay. For samples inhibited by DANA, sialidase was treated with DANA (30 mg/ml) for 2 hours prior to addition to TZM-bl cells. Thereafter, 200 TCID50 units of pNL4 IMCs were added and infection was measured after 48 hours. Sonicated *Gardnerella* cells were prepared from three independent bacterial cultures.

2.13 Visualisation of Viral Like Particles

Viral Like Particles (VLPs) were generated as described in Section 2.4 with the following changes: HEK293T cells were transiently co-transfected with 2.5 µg pCDNA3.1_gp160 E8 [126] and 5 µg eGFP-gag [The following reagent was obtained through the NIH HIV Reagent Program, Division of AIDS, NIAID, NIH: Plasmid pTM1 Expressing Gag-Green Fluorescent Protein (GFP) Fusion, ARP-4810, contributed by Dr. Paul Spearman] using 22.5 µg PEI. The culture medium was harvested, and filter sterilised and total protein was determined by BCA assay. VLPs equivalent to 1000 µg was used to infect the TZM-bls. TZM-bl cells were plated at a density of 5000 cells per well in order to acquire single cell images and incubated overnight. Cells and VLPs were treated with 20 U of *C. perfringens* NanI or *G. piotii* NanH3 for 8 hours before inhibitor was added and incubated for a further 2 hours. The positions of single cells were mapped using the Zeiss LSM 980 confocal microscope, and VLPs were added before images were taken every minute for 300 minutes. Zen 3.5 was used for image analysis. Mean fluorescent intensity (MFI) was calculated for 3 images per condition to find the average.

2.14 Biofilm quantitation assay

Gardnerella cultures, including ATCC 14018 as a positive control [127], were diluted to an OD_{600} of 0.2 and 200 µl was plated per well in a 96-well tissue culture plate. The plate was incubated at 37

°C under anaerobic conditions for 48 hours before the culture medium containing the non-adherent plankton was removed. The wells were washed with 200 µl PBS and fixed with 100 µl methanol at room temperature for 15 minutes. Once the methanol was removed, and the cells were washed twice with 200 µl PBS, 100 µl of 0.5 % crystal violet was added and incubated at room temperature for 20 minutes. The cells were washed a further two times with PBS to remove the unbound crystal violet. Cells were solubilised with 150 µl of 33 % acetic acid before 100 µl was measured spectrophotometrically at 550 nm. Biofilms were quantified relative to the ATCC 14018 strain. Each biological repeat was prepared from independent bacterial cultures.

2.15 Haemolysis assay

To assess the potential of *Gardnerella* cell-free culture medium to lyse human erythrocytes, a haemolysis assay was performed according to the protocol described by Evans*, et al.* [121]. Briefly, whole blood was centrifuged, and the plasma layer was removed leaving behind the erythrocytes. To test haemolysis, the equivalent of 50 µg total protein of each sample was added in a final volume of 10 µl to 190 µl of erythrocytes in PBS at a pH of 7.4 in a 96-well plate. The cells were incubated at 37 °C for 1 hour before centrifuging to remove the intact erythrocytes and the absorbance of the supernatant (100 µl) was measured in a spectrophotometer at 450 nm. Preparation of cell-free culture medium was conducted on three independent bacterial cultures.

2.15.1 Haemolysis inhibition assay

Cell-free culture medium was obtained as described in section 2.10 and the haemolytic assay was performed as described in section 2.15. However, 50 µg of each sample was incubated in the absence and presence of 0.4 mg/ml retrocyclin for 1 hour 30 minutes at 37 °C before the haemolysis assay.

2.15.2 Biofilm haemolysis

Bacteria were grown as biofilms as described in section 2.14. After 48 hours, supernatants collected, and filter sterilised to obtain cell-free culture medium. The concentration of total

protein was determined by BCA assay before the cell-free culture medium was used in the haemolysis assay as described in section 2.15.

2.18 Sialidase quantitation assay

The sialidase assay was adapted from the method described by Moncla*, et al.* [128]. Briefly, the sialidase substrate MUN, was prepared at a concentration of 0.5 mg/ml in H_2O. Prior to use, 1 M sodium acetate was added at 20 µl per 180 µl of MUN. For each reaction, 25 µl of MUN with sodium acetate and 25 µl of the sample was used. The reactions were added to a black 96-well plate and fluorescence was measured at emission 415-445 and excitation 365 nm for 1 hour at 37 ℃ using a GloMax 96 Microplate Luminometer (Promega). A standard curve was generated using the *C. perfingens* sialidase (0.9 – 0.01 U/µl) to determine the enzyme concentration of recombinant *Gardnerella* NanH3.

2.19 Cloning

2.19.1 PCR

Primers flanking the *nanh3* gene were designed with the reverse primer annealing to the template immediately before the start of the transmembrane domain. The forward and reverse primers included 5' extensions for the recognition sites of restriction enzymes (RE), *BamH*I and *Spe*I. Cloning of *nanH3* into the vector's multiple cloning site ensured that it was in frame with the N-terminal His-tag. DNA was extracted from GP 74 culture using the gDNA extraction kit (Zymo) and 50 ng of DNA was used in a PCR as outlined in section 2.6. PCR samples were pooled and non-polymerised nucleotides were removed using a DNA clean up kit (Qiagen). The PCR product and the pProEx-HTa vector were digested with *Bam*HI and *Spe*I (ThermoFisher) at 37 °C for 1 hour and the appropriate bands were excised and purified (Qiagen) from a 0.8 % agarose gel run at 80 V for 1 hour and 30 minutes.

2.19.2 Ligation

RE digested fragments were ligated in a 3:1 (insert: vector) ratio using T4 ligase according to the manufacturer's instructions (ThermoFisher). JM109 cells (Promega) were transformed with the ligation reaction and a single cfu was amplified using the NanH3 primers and cycling conditions explained in section 2.6. PCR products were separated on a 0.8 % agarose gel run at 80 V for 1 hour and 30 minutes and the 2000 bp band identified the cfu that were positive for the insert. Two positive and one negative cfu were selected for plasmid extraction (Qiagen) and digestion with *Bam*HI and *Spe*I. DNA was digested at 37 °C for 1 hour and ran on a 0.8 % agarose gel for 1 hour 30 minutes.

2.19.3 Protein expression and column purification

Luria broth (LB) with carbenicillin (100 mg/ml) was inoculated with a single cfu and grown for approximately 16 hours at 37 °C, with shaking. The plasmid DNA was extracted (Qiagen) and used to transform BL21 cells (NEB) for protein expression. Cultures were incubated at 37 °C for approximately three hours, with shaking before gene expression was induced by 0.2 mM IPTG (ThermoFisher) and incubated for a further 4 hours. Thereafter, the cultures were centrifuged at 7 000 rpm for 10 minutes and the wet weight of the cell pellets were determined so that the appropriate amount of BugBuster [5 ml per g (Merck)], Benzonase [1 µl per ml of BugBuster (Merck)] and PMSF (1 mM final concentration) were added to in the cells. Once resuspended, the solution was allowed to gently shake at room temperature for 20 minutes before centrifugation for 20 minutes at 16 000 g at 4 °C. The cell lysate was applied to a 10 ml column (Bio-Rad) with 2 ml Ni-NTA agarose, after equilibration with 2 x 4 ml equilibration buffer (ThermoFisher) and the flow through was collected. The column was washed twice with 4 ml of wash buffer 1 (20 mM imidazole) followed by twice with 1 ml wash buffer 2 (30 mM imidazole) and finally washed twice with 1 ml buffer 3 (50 mM imidazole). Proteins were eluted 5 times with 500 µl buffer containing 250 M imidazole. All wash and eluted fractions were collected and measured for total protein BCA assay and sialidase activity as described in Section 2.18.

2.19.4 SDS-PAGE, Western blot and Coomassie

Fractions were prepared for SDS-PAGE by adding loading dye (1% SDS, 4% glycerol, 1% β-mercaptoethanol and 0.01% bromophenol blue) and heating at 100 °C for 10 minutes prior to loading on the SDS-PAGE. The SDS-PAGE was prepared by creating a 5 % stacking gel and an 8 % resolving gel of 0.75 mm thickness. The samples were loaded onto duplicate gels and then run at 60 mA for approximately 3 hours. One gel was stained with Coomassie dye by shaking at room temperature for 1 hour, while the other gel was transferred at 100 V for 110 minutes onto a polyvinylidene difluoride (PVDF) membrane for Western blotting. After Coomassie stain, the gel was incubated at room temperature overnight in destaining buffer (Appendix I) and imaged the following day. For Western blotting, the membrane was blocked with 5 % skim milk blocking buffer for 1 hour at room temperature before exposure to a His-tag antibody (Cell Signalling Technology) followed by washing 3 times in 5-minute intervals with 1 x TBS-T. A horse-radish peroxidase conjugated anti-rabbit secondary antibody (Santa Cruz) was added and incubated for 1 hour before extensive washing with 1 x TBS. LumiGLO (WhiteSci) peroxidase chemiluminescent substrate was used to detect the bands using the iBright FL1500 imaging system (Invitrogen).

2.20 Mass spectrometry analysis

Mass spectrometry analysis (Sections 2.20.2 – 2.20.4) was carried out at Stellenbosch University's Biomedical Research Institute.

2.20.1 Preparation of gel slices

Bands of interest were excised and segmented into 1 cm x 1 cm sections and placed in 100 % acetonitrile until the slices became opaque. Thereafter, the samples were placed in a SpeedVac for 30 minutes to prepare for transportation.

2.20.2 In-gel digest

Samples were dehydrated and desiccated before reduction with 2 mM triscarboxyethyl phosphine (TCEP) in 25 mM NH_4HCO_3. Cysteine residues were thiomethylated with 20 mM S-Methyl methanethiosulfonate in 25 mM NH_4HCO_3 and the gel fragments were dehydrated and washed with 25 mM NH_4HCO_3 then dehydrated again. Proteins were digested with 20 ng/µl trypsin at 37 °C overnight. Water was used to extract peptides and resuspended in 30 µl 2 % acetronitrile, 0.1 % formic acid.

The residual reagents were removed by using a C_{18} stage tip. The bound samples were washed with 30 µl 50 % acetonitrile, 0.1 % formic acid. The eluate was evaporated, and the dried peptides were dissolved using 2 % acetonitrile, 0.1 % formic acid for subsequent liquid chromatography mass spectrophotometry analysis.

2.20.3 Liquid chromatography

Liquid chromatography was performed using the Ultimate 3000 RSLC (Thermo Scientific), with a 5 mm x 300 µm C_{18} trap column and a CSH 25 cm x 75 µm 1.7 µm particle size C_{18} column. The solvent system used was solvent A: 2 % acetonitrile, 0.1 % formic acid and solvent B: 100 % acetonitrile. Chromatography was performed at 40 °C and delivery to the mass spectrometer was via a stainless steel, nano-bore emitter.

2.20.4 Mass spectrometry

Mass spectrometry was carried out using the Fusion mass spectrometer (Thermo Scientific) with a Nanospray Flex ionization. Data was collected with spray voltage at 1.8 kV and ion transfer capillary at 290 °C. MS1 scans were done using the orbitrap detector at 120000 over a range of 375-1500 with AGC target at 5 E4. MS2 scans were done using monoisotopic precursor selection for charges +2 - +7. Fragment ions were detected in the orbitrap mass analyser at 30000 resolution and data was acquired in centroid mode.

2.21 Enzyme kinetics

The sialidase quantitation assay (Section 2.18) was used to determine the enzyme kinetics of both *C. perfringens* and *G. piotii* sialidases. This required various concentrations of substrate and constant enzyme concentration. MUN concentrations ranged from 0 – 1 µmol/ml and optimal enzyme concentrations were determined experimentally. Therefore, 10 U of *C. perfingens* NanI and 0.25 U of *G. piotii* NanH3 was added to the reaction mix at each substrate concentration and fluorescence was measured as described in Section 2.18. Initial velocity was determined by measuring the slope of the curve and plotted against substrate concentration. Michaelis-Menten kinetics and Lineweaver-Burke analysis identified the Km of the sialidase.

2.22 Statisical analysis

A one-way ANOVA with multiple comparison Dunnett's post-test were performed on the average of two or three independent experiments performed in triplicate as indicated. Association between variables such as expression, activity and impact on HIV infection were compared using linear regression. All statistical tests performed in this study were done in GraphPad Prism (Version 8, GraphPad Software, La Jolla, California, United States) and Microsoft® Excel®. Data were deemed significant if p-values were less than 0.05, 0.01, 0.001 and 0.0001, indicated as *, **, *** and ****, respectively.

Chapter 3: Vaginolysin

3.1 Introduction

The consistent isolation of *Gardnerella* from the FRT of BV-positive women suggests that it could be responsible for the displacement of optimal *Lactobacilli* spp. such as *L. crispatus*. However, the presence of *Gardnerella* in vaginal samples from BV-negative women suggests that it is unlikely to be the only aetiological agent of BV. Alternatively, *Gardnerella* genome species may encompass both pathogenic and non-pathogenic strains, with only the former linked to a non-optimal FRT microbiome and adverse health outcomes. Recently, it has been shown that *Gardnerella* comprises *G. leopoldii*, *G. piotii*, *G. swidsinskii*, *G. vaginalis* and nine genome species [69]. The relative abundance of *G. vaginalis*, *G. swidsinskii*, and *G. piotii* was associated with Amsel/Nugent-BV [129] potentially due to differences in virulence [130].

The association between BV and HIV acquisition suggests that pathogenic strains of *Gardnerella* could be a factor predisposing individuals to HIV infection and that virulence determinants could play a direct role in HIV transmission. The virulence potential of *Gardnerella* spp. is largely defined by the formation of biofilms, and the activity of VLY and sialidase. VLY is a species-specific cytolysin that lyses human cells at high concentrations and causes membrane blebbing at low levels [68]. VLY has recently been divided into five groups based on differences in their undecapeptide sequences with Type 1A associated with increased toxicity. This suggests that some strains could be more cytotoxic than others depending on VLY genotype [131]. Alternatively, the cytotoxicity of VLY in the FRT could be due to the relative abundance of VLY-positive strains and/or whether some strains express *vly* better than others. It was shown that expression levels were proportional to increased cytotoxicity *in vitro* [81] depending on whether strains were isolated from BV-positive or BV-negative women [48,132].

At low concentrations, VLY might increase HIV infection by causing membrane restructure. Membrane blebbing has been linked to cell healing and repair in response to mechanical damage such as pore formation [133] by inducing exocytosis and endocytosis [134]. Although controversial,

endocytosis might enhance HIV infection [38], suggesting that low expression of VLY might directly favour transmission. Therefore, VLY might have conflicting influences on HIV transmission based on its concentration levels in the FRT because endocytosis could enhance viral entry [135] and cell lysis could either decrease infection by lysing infected epithelial cells and/or support immune evasion by lysing neutrophils [136].

Studies have revealed that *Gardnerella* spp. differentially express virulence genes depending on planktonic or biofilm growth [86,123]. For example, it was demonstrated that *vly* expression was downregulated when *Gardnerella* spp. were cultured as biofilms [123]. The importance of biofilms is not only restricted to changes in expression of virulence factors but also facilitates adherence of bacteria to the FRT epithelium, antibiotic resistance, metabolic efficiency and protection from host defence mechanisms [87]. Therefore, biofilm formation may be considered a crucial part of *Gardnerella* virulence and strains that differ in their ability to form biofilms might vary in virulence potential.

This study determined whether strains varied in biofilm production, and whether this was linked to expression and haemolytic activity of VLY. Finally, we investigated whether *vly* expression levels and haemolytic activity was associated with HIV infection *in vitro*.

3.2 Results

3.2.1 Genotypes

Twenty-one strains were isolated from five BV-positive (UC058, UC068, UC074, UC079 and UC094) and one BV-intermediate patient (UC092) from which 1, 15, 1, 1, 2 and 1 strains were identified, respectively (Table 3.1). Sequencing of *cpn60* confirmed that 16 strains belonged to *G. vaginalis*, two were *G. piotii* and two were *G. swidsinskii* (work done by Hannah Livingstone, Jo-anne Passmore). PCR with *Gardnerella*-specific primers was routinely conducted on single colony-forming units (cfus) prior to inoculating broths to ensure bacterial cultures were free of contamination (data not shown). PCR of the genomic DNA from one strain (GV 5), was not positive

for the 250 bp band, suggesting it belonged to a different genus and was therefore used as a *Gardnerella*-negative control in subsequent experiments.

Table 3.1: Strain identities and genotypes

Strain ID number	[a]Strain name	BV status	Species	Number of strains	Strains positive for *vly* (%)	strains positive for *nanh1* (%)	Strains positive for *nanh3* (%)
ATCC 14018	[b]ATCC	Type strain	*G. vaginalis*	16	100 %	100 %	0 %
UC068 2-1	GV 1	BV-positive					
UC068 2-2	GV 2						
UC068 2-4	GV 4						
UC068 2-5	GV 5						
UC068 2-7	GV 7						
UC068 2-10	GV 10						
UC068 2-11	GV 11						
UC068 2-12	GV 12						
UC068 2-13	GV 13						
UC068 2-14	GV 14						
UC068 2-15	GV 15						
UC068 2-17	GV 17						
UC068 2-18	GV 18						
UC068 2-19	GV 19						
UC068 2-20	GV 20						
UC079 1-3	GV 79						
UC092 1-4	GV 92	BV-Intermediate					
UC058 1-1	GS 58	BV-positive	*G. swidsinskii*	2	100 %	0 %	0 %
UC094 3-2	GS 94						
UC074 1-2	GP 74		*G. piotii*	2	100 %	100 %	100 %
UC094 3-3	GP 94						

[a] Each strain was provided with a name: initials of species (GV: *G. vaginalis*; GP: *G. piotii*; GS: *G. swidsinskii*) followed by strain ID

[b] ATCC 14018

3.2.2 Expression of vaginolysin

PCR with *vly*-specific primers showed that all strains produced an amplicon with a molecular weight of 156 bp (Figure 3.1), similar to an earlier study that showed most *Gardnerella* spp. were positive for the gene [131].

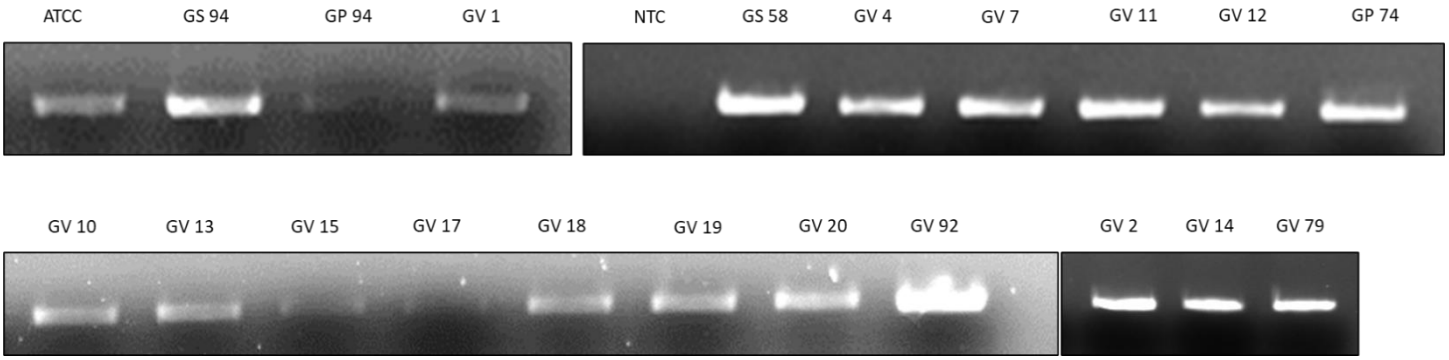


Figure 3.1: Presence of the *vly* gene. PCR using primers specific for *vly* was conducted on all strains. gDNA was obtained using Proteinase K and 2 µl of gDNA was used for a PCR. Products were run on a 1.2 % agarose gel for 1 hour at 90 V. The expected band of 156 bp for all strains. NTC – no template control indicated no contamination.

Despite the presence of the gene, *vly* expression levels were highly variable. When grown under planktonic conditions, six *G. vaginalis* and one *G. swidsinskii* strain expressed *vly* significantly higher than ATCC whereas five strains' expression was significantly lower than the control (Figure 3.2). When strains were grown as biofilms, differences became more marked, and 16 strains expressed *vly* differently with the majority of these strains showing significantly reduced *vly* transcription. Interestingly, *G. swidsinskii* and *G. piotii* expression of *vly* was more consistent than *G. vaginalis* strains, with three out of four having negligible expression, irrespective of whether bacteria were grown as biofilms or plankton. Overall, when bacteria were grown as biofilms, *vly* expression was highly variable with most strains showing a significant reduction in expression of the virulence factor.

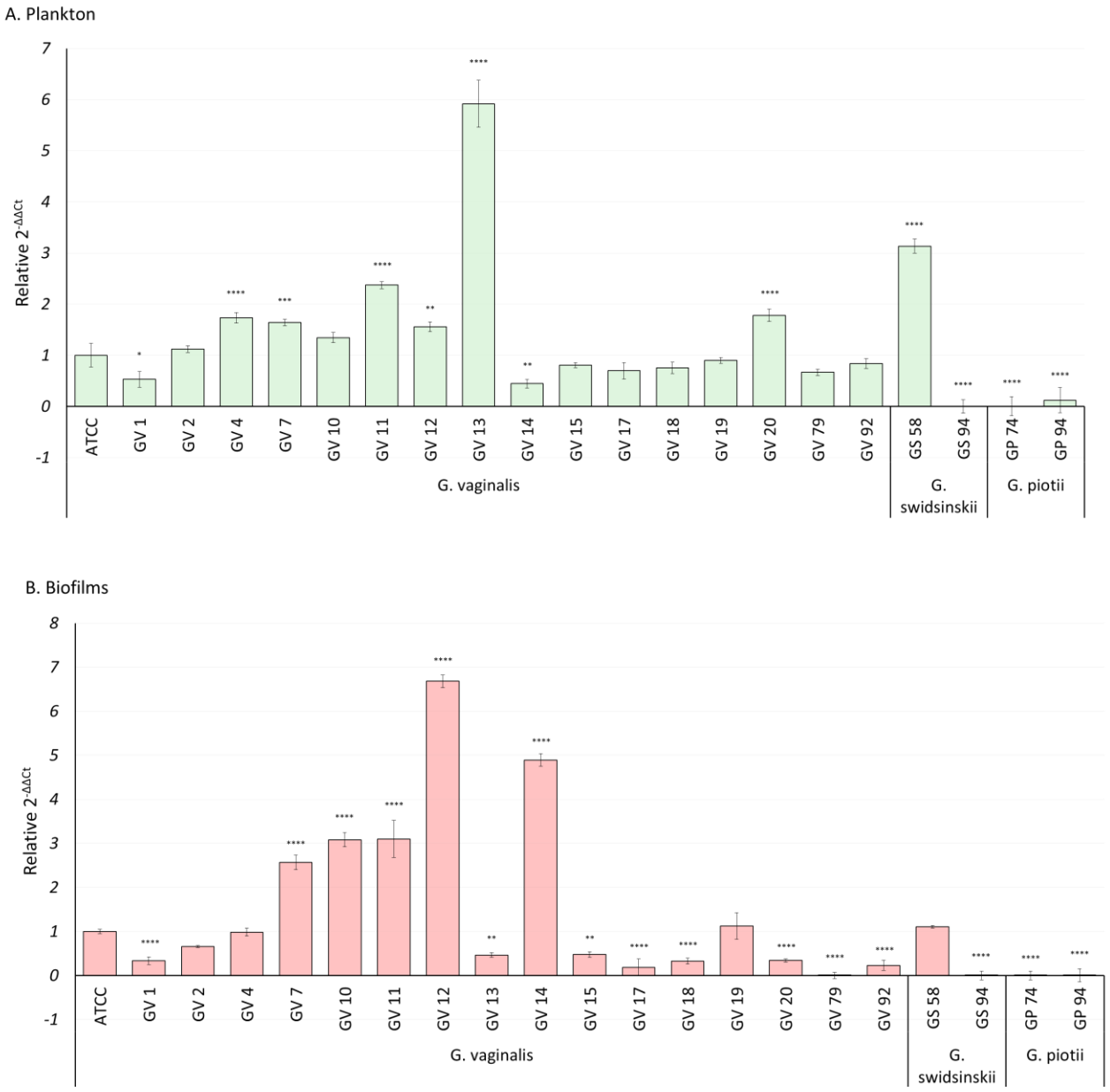


Figure 3.2: Expression of *vly* in *Gardnerella* strains under planktonic (A) and biofilm (B) conditions. *Gardnerella* strains were cultured in NYCIII medium as A) plankton or B) biofilms for 48 hours at 37 °C under anaerobic conditions. RNA was extracted from cell pellets and 90 ng of RNA was used to synthesize cDNA for all samples. SYBR green was then used in a qPCR with primers specific to *vly* to determine expression levels. All measurements are indicated relevant to ATCC strain 14018. Error bars represent the standard deviation of average $2^{-\Delta\Delta}$CT values of three biological repeats. Statistical analysis was performed on Graph Pad Prism with Dunnett's multiple comparisons test (*, < 0.05; ** <0.01; *** < 0.001; **** < 0.0001).

3.2.3 Comparison of *vly* promoter

As strains expressed *vly* to varying degrees, the promoter regions of six strains were compared to determine whether sequence variation was associated with expression levels. The 273 bp region upstream of the start codon was identical for each species and the *G. vaginalis* sequences only differed from the *G. piotii* sequences by 13 nucleotides. On the other hand, the sequence of two *G. swidsinskii* strains were very different, suggesting functional significance.

Prediction of bacterial promoter (BPROM) [137] analysis indicated that *G. vaginalis* and *G. piotii* strains had identical -10 box and -35 box sequences, contrary to GS 58 which was very different although identical to a known *G. swidsinskii* strain (Accession number: GCA_023277705.1) (Figure 3.3). There seemed to be a relationship between promoter sequence and expression when GV 79, GP 74 and GP 94, with identical promoter elements, were grown as biofilms because all had barely detectable *vly* expression. However, ATCC, with identical -10 and -35 sequences, expressed *vly* to detectable levels. On the other hand, GS 58, with a unique putative promoter sequence, had high relative *vly* expression, irrespective of whether grown as plankton or biofilms suggesting that the unique promoter elements might affect the expression levels of the virulence determinant. The – 10 box of the *G. swidsinskii* strains was identical to the consensus sequence of strong promoters, TTGACA [138] while the promoter element of GV 79, GP 74 and GP 94 differed by two nucleotides. Moreover, *G. piotii* and *G. vaginalis* sequences had a guanine instead of a thymine at position – 34 and Kobayashi*, et al.* [138] showed that a substitution at this site caused a marked decrease in promoter strength [138]. However, this does not hold true for the ATCC and GS 58 strains as they did not share unique sequences that might explain their high expression levels. Without sequencing the promoter of the remaining strains and further analysis, we cannot confirm whether differences in promoter sequences drive changes in expression of *vly*.

Figure 3.3: Comparison of *vly* putative promoter regions. Sequences comprising 273 nucleotides upstream of the start codon of vly of 4 strains were aligned to ATCC 14018 and JNFY3, a known *G. swidsinskii* strain, using BioEdit and BPROM analysis identified putative -10 and -35 box sequences. - 10 box outlined in blue, – 35 box outlined in red. Green arrows indicate – 34 where guanine is present in *G. vaginalis* sequences while thymine is present in *G. swidsinskii* sequences.

3.2.4 Activity of VLY

VLY activity was compared across all strains using a human erythrocyte lysis assay as previously described [68,121]. Strains were grown as plankton (Figure 3.4 A) and biofilms (Figure 3.4 B) to determine whether the difference in growth conditions affected haemolytic activity. As expected, GV 5, which does not belong to the *Gardnerella* genus and does not possess *vly*, showed significantly lower haemolytic activity compared to the ATCC strain. Comparison between GV 5 and the NYC III medium-only control indicated that the strain had low levels of haemolytic activity, suggesting the presence of bacterial compounds with cytotoxic activity.

Relative to ATCC, haemolysis occurred in the presence of the planktonic cell-free culture medium of all strains but was significantly higher in seven strains (Figure 3.4 A). However, only three of the seven strains had expression levels significantly higher than ATCC (Figure 3.2 A). When cells were grown as biofilms, GV 10, GV 11, GS 58 and GS 94 had the highest relative levels of haemolysis but only GV 10 and GV 11 were high VLY expressors. On the other hand, GS 58 expressed VLY similar to ATCC and GS 94 expression was barely detectable.

Together, these results suggest that better expression of *vly* might not translate into higher haemolytic activity. There was no correlation between haemolytic activity and VLY expression and

differences in haemolysis between strains were not consistent when bacteria were grown as plankton or biofilms, confirming the influence of growth condition on the phenotype of *Gardnerella* strains.

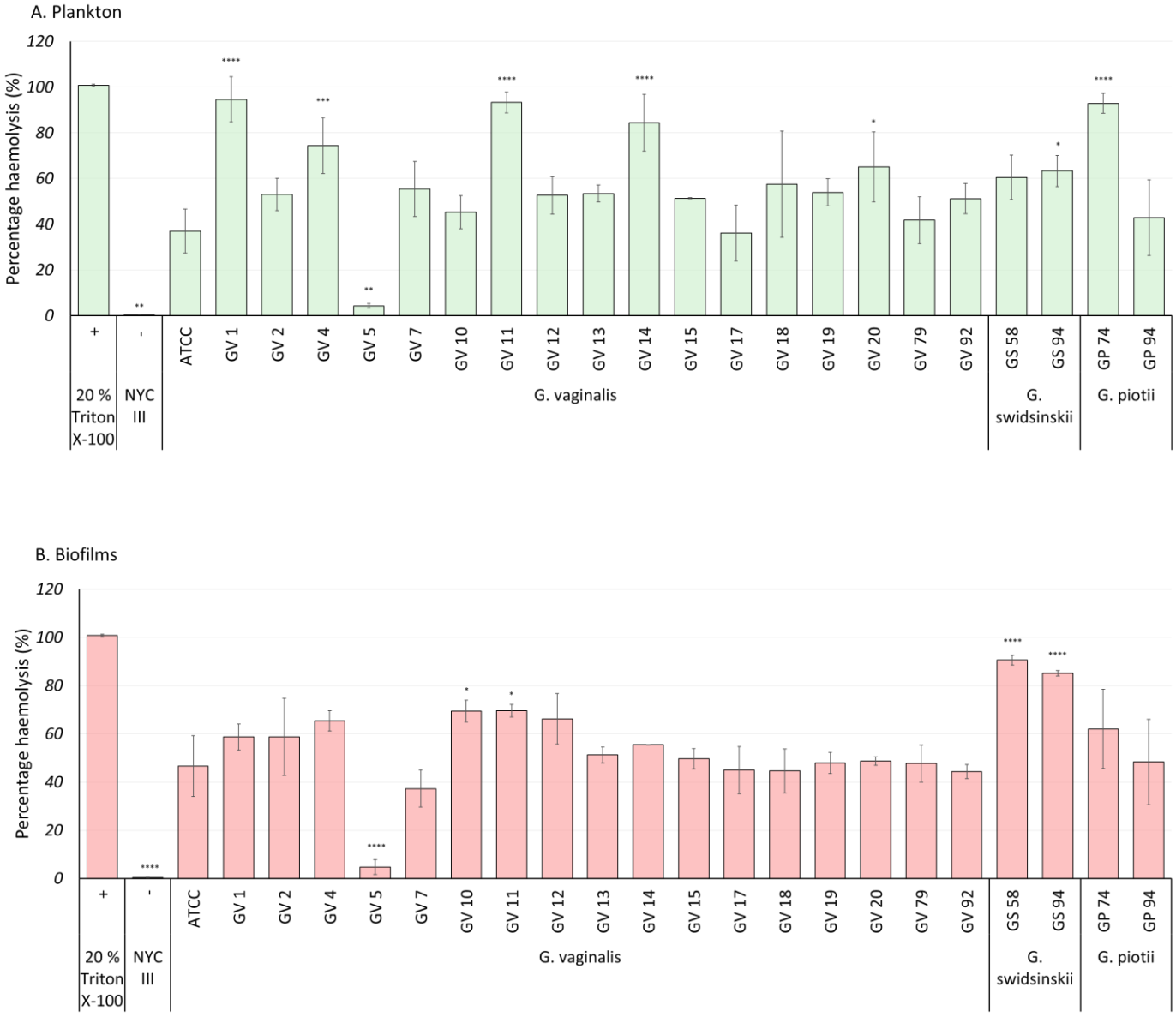


Figure 3.4: Haemolytic activity of *Gardnerella* strains. *Gardnerella* strains were cultured in NYCIII medium for approximately 48 hours at 37 ° C under anaerobic conditions as either A) plankton or B) biofilms. The supernatants were filter sterilised and a BCA assay was used to determine the protein concentration of each sample. A total of 50 µg was then added to human erythrocytes in 96-well plates to observe haemolysis. The plates were incubated at 37 ° C for 1 hour before measuring haemolysis by absorbance at 450 nm. "NYCIII" determined that the medium itself was not causing any haemolysis while 100 % lysis "+" was demonstrated by 20 % Triton-X 100. Error bars represent the standard deviation between three biological repeats. Statistical analysis was performed relative to ATCC in Graph Pad Prism using Dunnett's multiple comparisons test (*, < 0.05; ** <0.01; *** < 0.001; **** < 0.0001).

The haemolysis assay was repeated in the presence of retrocyclin (RC-101), a compound known to inhibit VLY, to confirm that the haemolytic activity associated with the culture medium of *Gardnerella* strains was due to VLY [139]. RC-101 reduced haemolytic activity compared to the untreated sample, with the exception of GV 5 (Figure 3.5), confirming the role of VLY. We cannot exclude the possibility that RC-101 might not be specific for VLY given its function as a general anti-chaperone of toxins [140].

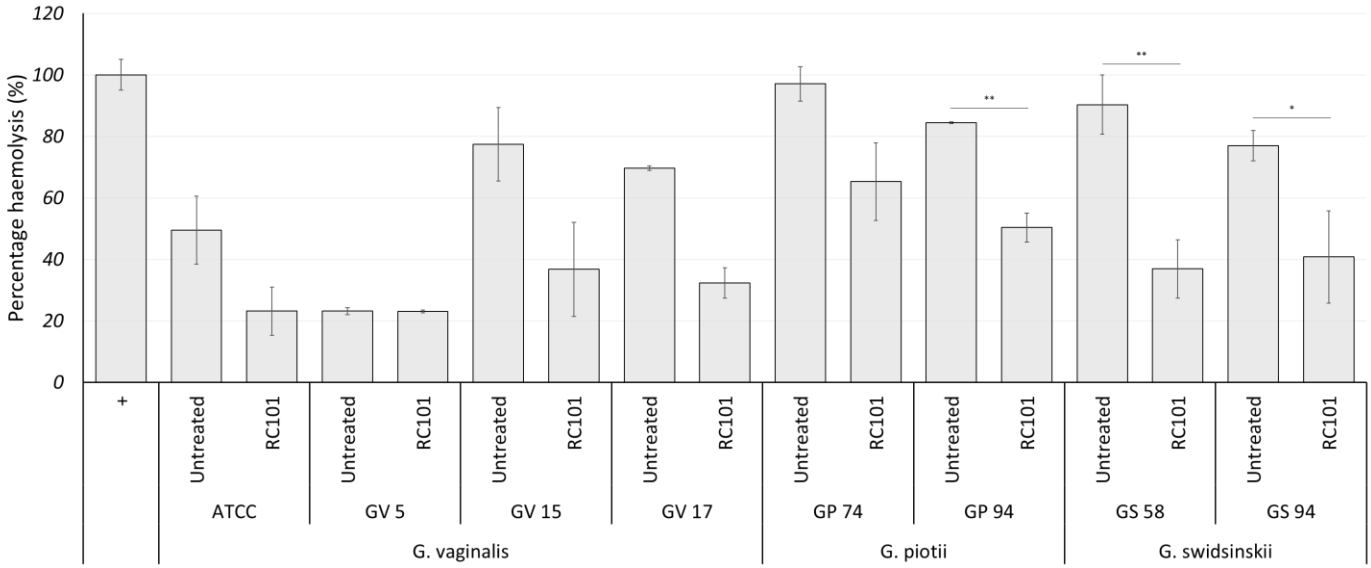


Figure 3.5: Inhibition of haemolytic activity of *Gardnerella* strains. *Gardnerella* strains were cultured in NYCIII medium for approximately 48 hours at 37 ° C under anaerobic conditions. The cell-free culture medium was filter sterilised and the equivalent of 50 µg total protein was added to human erythrocytes in the presence and absence of RC-101 (0.4 mg/ml). Haemolysis was detected at a wavelength of 450 nm. Graph Pad Prism 8 was used to perform Dunnett's multiple comparisons test, relative to + (*, < 0.05; ** <0.01; *** < 0.001; **** < 0.0001).

3.2.6 Biofilm formation by *Gardnerella* spp.

When strains were grown as biofilms there was a reduction in *vly* expression (Figures 3.2 B). We hypothesised that strains better able to form biofilms would have lower *vly* expression potentially manifesting as weaker haemolytic activity. Alternatively, as it was previously suggested that high VLY activity promoted biofilm formation [141]. Regardless, there might be a directly proportional relationship between the two virulence determinants. When bacteria were cultured in 96-well

plates for biofilm formation [92,123,132,139], there was variation between strains (Figure 3.6). GV 10, GV 13 and GV 17 produced significantly lower biofilm levels compared to ATCC. Conversely, GV 79, GV 92, GS 58, GP 74, and GP 94 were better biofilm-producers, albeit only approximately 1.4-fold higher than ATCC (p = > 0.05) and with only GP 74 reaching significance. The two *G. piotii* strains, GP 74 (p = 0.017) and GP 94 (ns) that were among the higher biofilm-forming strains, had negligible expression of *vly* and variable haemolytic activity. Therefore, there was no correlation between biofilm levels and *vly* expression or VLY haemolytic activity (R^2 = 0.1238 and R^2 = 0.1175, respectively) (data not shown), suggesting that there was no clear relationship between the two virulence factors. It is possible that the high variation between biological repeats, caused by the complexity of biofilm formation, [142] might have affected the outcome of the experiment.

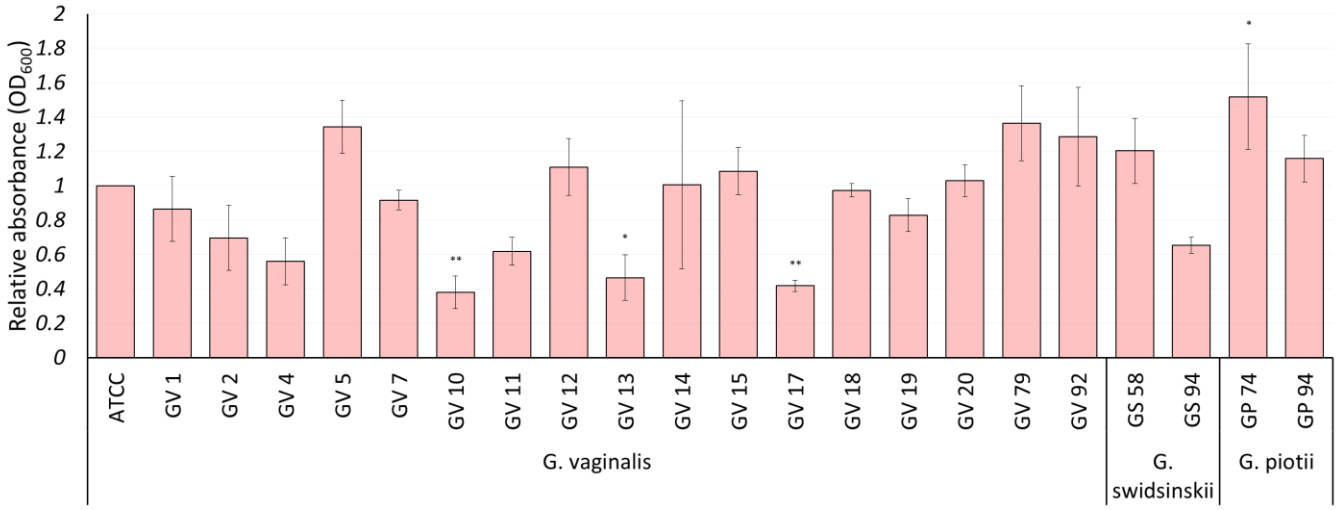


Figure 3.6: Biofilm formation in *Gardnerella* strains. *Gardnerella* strains were cultured for 48 hours under anaerobic conditions. The cultures were diluted to an OD_{600} of 0.2 and plated in 96-well plates. The plates were then incubated at 37 ° C for 48 hours under anaerobic conditions. Thereafter, the planktonic fraction was removed while the biofilm remained. Crystal violet was used to stain the biofilm cells and measured spectrophotometrically at 590 nm. Error bars are a representation of the standard deviation between three biological repeats. Statistical analysis was performed using Graph Pad Prism with Dunnett's multiple comparisons test (*, < 0.05; ** <0.01; *** < 0.001; **** < 0.0001).

3.2.7 The effect of cell-free medium on HIV infection

The cell-free culture medium of most strains grown as plankton did not affect HIV infection and cell viability. Only four strains significantly reduced cell viability with concomitant decreased HIV

infection. There was a 40 % decrease in HIV infection in the presence of GV 79 culture medium ($p < 0.001$), approximately 30 % reduction for GP 94 (ns), and approximately 90 % inhibition associated with GS 58 and GS 94 ($p < 0.0001$) (Figure 3.7). Interestingly, the *G. swidsinskii* strains decreased cell viability by approximately 90 % and this species has been shown to be highly cytotoxic (Figure 3.7 B) [131]. All other strains had no effect on cell viability, suggesting that VLY concentration was too low to lyse the reporter cell line. Of these, five strains significantly increased infection (GV 2, GV 15, GV 17, GV 19 and GV 92), and all had similar relative VLY expression and activity. Strains were identified as having a positive, but non-significant, effect on HIV infection if the means were greater than two standard deviations of the control (Table 3.2) [143]. The strains that fell within this category had highly variable VLY expression and activity. The rest of the isolates had no effect on HIV infection but also had very different VLY expression and activity relative to ATCC. Therefore, strains could be loosely grouped according to their impact on HIV infection and cell viability and those that significantly increased HIV infection seemed to have very similar VLY levels relative to ATCC in the cell-free culture medium (Table 3.2).

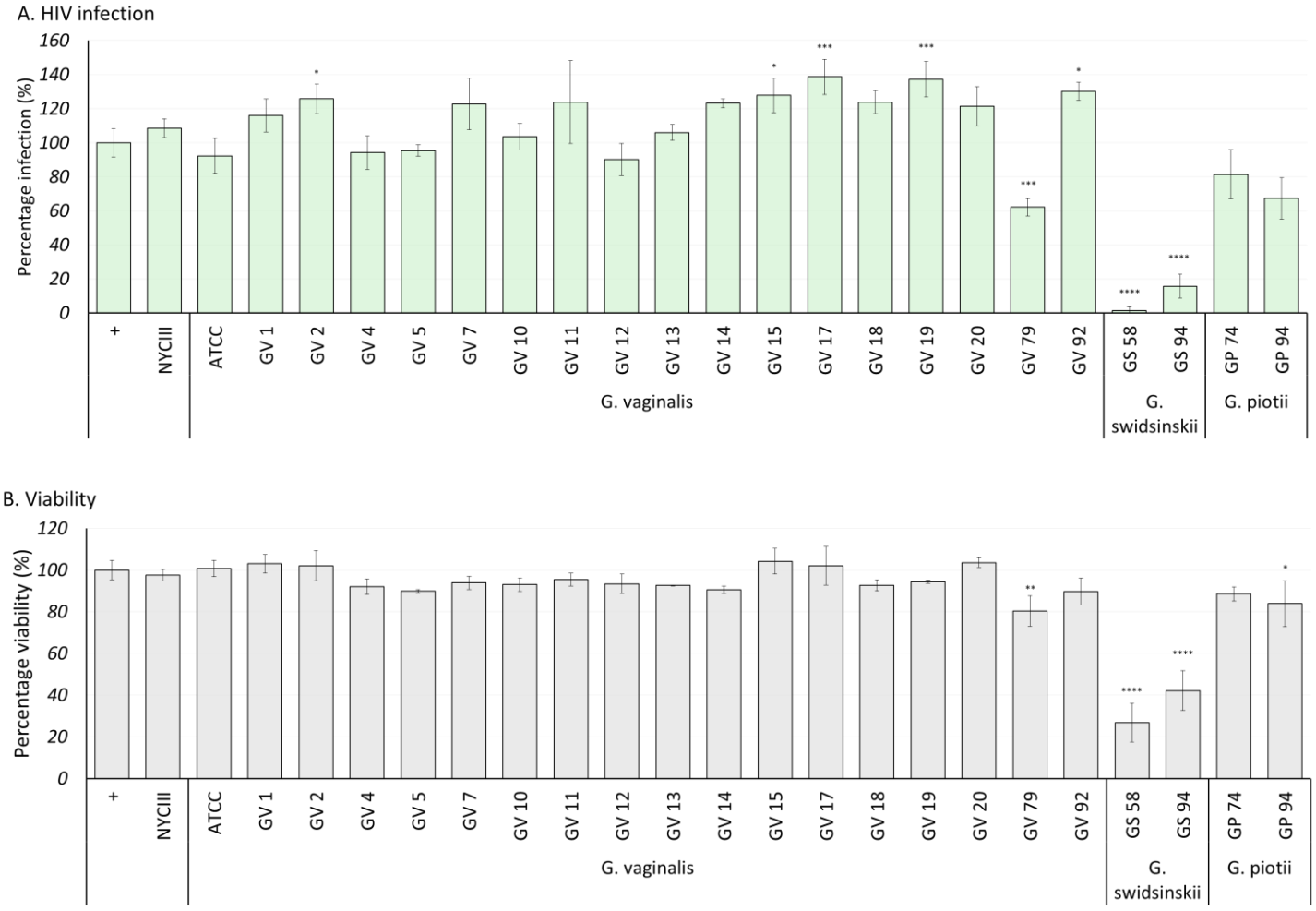


Figure 3.7. Effect of *Gardnerella* cell-free culture medium on HIV infection and cell viability. Culture medium of *Gardnerella* strains equivalent to 10 µg of total protein were added to TZM-bl cells in the presence of 200 TCID50 units of pNL4 IMCs. A) HIV infection was monitored by changes in luminescence and measured in relative light units and B) cell viability was determined by MTT assay. All values were calculated relative to ATCC. Positive control (+) represented virus-only and the negative control (NYC III) was a medium-only control. Error bars represent the standard deviation of the mean of three biological repeats. Statistical analysis was performed using Graph Pad Prism with Dunnett's multiple comparisons test relative to the ATCC strain (*, < 0.05; ** <0.01; *** < 0.001; **** < 0.0001).

Table 3.2: Effect of planktonic *Gardnerella* spp. culture medium on HIV infection

Strain	[a]Effect on HIV infection	[a]Effect on cell viability	[b]Difference in VLY expression	[b]Difference in Haemolysis	[b]Difference in Biofilm formation	Comment
ATCC	ND	ND	0	0	0	No effect on HIV infection and cell viability
GV 1	ND	ND	-*	+****	ND	
GP 74	ND	ND	-****	+****	+*	
GV 13	ND	ND	+****	ND	-*	
GV 4	ND	ND	+****	+***	-^{ns}	
GV 12	ND	ND	+**	ND	ND	
GV 10	ND	ND	ND	ND	-**	
GV 14	+^{ns}	ND	-**	+****	ND	Increased infection with no effect on cell viability with variation in VLY
GV 11	+^{ns}	ND	+****	+****	-^{ns}	
GV 20	+^{ns}	ND	+****	+*	ND	
GV 7	+^{ns}	ND	+***	ND	ND	
GV 18	+^{ns}	ND	ND	ND	ND	Increased infection with no effect on cell viability and no variation in VLY
GV 2	+*	ND	ND	ND	-^{ns}	
GV 15	+*	ND	ND	ND	ND	
GV 17	+***	ND	ND	ND	-**	
GV 19	+***	ND	ND	ND	ND	
GV 92	+*	ND	ND	ND	+^{ns}	
GV 79	-***	-**	ND	ND	+^{ns}	Decreased infection due to cytotoxicity caused by VLY or another factor
GS 58	-****	-****	+****	+^{ns}	+^{ns}	
GS 94	-****	-****	-****	+*	-^{ns}	
GP 94	-^{ns}	-*	-****	ND	+^{ns}	

[a]Differences were calculated relative to the positive control: IMC infection in the absence of cell-free culture medium. If means were greater than two standard deviations of the positive control, then they were classified as higher (+) and if the mean was lower by two standard deviations (SD) then the sample was classified as lower (-). If the means fell between +/- two SD then they were identified as similar to the control (ND) [143].

[b]Differences were calculated relative to the ATCC 14018 strain.

Statistical significance is shown as *, < 0.05; ** <0.01; *** < 0.001; **** < 0.0001 and ns, not significant but +/- two SD. When multiple clones were grouped, no p-value was indicated because the differences between ATCC and each strain were not consistently significant.

Correlation analysis was conducted on all variables but no correlations were found.

3.3 Discussion

Recently, there has been a focus on understanding whether the phenotypic and genotypic diversity within and between *Gardnerella* species or subtypes were associated with the presence or absence of BV, and whether this relationship was related to differences in virulence [69,81,144]. The genus, *Gardnerella,* consists of *G. vaginalis* (subgroup C/clade 1), *G. leopoldii* and *G. swidsinskii* (subgroup A/clade 4), and *G. piotii* (subgroup B/clade 2) and a further nine genomic species, confirming that the genus is much more diverse than originally believed [69]. Furthermore, as *Gardnerella* is found in the FRT of both BV-positive and BV-negative women, it was suggested that perhaps only some species or strains were pathogenic. Isolates from BV-positive women had better adherence, aggregation, and biofilm formation than a strain isolated from a BV-negative participant [92]. However, there have been no consistent associations of any species with BV [145] and whether this is due to the inherent complexity of the FRT microbiome or variation across study design remains unknown. Despite the absence of a clear association, *Gardnerella* spp. remain the prime candidate for initiating dysbiosis as it is one of the most common genera associated with BV and has virulence determinants such as biofilm formation and sialidase and VLY activity [146].

BV has been linked to an increase in HIV infection [117,135], and we hypothesised that *Gardnerella* spp., by virtue of its association with BV, might directly enhance HIV transmission. Therefore, this study determined whether there was a relationship between the expression and activity of VLY and biofilm production of twenty *Gardnerella* strains, and their impact on HIV infection.

VLY is a cytolysin linked to immune evasion, inflammation, cytotoxicity, and changes in plasma membrane structure [101]. Each of these mechanisms could enhance *in vivo* HIV entry into epithelial cells and/or infection of $CD4^+$ T cells, as entry into host cells depends on endocytosis, macropinocytosis, transcytosis, transinfection and migration via abrasions of the epithelial barrier [147]. The action of VLY depended on its concentration: at low levels it caused membrane restructure and at high concentrations, it became cytolytic [101]. Similar to what was suggested for BV, the mechanisms by which VLY may contribute to HIV infection could be dependent on its

concentration in the FRT [68] and thus variation in *vly* expression was determined for each of the twenty strains.

All strains in this study carried *vly* although the presence of the gene did not translate into VLY expression. One *G. swidsinskii* and both *G. piotii* strains expressed barely detectable levels of *vly* regardless of whether bacteria were grown as plankton or biofilms (Figures 3.2 A & B). It has been suggested that *vly* was exchanged via horizontal (lateral) transfer and not part of the core *Gardnerella* genome [148]. Bacterial genes shared via horizontal transfer can be silenced through the action of the ubiquitous histone-like nucleoid structuring protein (H-NS) [149]. As far as we are aware, *Gardnerella* spp. does not express H-NS, however it is tempting to speculate that some strains might be better at gene silencing than others.

Similar to a previous study, *vly* expression was much lower when strains were grown as biofilms compared to plankton [123] and highly variable across species and intra-species strains [81]. Comparison of the promoter regions of *vly* indicated that there was no relationship between -10 and -35 box sequence variation and expression levels and thus other regulatory sites might be involved. Pleckaityte*, et al.* [81] found that differences in haemolysis did not coincide with levels of VLY in the culture medium, nor apparent sequence changes in the promoter regions [81]. VLY is mostly secreted into the culture medium but is also encapsulated in membrane vesicles that are taken up by epithelial cells. They were found to induce blebbing, activate the p38 signalling pathway and lyse erythrocytes better than whole *Gardnerella* cells [150]. These vesicles also carried other bacterial proteins that might contribute to erythrocyte haemolysis and hence skew the direct relationship between VLY expression and haemolytic activity. Without analysing sequences from additional strains, we cannot confirm the impact of promoter changes on VLY expression.

Even if gene silencing could explain the lack of expression in the presence of *vly*, it would not explain why there was no relationship between the level of haemolytic activity associated with the culture medium and VLY expression. GS 58 and GP 74, with negligible expression had high haemolytic activity. Harwich*, et al.* [92] found that two strains that differed in cytotoxicity had the same expression levels with nearly identical promoter sequences [92]. This may suggest that more

than one factor contributes to haemolytic activity or cell death is due to a different mechanism. VLY activates p38 MAPK signalling pathway which induces programmed cell death by apoptosis and necrosis [151]. *Gardnerell*a spp. might also produce an alternative cytotoxin resistant to RC-101 because haemolytic activity was reduced by only 50 % (Figure 3.5). Alternatively, the concentration of RC-101 might have been too low to achieve 100 % inhibition.

Another alternative is that either processing and/or structural variation of VLY could influence the haemolytic activity found in the culture medium. When the predicted signal peptides were compared, intra-species strains had identical sequences and the differences between species seemed to coincide with activity associated with planktonic cell-free culture medium. It has been demonstrated that lysine or arginine in the second position of the signal peptide results in a positive charge and more efficient secretion of multiple proteins [152]. The signal peptide of the *G. swidsinskii* strain had a lysine at that position whereas *G. piotii* sequence carried an asparagine. *G. swidsinskii* may therefore be secreted more efficiently resulting in higher haemolytic activity in some cases. However, *G. vaginalis* strains' signal peptide also carried a lysine but did not consistently have high activity compared to the other strains. It is possible that polymorphisms at other sites could be affecting the rate at which VLY is secreted or another mechanism of cellular transport such as membrane vesicles could also play a role [150].

Comparison of the coding region of VLY indicated conservation within species and low inter-species diversity with an identity of approximately 94 %. Two regions of VLY are critical for its function: the undecapeptide region, involved in membrane interaction, and the CD59 binding motif. VLY has been classified into three types (Type 1-3) according to polymorphisms within these regions. Types 1 and 2 were most common amongst isolates and Type 1 VLY was suggested to be more cytotoxic than Type 2 [131]. However, this was not consistent with our findings because the cell-free culture medium of GS 58, a Type 2 variant, had one of the highest haemolytic activities in biofilms, whereas GV 79 and GP 94 haemolytic activity ranged from 40-60 % and belonged to Type 1.

Overall, sequence analysis of the VLY promoter, signal peptide and coding region could not explain the lack of relationship between *vly* transcription and haemolytic activity. Another possibility is that the apparent variation in haemolytic activity might be due to another factor entirely. Harwich*, et al.* [92], found that the more cytotoxic strain was better at adherence to epithelial cells and suggested that other adhesins, such as biofilm associated protein (BAP), could be responsible. Although most VLY is secreted, adherence of bacteria to host cells is essential for haemolysis suggesting a role for membrane bound VLY [92]. A caveat of this study was that only secreted VLY was tested for haemolysis which might explain why variation in activity did not coincide with expression as the effect of adherence was not taken into account.

Harwich*, et al.* [92]determined that *Gardnerella* carried a gene encoding BAP which mediates binding of bacteria to host cells hence facilitating biofilm formation [92]. As adherence to epithelial cells is the first step to biofilm formation, all 20 strains were compared for the ability to form biofilms. Four had significantly lower, and GP 74 had significantly higher, biofilm-forming capacity than ATCC. Overall, there was extensive variation between strains and the ability to form biofilms was not associated with increased *vly* expression. In fact, *vly* expression was reduced under biofilm growth conditions compared to when grown as plankton, confirming previous findings [123]. For future studies, it would be interesting to determine whether BAP expression is associated with biofilm formation.

To determine whether strains affected HIV infection differently, bacteria were grown as plankton and TZM-bl cells were infected in the presence of the cell-free culture medium. Given that TZM-bl cells express CD59 and their membrane contains over 50 % cholesterol, we expected VLY to either cause blebbing, and enhance HIV infection or cytolysis and reduce infection, [153,154] depending on its concentration [68,135]. The majority of the strains had the same effect on HIV infection and cell viability as ATCC, suggesting that they were all similar in those respects. However, the VLY expression, haemolytic activity and biofilm formation were highly variable. Four strains, GV 79, GS 58, GS 94 and GP 94 apparently inhibited infection but as they also significantly decreased cell viability, the reduced infection was as a result of their high cytotoxicity. The *G. swidsinskii* strains, when grown as plankton, had high haemolytic activity relative to ATCC and

reduced HIV infection and cell viability more effectively than all of the other strains, suggesting a relationship between high VLY activity and increased cytotoxicity (Figure 3.7 A & B).

Strains that increased HIV infection could be grouped separately depending on the extent to which they influenced HIV infection without affecting cell viability. Those that significantly increased HIV infection, had a phenotype similar to the ATCC strain suggesting that they had more or less the same VLY expression and haemolytic activity. These strains could have common characteristics that might enhance HIV infection. However, without knowing the absolute concentration of VLY in the culture medium, we cannot conclude that the strains' effect on HIV infection was due to the virulence determinant. One caveat of our approach was that normalising by total protein might have influenced the outcome of the experiment as the strains might metabolise proteins at different rates, thus affecting the apparent total protein present after incubation. However, we hypothesized that the impact of bacterial metabolism on total protein would be negligible given the excess of protein present due to the horse serum.

Replacing the cell-free culture medium with purified VLY in an HIV infection assay would eliminate the impact of other proteins present in the cell-free culture medium. It would also allow for addition of more accurate concentrations of VLY that do not cause cell death. Furthermore, VLY could increase HIV infection by other mechanisms such as the disruption of the epithelium barrier and lysing of neutrophils, deregulating the immune response. A tissue model might provide more information on the mechanism of how VLY affects HIV infection [155].

Biofilm formation is highly complex and difficult to replicate in *in vitro* studies [142] which probably contributed to the high experimental error evident in biofilm-related experiments such as VLY activity (Figure 3.4 B). The different expression and phenotypic variation of VLY was not consistent when the strains were grown as biofilms or plankton which could be due to mixed cultures. The expression of *pgi,* a marker for biofilms was increased when cells were grown as biofilms (data not shown) which was contrary to what was expected [123]. *Gardnerella* and other bacteria regulate gene expression based on the mode of growth [123] and can switch between plankton and biofilms, which can result in a mixed culture of genotypically and phenotypically distinct bacteria [87]. This

could also contribute to the lack of association between *vly* expression, haemolytic activity, and biofilm formation when bacteria grown as plankton and biofilms were compared.

Overall, the results confirm the high phenotypic variation across strains from different *Gardnerella* spp. and highlight the need for more suitable models to investigate the role of VLY in HIV infection. Despite the challenges, we found *G. swidsinskii* strains were highly cytotoxic and could play a protective role *in vivo* by lysing epithelial cells and HIV permissive cells, such as Langerhans cells and macrophages. There was no clear relationship between VLY and HIV infection. However, by grouping strains according to their impact on infection, a potential pattern emerged. Strains that were highly effective at increasing HIV infection, had very similar relative VLY expression and haemolytic activity. On the contrary, other strains that had no or non-significant impact on HIV infection had highly variable VLY levels and biofilm forming abilities.

Chapter 4: Sialidase

4.1 Introduction

Sialidase activity in cervicovaginal fluid is associated with BV [96]. Sialidase is an important virulence factor produced by *Gardnerella* spp. and other pathogens in the FRT. The role of sialidase activity in pathogenesis is not yet fully understood but it has been linked to biofilm formation [3,156], impairment of immunity by degradation of IgA and mucin [52,98] and most importantly, increase in HIV infection [112,114]. It has recently been determined that there are three isoforms of *Gardnerella* sialidase, NanH1, NanH2 and NanH3 and that the latter two are responsible for the sialidase activity despite the *nanh1* gene being found in most *Gardnerella* spp. On the other hand, *nanh2* and *nanh3* are found only in *G. piotii* and *G. vaginalis* strains. While the largely inactive *nanh1* does not have a defined role as yet, it is possible that it has an unidentified function in the FRT.

As BV has been linked to enhanced HIV acquisition, the sialidase of *Gardnerella* spp. might play a direct role in increasing HIV infection. Firstly, it could act to remove the terminal sialic acids of mucin which contribute to the rigidity of cervical vaginal mucin (CVM). The loss of the sialic acid moieties may decrease the viscosity of CVM and thereby, *Gardnerella* sialidase could diminish its protective mechanism against HIV infection [157]. Secondly, it is possible that sialidases could potentially affect HIV infection through the removal of terminal sialic acids of complex N-glycans of Env. The efficiency of HIV infection is dependent on the type and position of N-glycans on Env [158,159] and the loss of sialic acids from complex N-glycans has been shown to increase infection [113]. It is thus possible that the sialidases could desialylate gp120 and enhance its binding to CD4 and CCR5. Infection was measured using TZM-bl cells, a HeLa luciferase reporter cell line expressing high levels of CD4, CCR5 and CXCR4 and are thus not normally permissive to HIV infection. Alternative models such as PBMCs and *ex vivo* tissue explants might have been more physiologically relevant, but these also have limitations such as door variation, differential impact of tissue type on Vly and the potential for PBMCs to produce inhibitors of HIV infection [160,161].

Lastly, the non-specific desialylation of viral and/or cell surface glycoproteins could facilitate attachment of the virus to the host cells by decreasing the repulsive forces between the negatively charged membranes [112,114]. TZM-bl cells (a HeLa cell derivative), the common reporter cell line used in HIV infection assays are sialylated [162] and thus desialylation may contribute to a reduction in negative repulsion and increase HIV infection. Irrespective of the exact mechanism of action of sialidase, it might play a significant role in the relationship between BV and HIV infection. It is thus important to determine whether *nanh1, nanh2* and *nanh3* expression levels and activity are associated with changes in HIV infection.

4.2 Results

The two isoforms, NanH2 and NanH3 are mainly responsible for the sialidase activity associated with *Gardnerella* [1]. However, *nanh2* was not detected in any of our strains, despite designing primers specific to a highly conserved region (data not shown). This suggested that *nanh2* was not present in any of our strains, similar to a previous study where *nanh2* was absent in their *G. vaginalis* isolates [163]. However, the absence of *nanh2* in *G. piotii* sequences could be due to sampling bias because there were only two strains in our sample set. In the absence of detectable NanH2, this study focussed its analysis on NanH1 and NanH3.

4.2.1 Analysis of NanH1

The gene for NanH1 is common to *Gardnerella* spp. An earlier study found *nanh1* present in 75 % of 120 strains [141] and Bulavaite*, et al.* [163] found it in 100 % of their samples. Similarly, PCR with *nanh1*-specific primers indicated that most of our strains (90 %) carried the *nanh1* gene excluding the *G. swidsinskii* strains (Figure 4.1). Interestingly, both *G. piotii* bands corresponding to the expected molecular weight of the *nanh1* PCR product appeared to be slightly smaller than those belonging to *G. vaginalis*. This is unlikely due to contamination as the negative control and the *G. swidsinskii* samples were not positive for the 636 bp amplicon. Furthermore, upon sequencing, it was found that GP 74 and GP 94 had two unique deletions absent in GV 79, which resulted in a 27 bp shorter fragment.

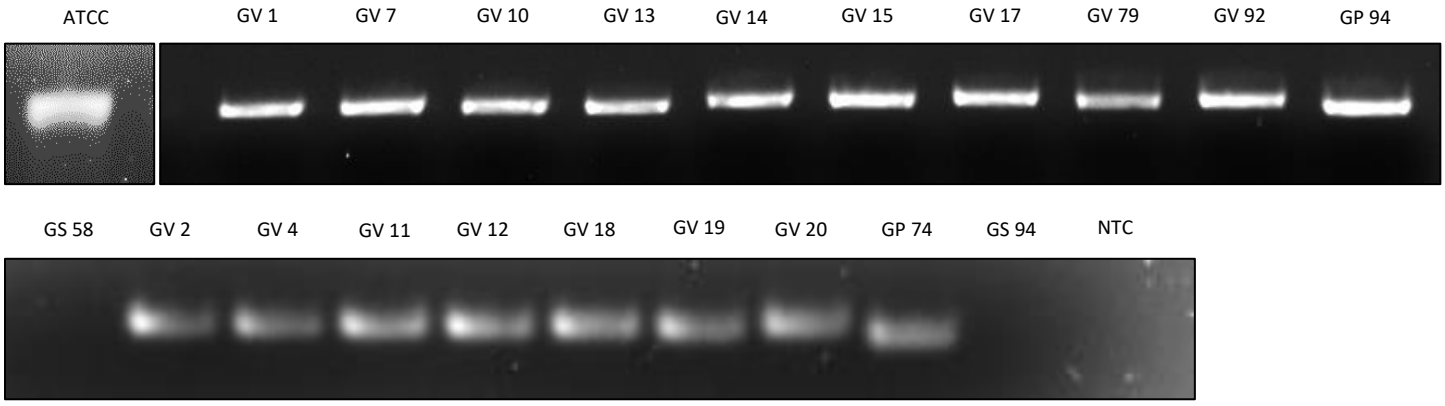


Figure 4.1: Detection of *nanh1*. *Gardnerella* strains were grown on NYCIII agar plates. A colony forming unit (cfu) of each strain was treated with Proteinase K and the genomic DNA was amplified using *nanh1*-specific primers. The PCR products were run on a 1.2 % agarose gel for 1 hour at 90 V. The band represents a 636 bp amplicon relative to a commercial molecular weight marker (NEB) (not shown). NTC represents the negative no-template control.

4.2.1.1 *nanH1* expression

Currently there is conflicting evidence as to the importance of *nanh1* expression levels on the onset of BV. Castro*, et al.* [132] found no variation in expression among Amsel/Nugent-BV and non-BV isolates, whereas Severgnini*, et al.* [164] found that pregnant women with Nugent-BV had higher levels of *nanh1* expression compared to both BV-negative and intermediate patients [132,164]. Study design might have influenced the outcome of the studies as pregnancy might affect the FRT microbiome [161,165,166]. Hardy*, et al.* [141] showed a correlation between expression of *nanh1* and biofilm formation. Therefore, all *G. vaginalis* and *G. piotii* strains positive for the gene were grown as biofilms and plankton to control for the potential confounding effect of culture conditions on *nanh1* expression.

All strains expressed *nanh1* to different levels depending on growth mode, with some strains expressing *nanh1* better when grown as plankton compared to biofilms or *vice versa* (Figure 4.2). Furthermore, the overall variation between strains was much greater when bacteria were grown as biofilms, suggesting it could play a role in biofilm formation (Figure 4.2). However, there was no association between the biofilm-forming ability of each strain (Figure 3.6) and *nanh1*

expression (data not shown). Interestingly, both *G. piotii* strains expressed *nanh1* to high levels when grown as biofilms and were among the higher biofilm-forming strains (Figure 4.2), suggesting that impact on biofilm formation could be limited to certain species. Overall, *nanh1* expression was very different between strains and the extent to which expression varied was most apparent when bacteria were cultured as biofilms.

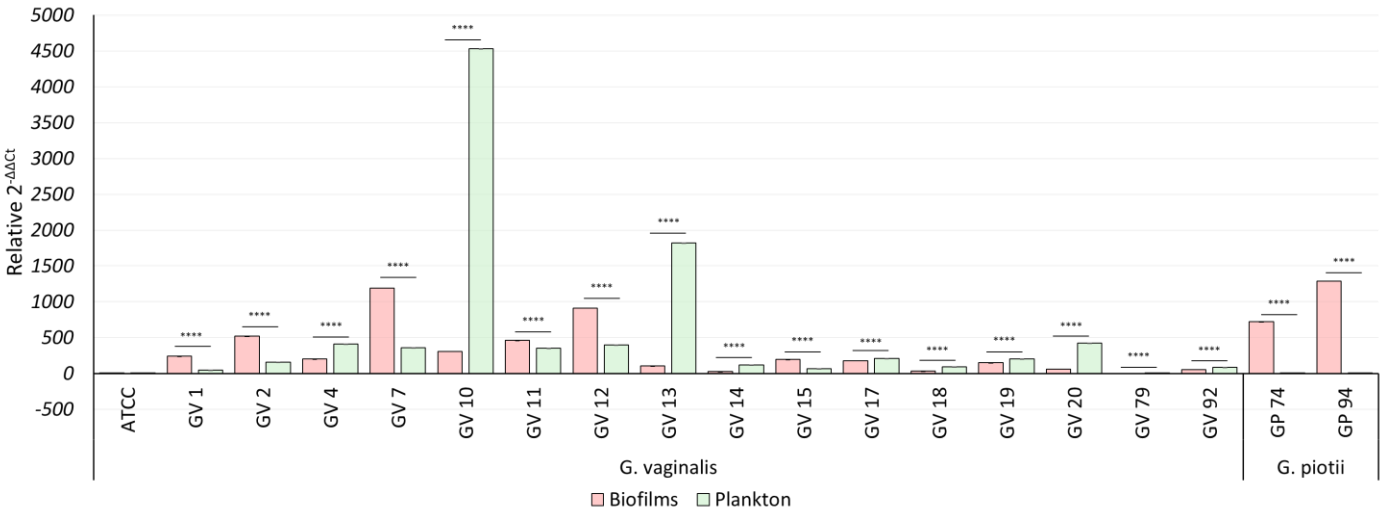


Figure 4.2. Expression of *nanh1*. *Gardnerella* strains were cultured in NYCIII medium as plankton and biofilms for approximately 48 hours at 37 ℃ under anaerobic conditions. Bacterial cultures were diluted to an OD_{600} of 0.8 before RNA was extracted, cDNA synthesised and qPCR performed using primers specific to *nanh1* and *16S* rRNA. The $2^{-\Delta\Delta_{CT}}$ was calculated for each biological repeat done in triplicate and fold-difference was calculated relative to the ATCC strain. Error bars represent the standard deviation of the average of three biological repeats. Statistical analysis was performed relative to ATCC in Graph Pad Prism using Sidak's multiple comparisons test (*, < 0.05; ** <0.01; *** < 0.001; **** < 0.0001).

4.2.1.2 NanH1 sialidase activity

Sialidase activity was measured using the fluorogenic substrate, MUN, as previously described [128]. In an earlier study, the activity of purified recombinant NanH1 on MUN was only detected at concentrations 80-160-fold higher than NanH3 [1]. Therefore, bacterial titre was maximised by culturing strains as plankton, in high volumes to increase the likelihood of detecting sialidase activity. However, when strains positive for *nanh1* were tested, sialidase activity was not detected (Figure 4.3). The concentration of sialidase might still have been too low to detect activity and/or

MUN might not be the preferred substrate of NanH1 [1]. It is possible that NanH1 might act on an unknown substrate in the FRT.

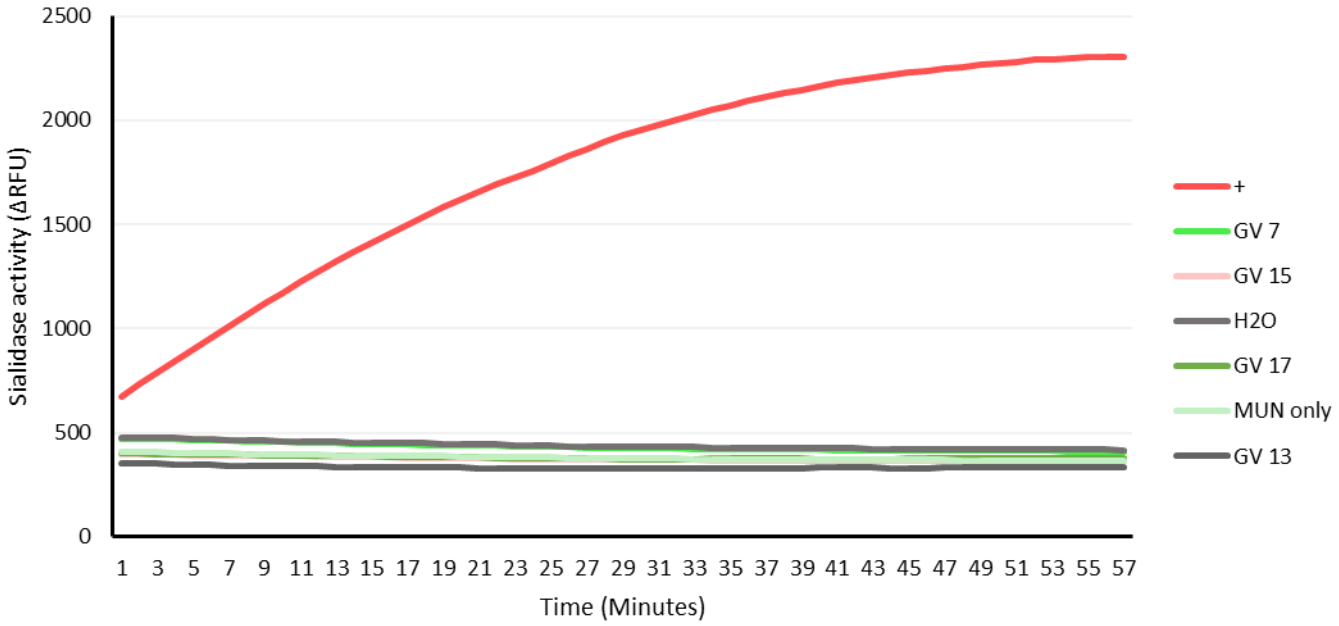


Figure 4.3. NanH1 sialidase activity. All *Gardnerella* strains positive for *nanH1* were cultured as plankton in NYCIII medium for approximately 48 hours at 37 ° C under anaerobic conditions before cells were pelleted, lysed, and sonicated and total protein measured. Sialidase activity associated with 50 µg of total protein was determined for each strain using 6 µg of MUN. The reaction was measured for 1 hour and the change in relative fluorescent units (ΔRFUI) is shown above. The positive control (+) indicates the activity of 10 U commercial sialidase of *C. perfringens* and the negative control is indicated by H_2O. The graph is a representative of 4 *nanh1*-positive strains and 2 biological repeats.

4.2.2 Analysis of NanH3

Previously, it was shown that the majority of *G. piotii* but only some *G. vaginalis* strains were positive for sialidase activity. The authors suggested that *nanh3* was transferred to *G. vaginalis* from *G. piotii* by horizontal gene transfer and was not an intrinsic part of the *G. vaginalis* genome [163]. In support of this suggestion, the gene for NanH3 was only detected in the two *G. piotii* strains in our sample set (Figure 4.4). Bulavaite*, et al.* [163] reported that one isolate in their study carried the *nanh3* gene but did not produce active enzyme [163]. It was therefore important to determine whether the two *G. piotii* strains expressed active enzyme.

Figure 4.4: Detection of *nanh3*. A colony forming unit (cfu) was obtained from the glycerol stocks of all strains and treated with Proteinase K. The genomic DNA was amplified using *nanh3*-specific primers. The PCR products were run on a 1.2 % agarose gel for 1 hour at 90 V to identify the expected 322 bp band. Only a subset of samples is shown: *G. piotii* (74 and 94) and *G. swidsinskii* (58 and 94). NTC indicates the no-template negative control. The molecular weight marker and *G. vaginalis* strains are not shown.

4.2.2.2 NanH3 expression

Like *nanh1*, *nanh3* expression levels varied between the two *G. piotii* strains with GP 74 showing significantly higher expression compared to GP 94 (Figure 4.5 A). On the contrary, when the strains were grown as biofilms, GP 94 expressed *nanh3* significantly better than GP 74 (Figure 4.5 B). Hence, growing strains as biofilms altered expression levels of *nanh3*. Genome sequencing identified approximately 150 bp upstream of the start codon which were identical between GP 74 and GP 94 and did not comprise putative promoter sequences, suggesting the transcription regulatory sites were further upstream.

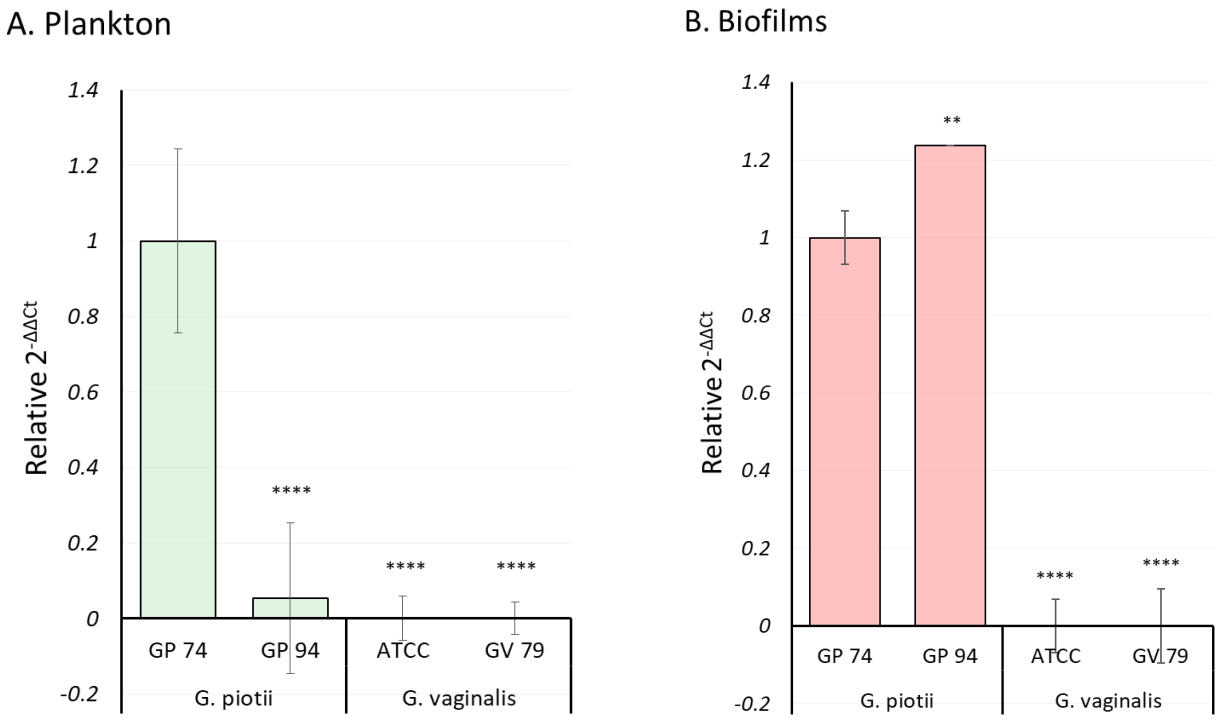


Figure 4.5: Expression of *nanh3*. *Gardnerella* strains were cultured in NYCIII medium as A) plankton and B) biofilm for approximately 48 hours at 37 ° C under anaerobic conditions. Bacterial cultures were diluted to an OD_{600} of 0.8 before RNA was extracted, cDNA was synthesized, and qPCR was performed using primers specific to *nanh3* and *16S* rRNA. The $2^{-\Delta\Delta}$CT was calculated for each biological repeat comprising 3 technical repeats and fold-difference was calculated relative to the ATCC strain. Error bars represent the standard deviation of three biological repeats. Statistical analysis was performed relative to GP 74 in Graph Pad Prism using Dunnett's multiple comparisons test (*, < 0.05; ** <0.01; *** < 0.001; **** < 0.0001).

Kurukulasuriya*, et al.* [10] identified an ORF of NanH3 consisting of a start codon, AUG, a predicted signal peptide, a poly(C) region, followed by a sialidase domain and transmembrane domain (TMD) (Figure 4.6). The poly(C) region would cause slipped-strand mispairing during replication of the bacteria, resulting in variation in homopolymer length. Depending on its length, the number of cytosines could cause a frameshift so that translation of active sialidase was initiated at an alternative start codon, GUG, 123 bp downstream of the canonical start codon [10]. GUG start codons are common in bacteria [167]. Main and alternate start codons have been identified within

the same ORF [168] but translation of transcripts starting with GUG are less efficient than AUG, and it has been suggested that this could be an important way of regulating gene expression [169].

Analysis of whole genome sequences obtained for GP 74 and GP 94, indicated that the poly(C) tract comprised of ten cytosines, which suggested that translation was initiated at GUG. Upon further analysis of the individual reads obtained by whole genome sequencing for GP 74 and GP 94, it was determined that although most of the reads had a poly(C) region of 10 cytosines, the length of the homopolymer was different for some (Tables S4.1 and S4.2). Therefore, bacterial culture may comprise of variants that expressed NanH3 with and without a signal peptide which might affect the level of sialidase activity associated with the culture medium. Nonetheless, the differences in transcript levels of GP 74 and GP 94 observed with qPCR would not be due to the shift in translation start site, suggesting regulation at transcriptional level. However, differences in translation efficiency would affect the level of sialidase activity.

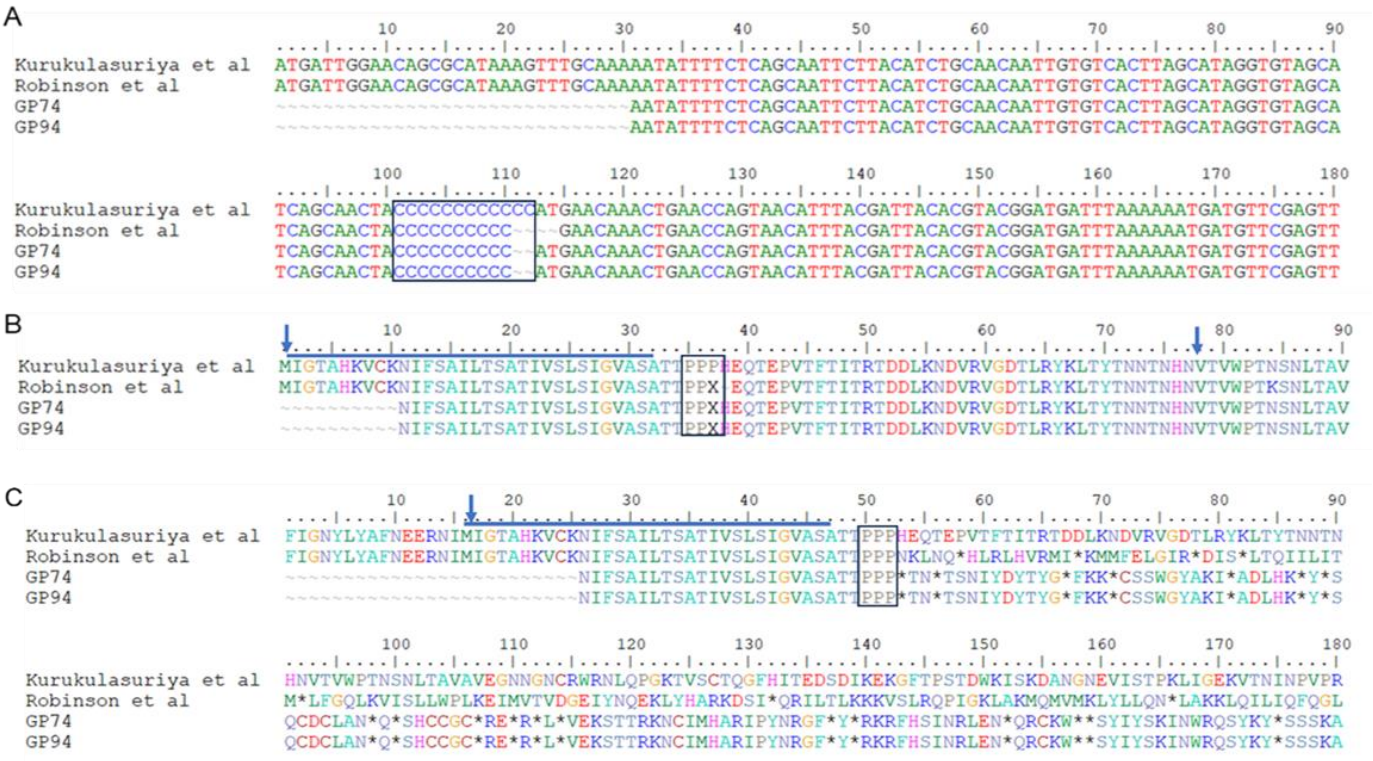


Figure 4.6: Comparison of genomic and protein Nanh3 sequences. Sequences analysed in [1] (JCP8151B) and [10] (W11) were aligned to GP 74 and GP 94 using BioEdit. A) DNA sequences indicating the stretch of cytosines [poly(C)] is highted by a box; B) translation of sequences in A) indicating the signal peptide (underlined) and the two start codons (arrows). The prolines encoded by the Poly(C) are boxed and the 'x' stands for spaces included in the sequences to maintain the open reading frame. C) Translation of sequences in A) but no spaces were included after the poly(C) to indicate the disruption of the ORF. The asterisks indicate stop codons.

4.2.2.4 NanH3 sialidase activity

Kurukulasuriya*, et al.* [10] showed that NanH3 consisted of an 812 amino acid protein, with a signal peptide and TMD, suggesting that the enzyme was secreted and localised to the surface of the cell [10]. However, as the length of the poly(C) tract determined whether the signal peptide was in frame with the sialidase domain, the ORF identified by Robinson*, et al.* [1] lacked a signal peptide. Similar to Robinson*, et al.* [1], sequence analysis of our isolates and NCBI sequences indicated that the coding region of GP 74 and GP 94 lacked a signal peptide. This was corroborated when activity was not found in the culture medium of GP 74 and GP 94 and only detected when whole cells were used in the enzyme assay (data not shown) and when cells were sonicated in the presence

of detergent (Figure 4.7). GP 74 sialidase activity was approximately two-fold higher than GP 94, most likely due to the difference in expression levels because the *nanh3* coding sequence of the two strains was identical (Figure S4.1). When the strains were grown as biofilms, sialidase activity was not detected, presumably because the bacterial titre was too low in the 96-well plates to produce detectable levels of enzyme (data not shown).

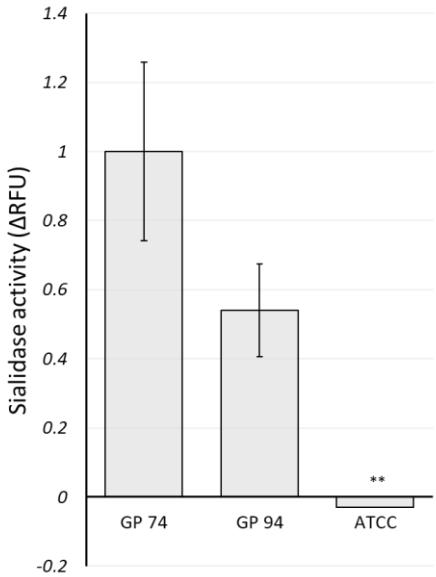


Figure 4.7: NanH3 sialidase activity. *Gardnerella* strains positive for *nanH3* were cultured in NYCIII medium as plankton for approximately 48 hours at 37 ° C under anaerobic conditions before sonication in the presence of detergent. Sialidase activity in cell lysates, corresponding to 50 µg of total protein, was determined for each strain using 6 µg of MUN. The reaction was measured for 1 hour and the change in relative fluorescent units (ΔRFU) are shown above. ATCC strain was included as a negative control because it did not carry the NanH3 gene. Error bars represent the standard deviation of three biological repeats (*, < 0.05; ** <0.01; *** < 0.001; **** < 0.0001).

4.2.2.5 Impact of *Gardnerella* cell lysates on HIV infection

Detection of NanH3 activity required sonication of bacteria in the presence of detergent. However, due to the nature of the infection assay, detergent could not be included when preparing the bacterial cell lysates. We hypothesised that even though the sialidase activity might be below detection level, it could still have an effect on HIV infection. Bacteria were thus sonicated in the absence of detergent and centrifuged to remove cell debris before the cell lysates were added to TZM-bl cells infected with HIV IMCs. However, the cell lysates reduced cell viability

at concentrations of total protein exceeding 10 ug and thus cell lysates with concentrations five-times less than those used in the enzyme activity assay were added to the infection assay (data not shown).

The cell lysates of nine strains significantly decreased HIV infection by approximately 45 – 30 %, including the *G. piotii* strains (Figure 4.8 A). The remainder of the strains had little to no impact on infection, with the exception of GV 18 which increased infection by 30 %. GV 18 lacked *nanh3* and its cell lysate tested negative for enzyme activity, suggesting that the increase was not due to sialidase. Most of the supernatants did not have a significant effect on cell viability with concomitant decrease in HIV infection other than GV 17, indicating that the effect on infection for these strains was unlikely due to cytotoxicity (Figure 4.8 B). It is also unlikely that the sialidase activity of GP 74 and GP 94 was responsible for the decrease in HIV infection as the same effect was observed for strains negative for *nanH3* and enzyme activity. Overall, the results suggested that the impact on HIV infection was not due to sialidase activity but caused by an alternative factor, the identification of which was beyond the scope of the study.

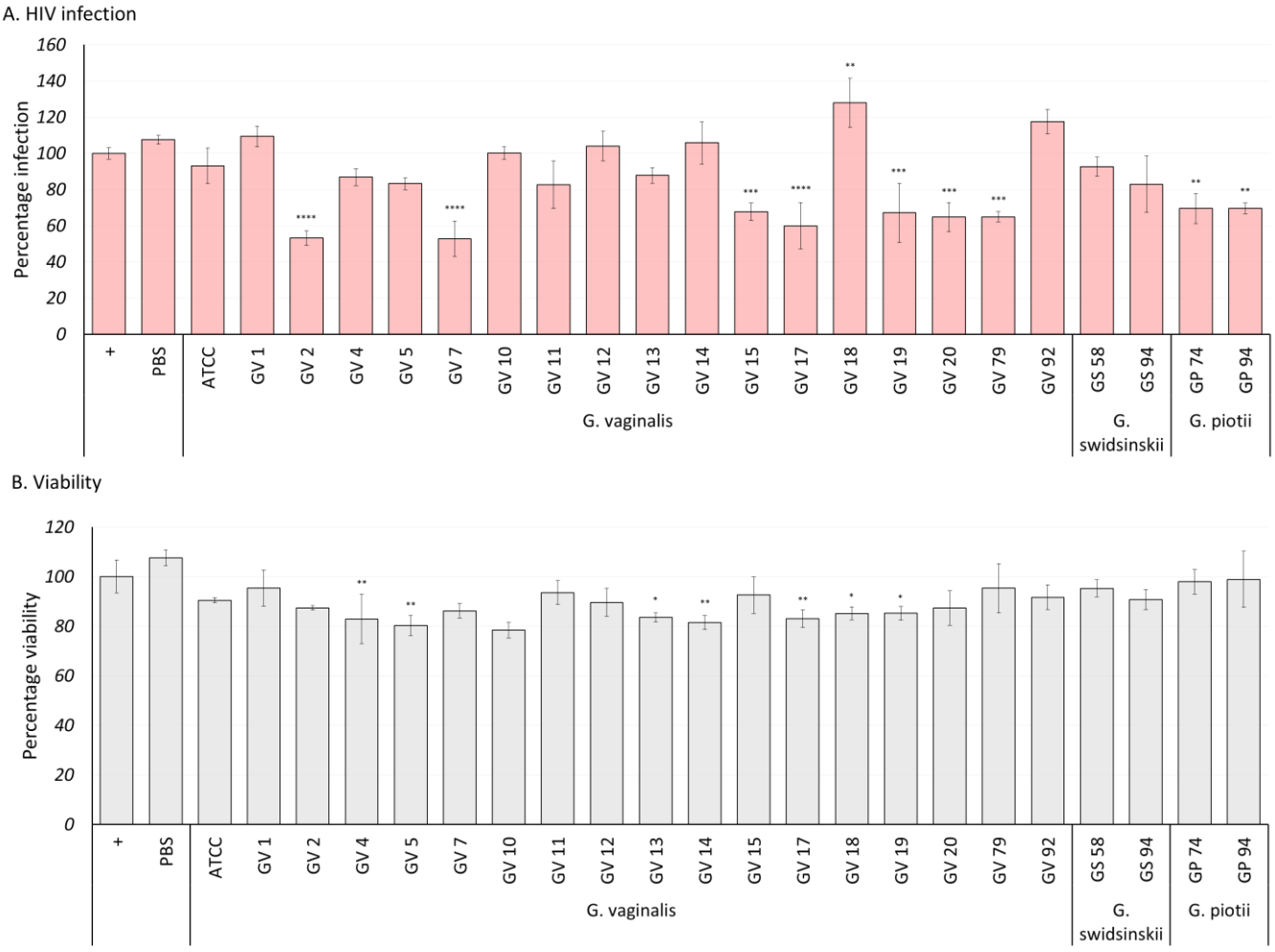


Figure 4.8. Effect of *Gardnerella* cell lysates on HIV infection and mammalian cell viability. *Gardnerella* strains were cultured in NYCIII medium as plankton for approximately 48 hours at 37 ° C under anaerobic conditions before sonication and centrifugation. The supernatants equivalent to 10 µg of total protein were added to the reporter cell line, TZM-bl, in the presence of 200 TCID50 units of pNL4 infectious molecular clones (IMCs) and A) infection and B) cell viability was measured after 48 hours. Infection was determined by luminescence and cell viability was determined by MTT assay. Infection and viability were calculated relative to the positive control (+) when no bacterial lysate was added. Error bars represent the standard deviation of the average of three biological repeats. Statistical analysis was performed relative to "+" using Graph Pad Prism with Dunnett's multiple comparisons test (*, < 0.05; ** <0.01; *** < 0.001; **** < 0.0001).

4.2.2.6 The effect of sonicated *Gardnerella* cells on HIV infection

Sialidase activity could be detected in sonicated *Gardnerella* cell pellets if the cell debris was not removed by centrifugation. To reduce the chance of cytotoxicity, a maximum of 10 µg total protein was added to the infection assay so that only 0.05 U and 0.015 U of GP 74 and GP 94 sialidase activity, respectively was added to the infected TZM-bl cells. The results showed that all strains, including a sialidase-negative strain, significantly increased HIV infection. This suggested that one or more cellular factors common to the three strains caused an increase in HIV infection. However, when a sialidase inhibitor was included, it was effective against only sialidase-producing strains, GP 74 and GP 94, and had no effect on GV 79 (Figure 4.9). Interestingly, despite GP 74 having 3-fold higher enzyme activity present compared to GP 94, they both increased HIV infection to the same extent. Perhaps the enzyme concentrations were too low to discern a differential impact on infection.

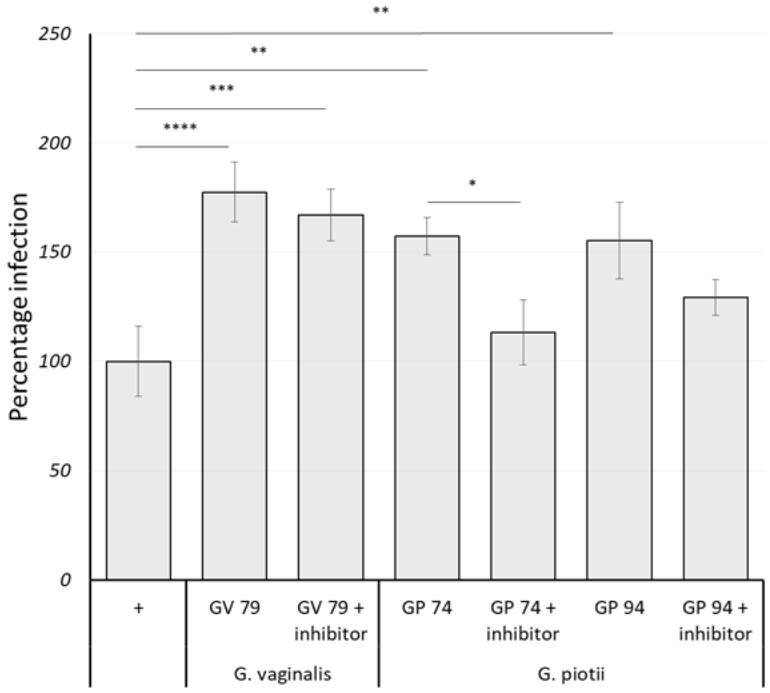


Figure 4.9: Effect of sialidase-negative and positive strains on HIV infection. GV 79, GP 74 and GP 94 cells were disrupted by sonication and the crude lysate equivalent to 10 µg of total protein was added to the reporter cell line, TZM-bl infected with 200 TCID50 IMC. Infection was monitored by luminescence and calculated relative to the positive control when crude lysate was not included. The sialidase inhibitor 2,3-dehydro-2-deoxy-N-acetylneuraminic acid (DANA) (30 mg/ml) was added to the infection assay. Bars represent the average of two biological repeats and error bars indicate standard deviation. Statistical analysis was performed relative to "+" using Graph Pad Prism with Tukey's multiple comparison test (*, < 0.05; ** <0.01; *** < 0.001; **** < 0.0001). GV 79 is an example of a *nanH3* and sialidase activity-negative strain whereas GP 74 and GP 94 produce NanH3.

The manner in which the cell lysates were prepared had a varying effect on HIV infection and cell viability. Sonicated and centrifuged GV 79 cells, that lacked the *nanh3* gene, significantly reduced HIV infection without reducing cell viability whereas if the cells were not centrifuged, GV 79 increased HIV infection. The extent to which the *G. piotii* strains increased infection was similar to GV 79, suggesting that other factors in the cell lysates were affecting HIV infection and potentially masking the ability of the enzyme to increase infection more robustly. The design of the assay might also limit the degree to which HIV infection is increased as IMCs could become more sialylated during replication so that the effect of desialylation on infection is lost over time.

4.3.1 The effect of *C. perfringens* sialidase on HIV infection

Despite extensive optimisation to prepare cell lysates with detectable sialidase activity, we were unable to conclusively distinguish the effect of sialidase from that of other cellular factors. Therefore, recombinant sialidase from *Clostridium perfringens* was added to the IMC infection assay to determine whether enzyme activity would affect *in vitro* infection. TZM-bl cells and IMCs were treated separately and together to determine whether desialylation of host cells, virus or both were required for an increase in HIV infection. The results showed that all treatment permutations increased infection by approximately two-fold. When a sialidase inhibitor was added, infection returned to the same level as the untreated positive control, confirming that the apparent increase was due to sialidase activity (Figure 4.10). Therefore, desialylation may contribute to a reduction in negative repulsion between the cell membrane and virus envelope and increase HIV infection.

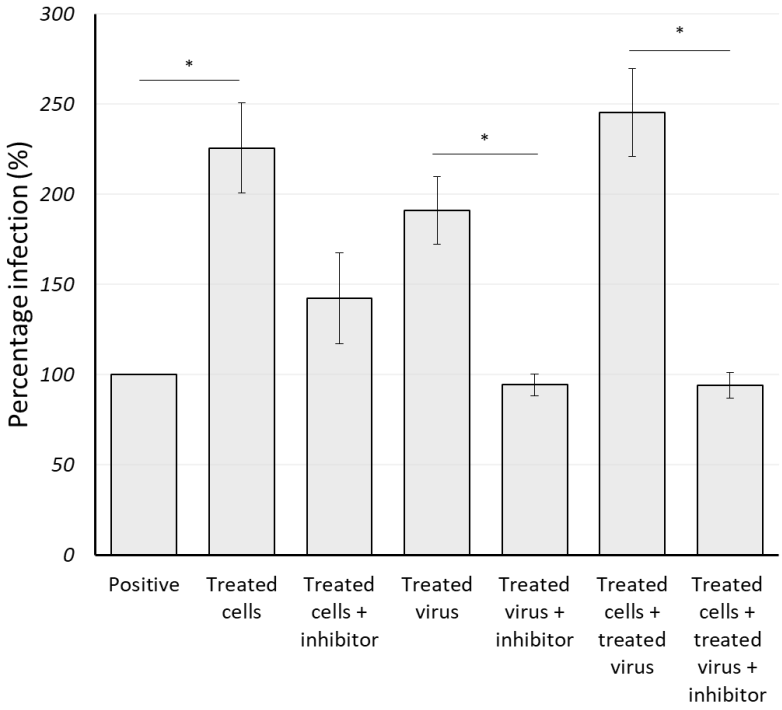


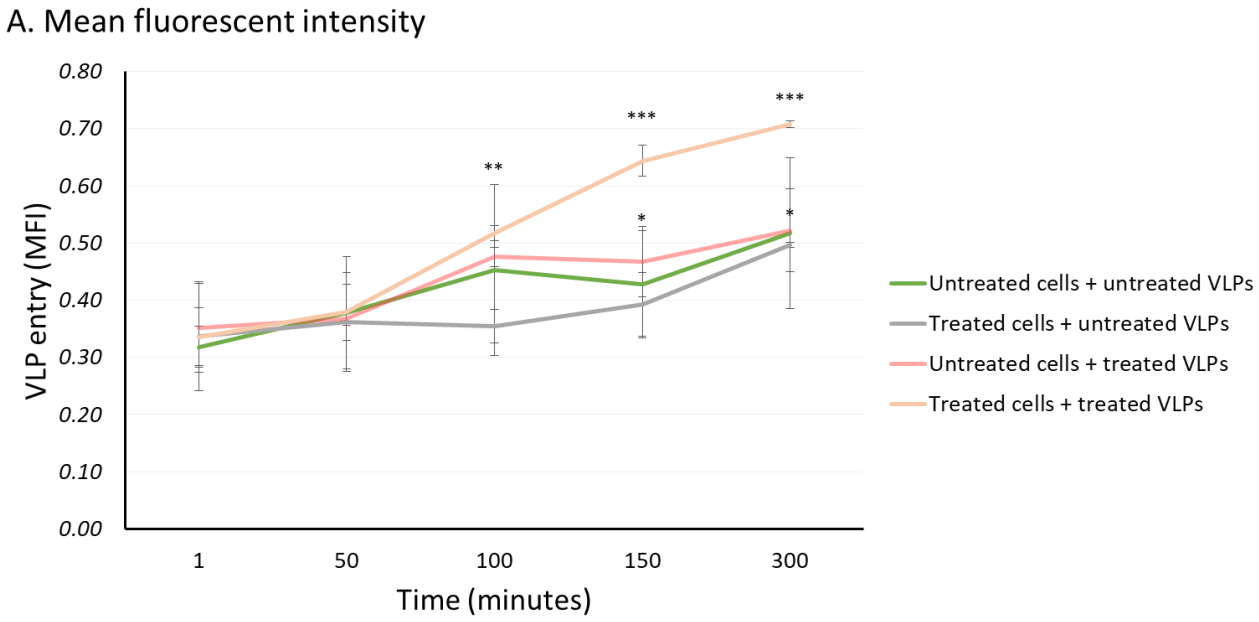
Figure 4.10: Effect of sialidase on HIV infection. TZM-bl cells were seeded in a 96-well plate at a density of 50 000 cells per well. The cells and HIV infectious molecular clones (IMCs) were treated with *C. perfringens* NanI (10 U) for 8 hours either separately or together before the reaction was inhibited by DANA and infection initiated by exposing the cells to virus. IMC replication continued for 48 hours and was monitored by luminescence (RLU). To confirm the activity of sialidase, NanI was first inhibited by DANA before addition to the cells and IMCs. Infection was normalized to the positive control untreated TZM-bl cells infected with untreated IMCs. Error bars indicate the standard deviation of the average of four biological repeats. Statistical analysis was performed using Graph Pad Prism with Dunnett's multiple comparisons test (*, < 0.05)

4.3.2 Effect of sialidase on viral-like particle entry of TZM-bl cells

Desialylation can act as a trigger for the rapid clearance of some affected glycoproteins although it is not known whether the half-life of all glycoconjugates is similarly regulated [170]. However, if true, this has ramifications on the study because desialylated moieties will be replaced by new glycoproteins with their full complement of sialic acids. Therefore, during the 48 hours of infection after desialylation, the negative charge associated with the cell surface could be fully regenerated, negating any effect sialidase treatment might have had on HIV infection. To limit the

potential impact of desialylation on the half-life of cell surface glycoproteins, TZM-bl cells were exposed to GFP-labelled viral-like particles (VLP), and entry was visualised for 300 min. The assay had the added advantage that sialylation of virus glycoproteins would not occur during replication.

TZM-bl cells and VLPs were treated separately and together with sialidase, and uptake was monitored by time-lapse fluorescence confocal microscopy (Zeiss LSM 980). It appeared as though fluorescent particles co-localised within the cell but without further analysis, the precise location cannot be confirmed (Figure 4.11 B). There was significantly higher mean fluorescence intensity (MFI) associated with cells when VLPs were treated alone and when VLPs and cells were treated together (Figure 4.11 A). It appeared as though uptake of VLPs was most efficient when both cells and viral particles were treated with sialidase as it was significantly higher than the untreated control after 100 minutes. This was contrary to what was observed for the infection assay as there was no apparent difference in IMC infection when cells and IMCs were treated together or separately. This could merely be due to differences between the two methods. Nonetheless, these findings suggest that sialidase can enhance virus-target cell interaction by desialylation.



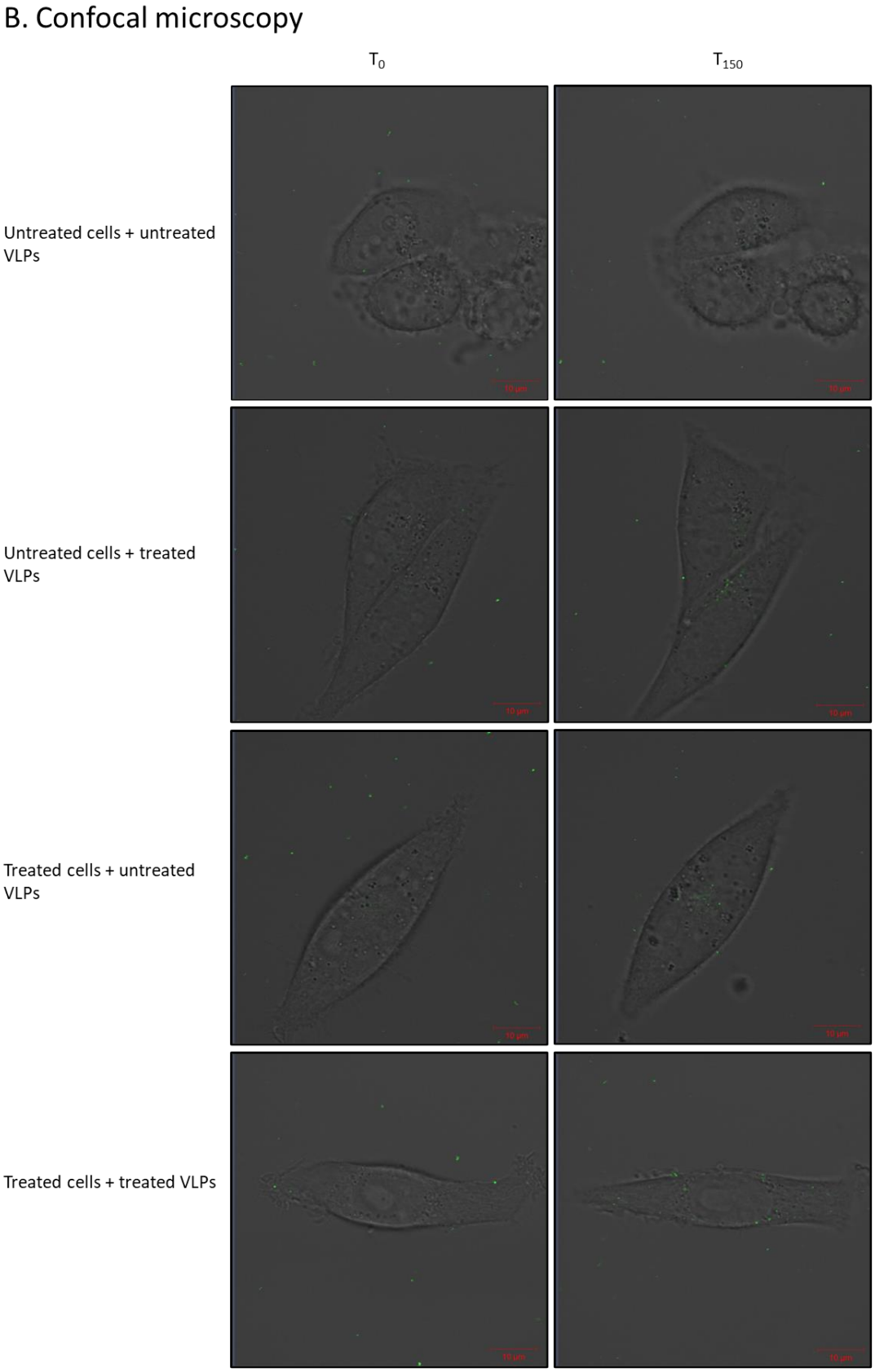


Figure 4.11: Effect of sialidase on virus-like particle entry of TZM-bl cells. TZM-bl cells and GFP-labelled virus-like particles (VLPs) were treated with *C. perfringens* sialidase (20 U) separately and together and fluorescence was measured by time-lapse confocal microscopy. A) Cell-associated relative fluorescent units (RFU) were determined at 1, 50, 100, 150 and 300 min and the error bars represent standard deviation of two biological repeats. Statistical analysis was performed using Graph Pad Prism, with two-way ANOVA relative to untreated control (*, < 0.05; ** <0.01; *** < 0.001). B) Representative images captured at 0 min and 150 min are shown to indicate internalization of fluorescently labelled VLPs when cells were untreated (untreated cells and untreated VLPs), treated VLPs, treated TZM-bls cells, and VLPs and cells treated together.

4.3 Discussion

Nugent/Amsel-BV increased the likelihood of HIV infection by 60 % [171] and Nugent-BV was associated with a three-fold increase in female-to-male transmission [103]. It is generally accepted that the association between BV and HIV acquisition is likely due to tissue damage, inflammation, increased proinflammatory cytokines and influx of HIV-permissive immune cells to the FRT [46,172,173]. However, it is also possible that BV could directly facilitate HIV infection by affecting the composition and characteristics of CVM [174] and the binding and entry of viral particles [35]. A great deal of attention has been focussed on *Gardnerella* spp. due to its consistent isolation from the FRT of BV-positive women and the suggestion that it is the most virulent of the BVAB [66]. One such virulence determinant is its sialidase activity which could favour the onset of BV by facilitating biofilm formation [141] as well as play a role in HIV infection by reducing the viscosity of CVM [157], desialylation of HIV Env [113] and/or diminish the negative charge associated with viral and host cell surface components to aid virus attachment [112,114]. This study characterised the genotype and phenotype of 20 clinical *Gardnerella* strains to investigate the role of *Gardnerella* sialidase in HIV acquisition.

Similar to an earlier study, the presence of the three homologs of *Gardnerella* sialidase was associated with specific species [163] with all *G. piotii* and *G. vaginalis* strains positive for *nanh1,* the two *G. piotii* strains positive for *nanh3,* and *nanh2* not present in the genome of any strains. It was recently reported that *nanh2* was rarely found without *nanh3* and since our sample set only contained two *nanh3*-positive strains, biased sampling might have reduced our chances of detecting *nanh2* [10,163]. Quantitative PCR analysis of these genes revealed that expression varied widely between strains when bacteria were grown as biofilms and plankton, suggesting that *Gardnerella* strains may regulate *nanh1* and *nanh2* expression based on its growth mode. It has been shown that *Gardnerella* regulated gene expression depending on whether isolates were grown as biofilms or plankton [123]. There was no association between biofilm formation and *nanh1* expression. Moreover, *G. swidsinskii* strains, negative for *nanh1, nanh2* and *nanh3,* produced biofilms if not equally, then better than other strains (Figure 3.6). The data suggested that sialidase activity is not a determining factor in biofilm formation. However, the addition of *C.*

perfringens sialidase to cultures of strains grown as biofilms might have provided more insight into understanding the role of sialidase in biofilm formation.

NanH1 is considered largely inactive whereas NanH3 is associated with sialidase activity [1]. The first detectable sialidase activity assay was carried out on whole cells and only the *nanh3*-positive *G. piotii* strains tested positive, suggesting the enzyme was cell-associated. Sequence analysis confirmed that NanH3 lacked a signal peptide and would thus not be transported to the cell surface for secretion into the culture medium, explaining why the cell-free culture medium was negative for sialidase activity. This was contrary to a recent review that reported NanH3 was located on the surface of the bacteria and shed into the culture medium following cell death or proteolytic cleavage [95]. The presence of a TMD would support the tethering of NanH3 to the cell membrane, but in the absence of a signal peptide, it is possible that the protein is not correctly processed, or forms aggregates due to hydrophobic interactions between the TMD of proteins. The unexpected localisation of the enzyme most likely explains why enzyme activity was only detected in the soluble fraction when detergent was included during sonication or sonicated cells were not centrifuged.

When the sonicated cells of the *G. piotii* strains were added to TZM-bl cells infected with IMCs, infection increased. Furthermore, when a sialidase inhibitor was added, infection was reduced. However, unexpectedly, a NanH3-negative *Gardnerella* strain, GV 79, increased HIV infection to the same extent as the *G. piotii* strains, although the inhibitor did not have any noticeable effect on the NanH3-negative strain. Therefore, the findings support the hypothesis that sialidase increases HIV infection, but there seems to be an additional factor in the cell lysates that is also enhancing infection.

The complex composition of the cell lysate, the challenge to detect sialidase activity and the inability to add detergent-rich fractions to the HIV infection assay, limited the ability to determine conclusively whether *G. piotii* NanH3 increased HIV infection. Therefore, commercially available *C. perfringens,* NanI, was used to demonstrate the effect of sialidase on HIV infection. When TZM-bl cells and virus were treated separately and together with sialidase, HIV infection increased.

When the assay was repeated in the presence of a sialidase inhibitor, infection was reduced to the same level as the positive control. This confirmed that sialidase increased HIV infection most likely by reducing the negative repulsive force between the virus particles and the TZM-bl cells.

IMC replication during the 48 hours of incubation could result in the re-sialylation of the cells and the virus. To circumvent this potentially confounding effect, GFP-labelled Gag VLPs with R5-tropic Env were added to TZM-bl cells and entry was monitored by time-lapse microscopy. Upon entry, the VLPs seemed to accumulate around the nucleus, as previously reported [175]. After 50 minutes, desialylated VLP particles were able to enter TZM-bl cells treated with sialidase more efficiently than when the viral particles and target cells were treated separately. This result did not align with our earlier finding that the separate or simultaneous desialylation of IMCs and TZM-bl cells had the same effect on infection because the increase in VLP entry did not require desialylation of the cells. This suggests that either the loss of surface negative charge of the virus was sufficient to overcome the charge repulsion between the VLPs and TZM-bl cells or removal of sialic acids from Env could enhance its binding with CD4 and CCR5 receptors. IMCs used for this study are X4-tropic meaning that Env binds to CXCR4 and not CCR5. Env interactions with CCR5 and CXCR4 can affect the efficiency at which HIV is able to infect $CD4^+$ T cells [176]. Studies have shown that VLPs may enter cells by different mechanisms, such as receptor-mediated endocytosis[177], however, investigation of the mechanism of entry was beyond the scope of this study. Furthermore, culture of TZM-bl cells during IMC infection was 48 hours whereas VLP entry was monitored for only 5 hours. This could have affected to what extent cell surface glycoproteins had regained their negative charge. Despite the multiple factors that might be responsible for the differential impact of sialidase on VLP entry and IMC infection, both assays indicated that NanI increased HIV infection.

In conclusion, NanH1 expression was highly variable across *Gardnerella* clinical isolates and was not associated with biofilm formation. Only the two *G. piotii* strains expressed *nanh3* and had cell-associated enzyme activity, confirming that the homologue is responsible for the sialidase activity associated with *Gardnerella* spp. Lastly, this study showed that NanI increased HIV

infection *in vitro* which suggests that NanH3 could play a similar role in the FRT to enhance the risk of HIV acquisition.

Chapter 5: Purification of *Gardnerella piotii* NanH3

5.1 Introduction

Sialidase was first described by Hirst [178] who discovered the influenza virus enzyme [178]. Since then, sialidase activity has been associated with many bacteria, viruses and mammals [179], including pathogens such as *Streptococcus pneumoniae, Vibrio cholerae* and *C. perfringens* [180]. Due to its consistent association with a number of diseases such as gas gangrene (*Clostridium*); septicaemia (*Streptococcus, Pneumococcus, Bacteriodes, Corynebacterium*); pneumonia (*Streptococcus*); peritonitis (*Clostridium*, *Bacteriodes*); meningitis (*Streptococcus B*) and cholera (*Vibrio cholerae*) [181] and its multiple functions in pathogen adherence, infection processes, biofilm formation, nutrition and immune escape [182], it is considered an important virulence factor [183]. Removal of sialic acids from host cell surfaces can also affect cellular signalling pathways and host-pathogen interactions [184,185].

Bacteria usually produce multiple sialidases with *C. perfringens* (NanH, NanI, NanJ) and *S. pneumoniae* (NanA, NanB and NanC) both encoding three homologs with different substrate specificities, localisation and function [182,186]. The substrate specificity of sialidases is dependent on the type of glycan rather than the structure of the protein component of glycoproteins [186] and the enzymes are able to distinguish between Neu5Ac and Neu5Gc as well as the linkage to the next residue; 2-3, 2-6 or 2-8 [187]. NanI preferentially cleaved glycosidic bonds in the order of α-2,3 > α-2,6 > α-2,8 linkages, NanJ preferred α-2,6 > α-2,8 > α-2,3 sialic acid linkages and NanH activities showed a preference for α-2,8 > α-2,3 > α-2,6 linkages, indicating that sialidases can act on a number of different substrates, albeit on some better than others [188]. However, when *Gardnerella* sialidases, NanH2 and NanH3 were incubated with 3'-sialyllactose and 6'-sialyllactose, there was no preference for one particular type of bond [1]. The authors then compared their specificity for N- and O-linked glycans, by using IgA and bovine submaxillary gland, respectively. These experiments revealed that both NanH2 and NanH3 cleaved N- and O-linked glycans effectively. However, when 7-O and 9-O cleavage was compared, NanH2 was more effective at cleaving 9-O-acetylated sialic acids compared to NanH3 [1]. NanH2 and NanH3 from

Gardnerella were able to cleave sialic acids from all substrates tested which demonstrates the likelihood that they are capable of acting on substrates relevant to the vaginal mucosa [1].

Sialidases have been grouped according to their size [189] with the molecular weight of small enzymes in the region of 40 kDa and large homologs approximately 80 kDa due to the presence of additional carbohydrate-binding domains [190]. The NanH and NanI sialidases of *C. perfringens* are 43 kDa and 77 kDa, respectively, with only the latter secreted [191,192].

Robinson*, et al.* [1] reported that NanH3, was a 721 amino acid protein, lacking a signal peptide, contrary to another study which showed that the sialidase comprised of 812 amino acids including a signal peptide [10]. Therefore, some bacteria would secrete NanH3 while the sialidase of others would be cell-associated. This could have ramifications on the function of the enzyme. This study examined the impact of recombinant *Gardnerella* NanH3 activity on HIV infection.

5.2 Results

5.2.1 Cloning of NanH3

Analysis of whole genome sequences of GP 74 and GP 94 indicated that the poly(C) tract comprised of ten cytosines, which introduced premature stop codons that prevented translation of sialidase with a signal peptide (Figure 5.1). The alternative start codon of NanH3 was identified as GUG which coded for a valine. However, when the transcription machinery identifies GUG at the start of an ORF, as opposed to within a coding sequence, tRNAfMet recognises the codon and thus the N-terminal amino acid of all functional polypeptides begins with methionine. [193]. When GP 74 NanH3 was cloned, GUG was kept as the start codon as methionine would still be inserted at position 1 and thus not affect structure and function of the enzyme.

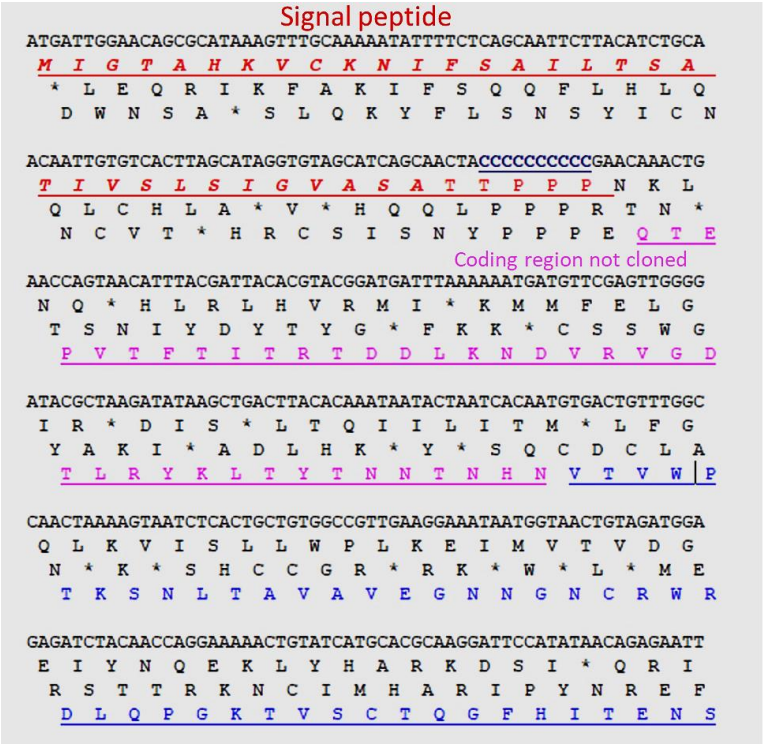


Figure 5.1 N-terminal sequence of GP 74 NanH3 clone: The 5' sequence of GP 74 was translated into three ORFs using DNA sequence analysis software, DNAMAN. The signal peptide and proline residues encoded by the cytosine polymer, poly(C) is underlined in red. The sequence underlined in blue represents the N-terminal end of GP 74 NanH3 with start codon, GUG. The sequence underlined in pink indicates the residues not translated and thus not cloned. Due to the length of the poly(C) tract, the ORF of NanH3 switches from ORF 1 to ORF 3. Asterisks indicate stop codons.

The *nanh3* gene was cloned into pProEx-HTa in-frame with a His-tag at the N-terminus and expressed in *E. coli* (Figure 5.2). The TMD was excluded from the clone as previous research demonstrated that purification of the gene without the TMD improved protein purification yields [1]. After transformation of *E. coli*, colony PCR identified positive clones (Figure S5.1) and restriction enzyme digestion confirmed the presence of the 2 000 bp insert (Figure S5.2). Sequencing of the clone and comparison to the GP 74 genomic *nanh3* sequence confirmed the absence of spurious mutations (data not shown). Alignment of the cloned sequence with JCP8151B NanH3 from Robinson*, et al.* [1] indicated 31 synonymous point mutations and two insertions of ten and six amino acid residues in clone GP 74 (Figure S4.1). The putative amino acids involved in catalysis were conserved, suggesting similar enzyme specificity [1].

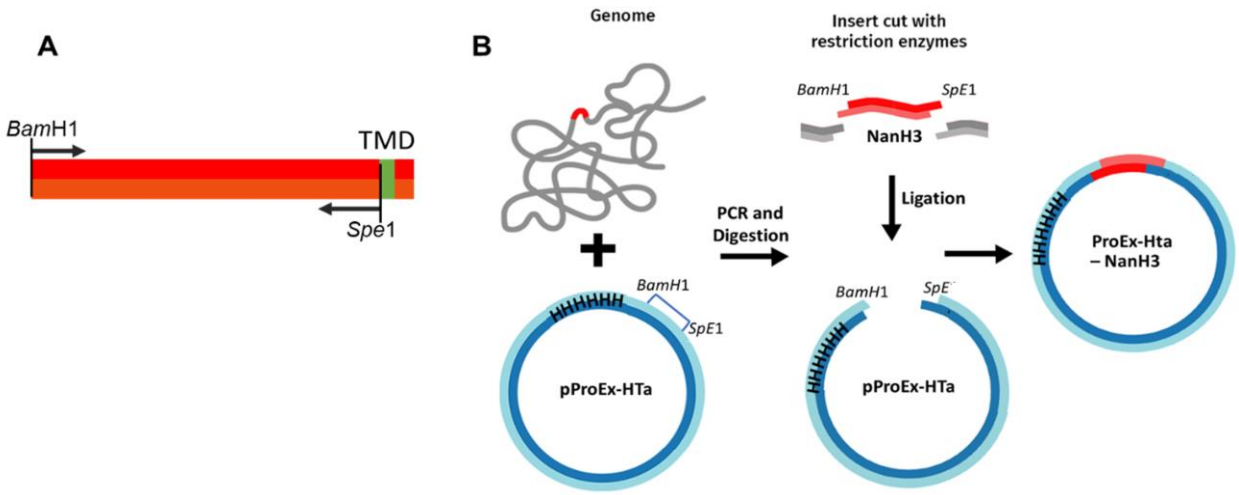


Figure 5.2. Cloning of NanH3. A) Primers were designed flanking *nanH3* with 5' extensions carrying the restriction enzyme sites for *BamH*I and *Spe*I. The reverse primer was complimentary to the sequence immediately upstream of the transmembrane domain (TMD) and the forward primer began at the start codon, GUG (valine) at the 5' end of the gene. B) Genomic DNA was extracted from the culture of GP 74, *nanH3* amplified and pProEx and the PCR product digested with the restriction enzymes. After ligation, the His-tag was in frame with NanH3 and the recombinant plasmid, pPROEX-HTa-NanH3, was used to transform *E. coli* cells.

5.2.2 Purification of NanH3

BL21 cells, engineered for high protein expression, were transformed with pProEx-HTa-NanH3 and NanH3 expression was induced by isopropyl β-D-thiogalactopyranoside (IPTG). The cell lysates were applied to a Nickel affinity column and the collected fractions were tested for total protein, enzyme activity and separated by SDS-PAGE for subsequent Coomassie staining and Western blotting. After optimisation, a band was detected that corresponded to the predicted molecular weight of 81 kDa [194] and Western blotting with a His-tag antibody confirmed that the band represented recombinant NanH3 (Figure 5.3). However, a prominent band, suspected to be BL21-associated, was detected at ~ 26 kDa. To ensure that this did not represent a truncated version of NanH3, an untransformed BL21 cell lysate sample was applied to the Ni^{2+} column. Subsequent Coomassie stained SDS-PAGE confirmed the presence of the 26 kDa protein and an enzyme assay confirmed that the eluted fractions were not positive for sialidase activity (Figure S5.3).

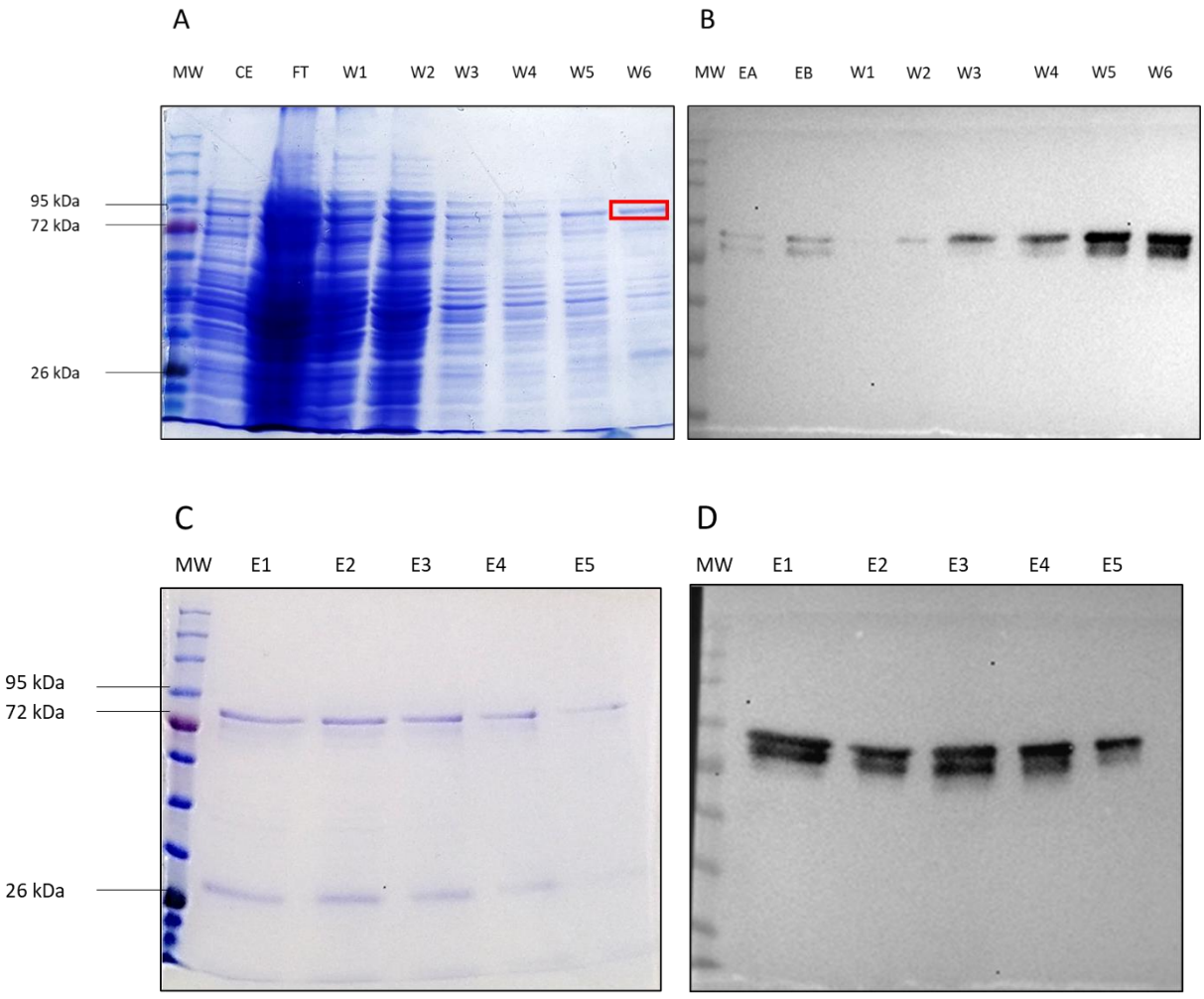


Figure 5.3. Affinity purification of His-tagged NanH3 Coomassie stain (A and C) and Western blot (B and D) of fractions collected from the Nickel affinity column. The fractions from the Ni^{2+} affinity column were separated by SDS-PAGE and used for A and C) Coomassie staining and B and D) Western blotting with an anti-His-tag antibody. The lanes are labelled: crude extract (CE), flow-through (FT), sequential wash steps 1-6 (W2-W6) and elution fractions (E1 – E5). EA and EB represent eluted fractions from previous purification attempts. The band corresponding to the predicted molecular weight of NanH3 is highlighted by a red box. Contaminating protein of ~ 26 kDa is indicated. MW – molecular weight marker.

SDS PAGE and Western blotting showed that although recombinant NanH3 eluted from the column during the wash steps, most eluted at higher imidazole concentrations (Figure 5.3) which corresponded to a slight increase in sialidase activity (Figure 5.4). However, although high levels of sialidase activity was detected in the flow-through and wash steps 1 and 2, Western Blotting did not detect high concentrations of NanH3 in these fractions. It has been shown that NanI was

cleaved within the N-terminus which, in this case, would have resulted in the loss of the His-tag [195]. A band, corresponding to a molecular weight just below 80 kDa, was consistently found associated with the larger band in previous purification attempts. Similar to the larger band, it was not detected in the wash fractions 1-3 despite exceedingly high enzyme activity (Figure 5.3). The appearance of a smaller band suggested that NanH3 might be proteolytically cleaved at more than one site.

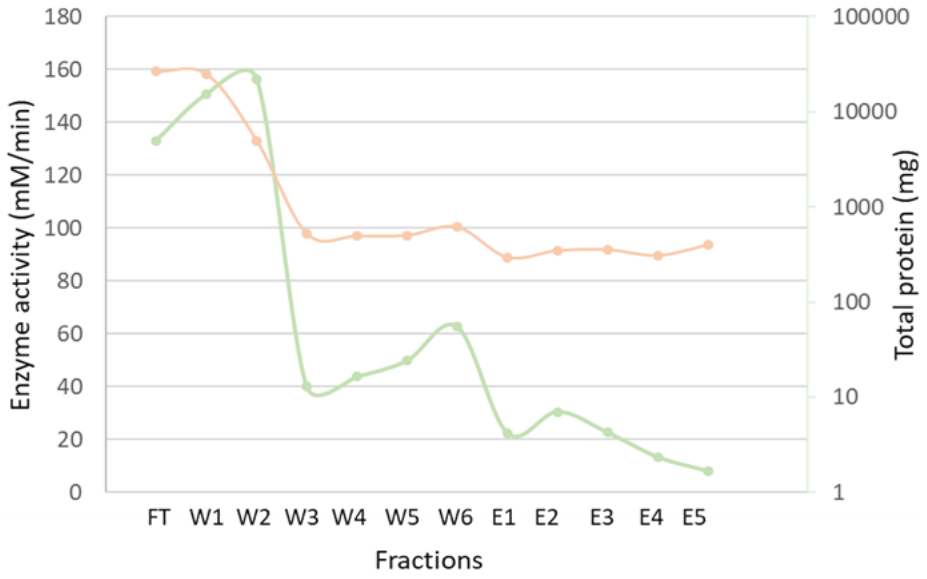


Figure 5.4. Elution profile of Total Protein and NanH3 sialidase activity Total protein (green) and sialidase activity (pink) was determined for each fraction: flow-through (FT), sequential wash steps 1-6 (W1-W6) and elution fractions (E1 – E5).

5.2.2 Mass spectrometry analysis

The doublet coinciding with approximately 80-75 kDa was excised from the Coomassie stained gel, sliced into 1 cm x 1 cm sections, dehydrated and transported to the mass spectrometry (MS) facility for analysis. The results obtained showed that 91 % of the sample identified as an 81 kDa exo-alpha-sialidase from *G. vaginalis*. None of the other identified proteins were associated with sialidase activity (Table 5.1). This confirmed that the protein of interest was successfully purified, and that the contaminants present were unlikely to be responsible for the sialidase activity

detected. Furthermore, the relative abundance indicated that both bands comprising the doublet, represented NanH3 and the smaller band was likely the result of proteolytic cleavage.

Table 5.1: Mass spectrometry results of recombinant NanH3.

#	Identified Proteins	Accession Number	MW (kDa)	Quantity relative to total spectra	Probability of correct protein identified
1	Cluster of Exo-alpha-sialidase [a]*Gardnerella vaginalis*	I4M8F1_GARVA	81	660	100
2	Cluster of Uncharacterized protein; *Gardnerella vaginalis*	I4LZR2_GARVA	353	12	100
3	Cluster of Repeat protein; *Gardnerella vaginalis*	A0A3D8TNF4_GARVA	338	15	100
4	Cluster of Long Rib domain-containing protein *Gardnerella vaginalis* JCP8151B	S4GNI3_GARVA	194	6	100
5	Cluster of Heavy metal translocating P-type ATPase (Fragment); Bifidobacteriaceae bacterium	A0A3E2CUP8_9BIFI-DECOY	?	6	100
6	Glycosyltransferase RgtA/B/C/D-like domain-containing protein; *Gardnerella vaginalis* JCP8108	S4GQA7_GARVA	67 kDa	4	100
7	Uncharacterized protein; *Gardnerella vaginalis*	D6SYR9_GARVA	356	5	100
8	Cluster of alpha-amylase; Bifidobacteriaceae bacterium	A0A3E2CT63_9BIFI	46	6	100
9	MaoC-like protein, *Gardnerella vaginalis*	A0A133NRF1_GARVA	343	7	100
10	Peptidase; *Gardnerella vaginalis* OX=2702	A0A3D8TPG9_GARVA	124	5	100

[a] Although proteins were identified as *G. vaginalis*, it is likely that the database has not yet been updated with the most recent species names i.e. *G piotii*.

5.2.3 Enzyme kinetics

The Michaelis constant (K_m) of purified, recombinant NanH3 and *C. perfringens,* NanI, was determined using MUN as a substrate (Figure 5.5). The analysis indicated that the K_m of NanH3 was between 0.21 – 0.27 mM while the K_m for the *C. perfringens* sialidase was 0.17 mM. The calculated K_m of *C. perfringens* was similar to earlier publications which reported K_m values of 0.19 mM and 0.165 mM [191,196]. Therefore, NanI could have marginally higher binding affinity for MUN, compared to NanH3. Overall, the results demonstrated that translation from the alternative start codon and truncation of the TMD resulted in a functional enzyme.

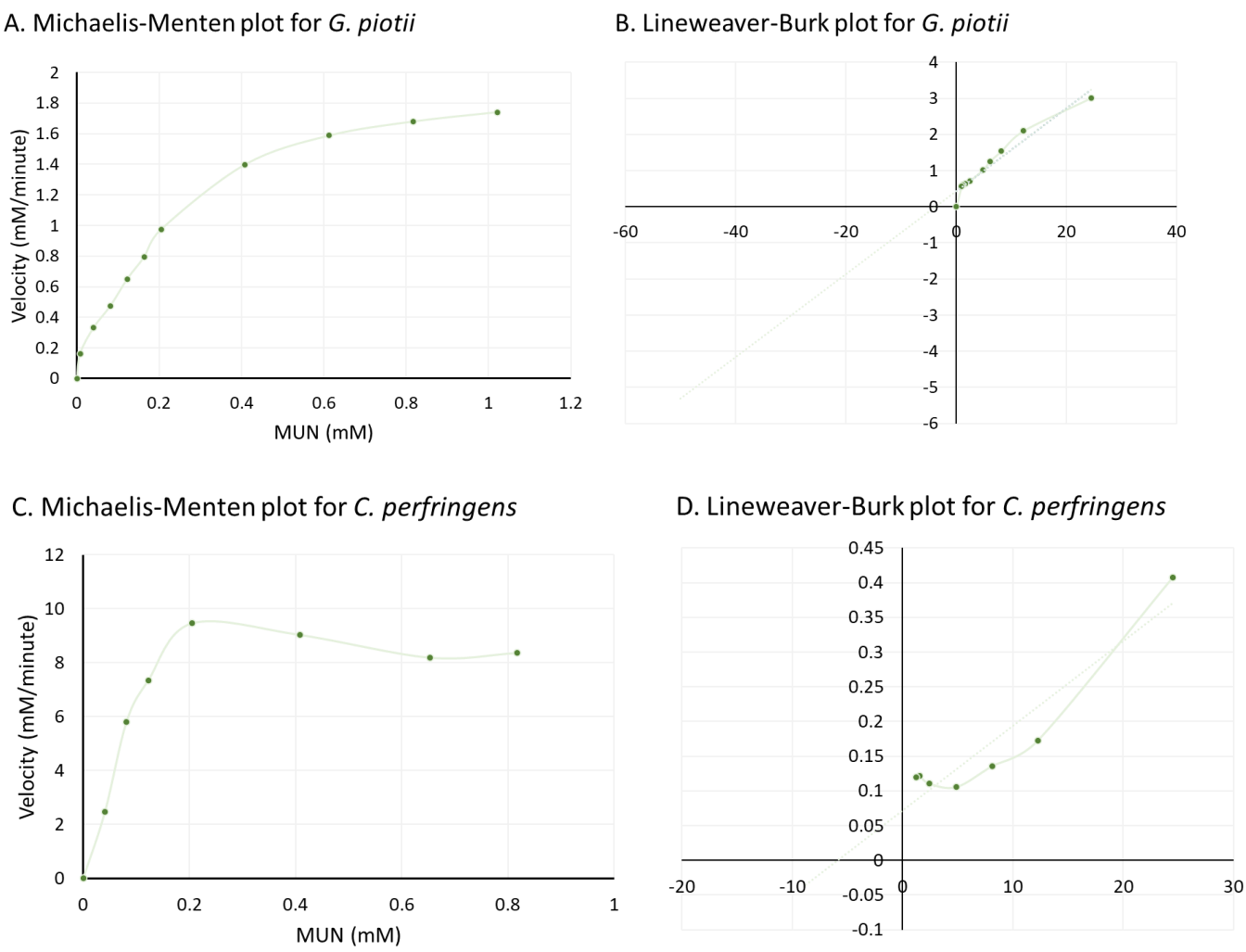


Figure 5.5: Michaelis-Menten and Lineweaver-Burk plots for *G. piotii* and *C. perfringens.* A) Michaelis-Menten plot for *G. piotii* sialidase (0.25 U) B) Lineweaver-Burk plot for *G. piotii* sialidase C) Michaelis-Menten plot for *C. perfringens* sialidase (10 U). D) Lineweaver-Burk plot for *C. perfringens* sialidase. Initial reaction velocity of recombinant NanH3 and *C. perfringens* NanI was determined for different concentrations of MUN substrate to determine the K_m of each enzyme.

5.2.4 Predicted structure

Bacterial sialidases have variable length, a range of molecular weights (~ 40–115 kDa), and low sequence identity. Despite high diversity, there are two conserved motifs: RIP (Arg-Ile-Pro), and multiple 'Asp-box' (most commonly Ser/Thr-X-Asp-[X]-Gly-X-Thr-Trp/Phe) sequences [197]. Submission of the NanH3 sequence to the protein homology/analogy recognition engine (PHYRE V2.0) [8] identified a predicted three-dimensional model with the best coverage (58 % of sequence) and with the highest identity (42 %) to *M. viridifaciens* sialidase [198]. Alignment of GP 74 with *M. viridifaciens* sialidase indicated low sequence homology despite structural similarity. GP 74 residues 106-563 aligned with *M. viridifaciens* sequence from position 59 to 490 which corresponded to the six-bladed-propeller catalytic sialidase domain and part of the β-sandwich linker domain (Figure 5.6 A and B). GP 74 did not comprise a galactose-binding domain and further analysis of the C-terminus of NanH3 did not yield any other structural similarity to other functional domains. Four of the five Asp boxes located in the β-propeller region of *M. viridifaciens* were conserved in NanH3 but the fifth did not have the consensus motif as the Thr was replaced by a Ser in the second last position (Figure S5.4). NanH3 was less similar to NanI, with 32 % identity. The solved structure of *C. perfringens* NanI comprised of a β-propeller domain and a small-barrel domain [199] (Figure 5.6 F) PHYRE analysis indicated that NanH3 structural prediction using NanI as a template, did not have a structure similar to the small-barrel domain but instead seemed to carry a structure more similar to *M. viridifaciens* β-sandwich linker domain. All predicted structures for NanH3 comprised of a six-bladed-propeller catalytic sialidase domain, confirming its enzymatic activity.

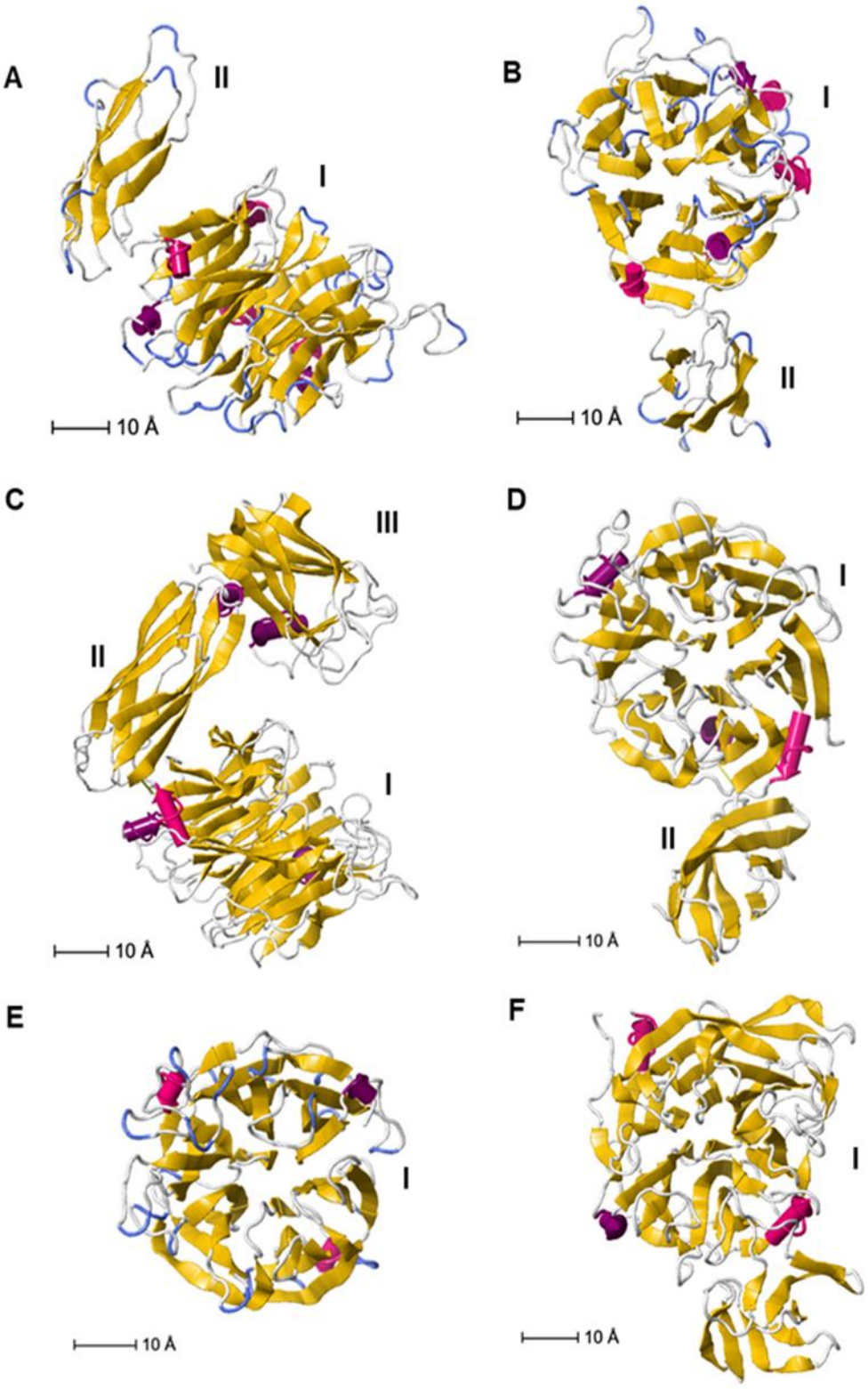


Figure 5.6 Predicted structure of NanH3: GP 74 protein sequence was submitted to the protein homology/analogy recognition engine (PHYRE V2.0) and structures were predicted using the solved structures for *M. viridifaciens* sialidase (PDB 2BER) and *C. perfringens* NanI (PDB 2VK7) as templates. *M. viridifaciens* sialidase comprises of a β-propeller domain (I); a β-sandwich linker domain (II); and a β-sandwich galactose-binding domain (III) A) Side-view of NanH3 with *M. viridifaciens* sialidase as template; B) Top-view of NanH3 with *M. viridifaciens* sialidase as template; C) Side-view of *M. viridifaciens* sialidase; D) Top-view of *M. viridifaciens* sialidase with galactose binding domain hidden (III); E) Top-view of NanH3 with *C. perfringens* NanI as a template and F) Top-view of *C. perfringens* NanI.

5.2.5 *Gardnerella* sialidase effect on HIV infection

Since the *C. perfringens* sialidase increased HIV infection (Figure 4.10), recombinant NanH3 was used to treat TZM-bl cells separately and together with IMCs. A standard curve of *C. perfringens* sialidase activity was used to determine the concentration of purified sialidase so that a total of 10 U was kept constant for both enzymes. The results indicated that *Gardnerella* sialidase was able to significantly increase HIV infection by approximately 1.5-fold, presumably by acting on the target cells (Figure 5.7). However, unlike NanI, there was no increase seen in HIV infection when the virus was treated separately, and when cells and virus were treated together. The latter result was unexpected because HIV infection should have increased due to the action of NanH3 on cells alone. Perhaps NanH3 modified the N-glycan structures required for Env function so that HIV infection was reduced, thus counteracting the positive effect of sialidase treatment of TZM-bl cells. Irrespective of the mechanism, NanH3 sialidase activity increased HIV infection as the effect was abrogated by the inhibitor. Furthermore, there was no increase in infection in the presence of an eluted fraction from when BL21 cell lysates were applied to the N^{2+} column, (Figure S5.3) suggesting that contaminants co-purified with recombinant NanH3 did not affect HIV infection.

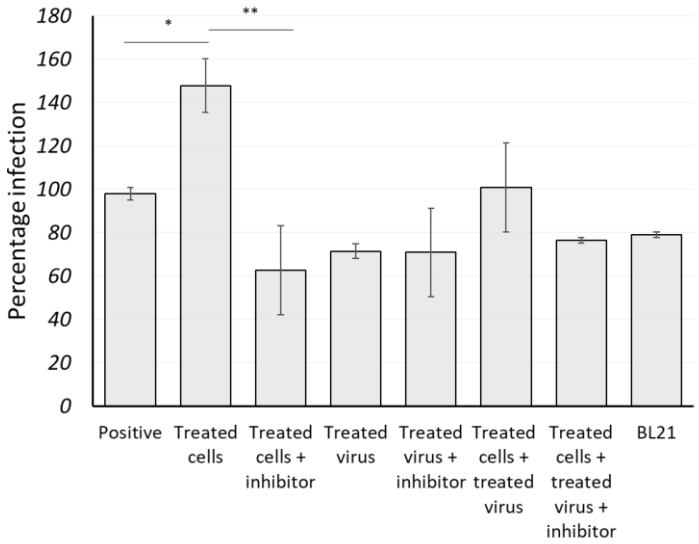


Figure 5.7: Effect of purified Gardnerella sialidase on HIV infection. Target cells, virus or both were treated with 10 U of purified *Gardnerella* sialidase prior to the HIV infection assay. An inhibitor of sialidase activity, 2,3-dehydro-2-deoxy-N-acetylneuraminic acid (DANA) (30 mg/ml) was added to each treated sample. A fraction of BL21 cell lysates eluted from the N^{2+} affinity column was also added to the infection assay. Error bars represent the standard deviation of two biological repeats. Relative luminescence units (RLU) were normalized to the positive control when cells and virus were untreated. Statistical analysis was performed relative to the positive control in Graph Pad Prism using Tukey's multiple comparisons test (* < 0.05; ** <0.01; *** < 0.001; **** < 0.0001).

5.2.6 *Gardnerella* sialidase effect on VLP entry

To establish whether NanH3 would have a similar effect on VLP uptake by TZM-bl cells as it had on IMC infection, TZM-bl cells were treated with recombinant NanH3 prior to the addition of VLPs. Despite technical error, VLP uptake was apparently higher when TZM-bl cells were treated with NanH3, confirming the results obtained by the infection assay (Figure 5.8).

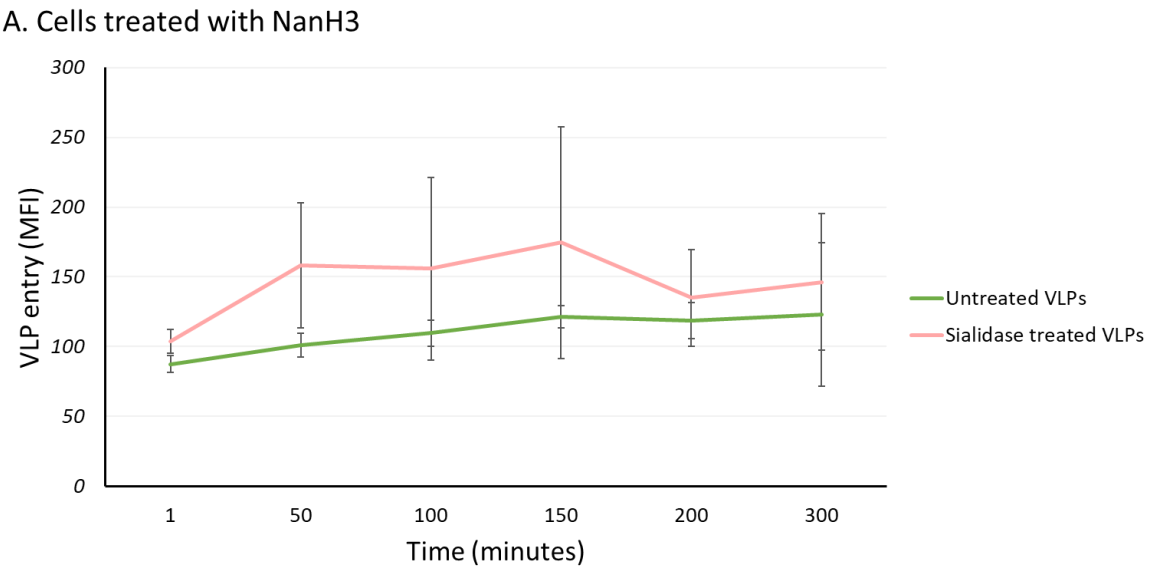


Figure 5.8: Effect of *NanH3* on virus-like particle uptake by TZM-bl cells. TZM-bl cells were treated with recombinant NanH3 (20 U) prior to the addition of GFP-labelled virus-like particles (VLPs) and uptake was measured by time-lapse confocal microscopy. Cell-associated mean fluorescent units (MFI) were determined at 1, 50, 100, 150 and 300 min and the error bars represent standard deviation of three technical repeats from a single experiment. Untreated TZM-bl cells are indicated in dark green, NanH3-treated TZM-bl cells are shown in pink.

5.3 Discussion

Sialidases are characterised by the presence of a six-bladed β-propeller catalytic domain [181] responsible for cleavage at α-2,3, α-2,6 and α-2,8 linkages of sialic acid residues of glycoconjugates. Some enzymes have additional domains that play a role in substrate binding and recognition [200]. Bacteria usually produce multiple sialidase isozymes, and enzymes have low protein sequence homology (<30 %). Although enzymes share structural features, they vary in biochemical properties, such as linkage and substrate specificities [183] and sensitivity to pH [188].

To determine whether the activity of GP 74 NanH3 was similar to *C. perfringens* NanI, the sialidase was expressed in BL21 cells and purified by His-tag affinity chromatography. Recombinant NanH3 was successfully enriched, although SDS-PAGE indicated the presence of *E.coli* contaminants. Notably, a 26 kDa protein that, despite low levels of imidazole included in the wash buffer, consistently co-eluted with NanH3. The contaminant was also present when BL21 cell lysates were applied to the N^{2+} column. A number of *E.coli* proteins are known to bind with high specificity to the N^{2+} column, such as Peptidoylproline cis–trans isomerase (SlyD) [201] with a predicted molecular weight of 26 kDa [202]. Although GP 74 NanH3 was enriched by approximately 30-fold under these conditions, most of the enzyme activity was lost in the flow-through and during washing. This was likely due to non-specific proteolysis which removed the His-tag without affecting enzyme activity. The loss of the His-tag would prevent binding to the N^{2+} and detection by Western blotting.

NanH3 appeared as a doublet comprising bands of approximately 81 and 75 kDa by Western blotting, suggesting that if truncated, additional cleavage must have occurred at the C-terminus to result in an intact His-tag. Sialidases have been shown to be cleaved at the C-terminus [183]. The *M. viridifaciens* recombinant sialidase was cleaved into three different isoforms when expressed in *S. lividans*, suggesting that expression by alternative bacterial species could impact processing of the sialidase [203]. Alternatively, NanH3 could present as a doublet due to post-translational modification, such as glycosylation or phosphorylation [204,205] although this is unlikely as the predicted size of the larger band corresponded to the molecular weight of recombinant NanH3.

Furthermore, as far as we are aware, there have been no reports on the phosphorylation or glycosylation of bacterial sialidase. Irrespective of the appearance of two bands corresponding to NanH3, MS confirmed that a glycoside hydrolase had been successfully purified.

Structural prediction of NanH3 indicated the highest similarity to *M. viridifaciens* sialidase, confirming that the GP 74 clone encoded for an enzyme with sialic acid-hydrolysing activity. However, proteolytic truncation of NanH3 could disrupt its tertiary structure and hence, function. The three isoforms of recombinant *M. viridifaciens* sialidase had molecular weights of 68 kDa, 52 kDa and 41 kDa, with the latter two, products of proteolytic digestion of the larger enzyme. All three had the same catalytic activity, indicating that the 41 kDa form contained the active domain and that proteolysis did not affect activity [203]. Given the similarity to *M. viridifaciens* sialidase, it is possible that when expressed in BL21 cells NanH3 is proteolytically truncated resulting in two active enzymes, differing in molecular weight. It would be interesting to confirm whether these two forms are produced by GP 74 in culture. It was shown that when *M. viridifaciens* was exposed to colominic acid the soluble 41 kDa sialidase was secreted but when the bacteria were treated with milk casein the 68 kDa soluble sialidase was secreted [206]. Perhaps, the 68 kDa isoform was secreted due to its galactose-binding domain that might play an important role in casein metabolism. It is possible that the two isoforms representing NanH3 might be differentially expressed and play alternative roles depending on the conditions, potentially even affecting HIV infection differently.

The 68 kDa *M. viridifaciens* sialidase comprised of 601 residues divided into three domains: the 41 kDa canonical, six-bladed-propeller catalytic sialidase domain (residues 47–402); a β-sandwich linker domain (residues 403–505); and a β-sandwich galactose-binding domain (residues 506–647) [181]. The six blades of the propeller each comprise of a four-stranded antiparallel-sheet. The conserved Asp-boxes are located between the third and fourth strands of each sheet and it has been suggested that they are responsible for the fold between the strands, and/or play a role in protein secretion. The five Asp-boxes and RIP motif were mostly conserved between the two structures. The structure of NanH3 was predicted to comprise the canonical, six-bladed-propeller

catalytic sialidase domain and a smaller domain that resembled the β-sandwich linker domain. The function of the smaller domain is unknown.

Enzymatic characterisation determined that *C. perfringens* NanI K_m was 0.17 mM which was similar to previous findings [191,196]. NanH3 had a slightly higher K_m for MUN compared to NanI, suggesting NanI has a slightly higher affinity for MUN than NanH3. The V_{max} of NanI and NanH3 was, 17.4 mM/min and 0.4 – 1.2 mM/min, respectively and when k_{Cat} was calculated using the equation V_{max}/total enzyme, NanH3 turnover number (how many substrate molecules are transformed into products per unit time by a single enzyme) was 3-fold higher than NanI. Catalytic efficiency calculated with the equation k_{Cat}/K_m indicated that NanH3 was 1.3-fold better at hydrolysing MUN compared to NanI. However, the relevance of these enzyme kinetic parameters to the physiological function of NanH3 in the FRT is unknown.

A caveat of the study was that physiologically relevant substrates were not tested and thus we could not determine whether NanH3 substrate specificity was similar to NanI. NanI was shown to prefer α-2,3 linkages over α-2,6 bonds [188] whereas NanH3 cleaved both 2,3'-sialyllactose and 2,6'-sialyllactose equally well [1]. If NanI and NanH3 differ in catalytic efficiency for terminal sialic acid 2,3 linkages, they would not desialylate glycoconjugates with similar efficiency. HIV Env has been shown to carry N-glycans enriched with either α-2,3 or α-2,6-linked bonds, depending on 1) whether live virus or pseudoviruses were used in the *in vitro* assay and 2) the type of cell used to produce the viral particles. HEK293T cells sialylated N-glycans of pseudoviruses with only α-2,3-linked sialic acids, suggesting that if NanH3 recognised and hydrolysed α-2,3-linkages less efficiently than NanI, it would not affect HIV infection to the same extent considering that the IMCs were produced in HEK293T cells [31]. Additional substrates must be tested in order to determine the relevance of NanH3 to FGT-specific glycoconjugates as well as virus and host cell glycoproteins and glycolipids.

When TZM-bl cells were treated with recombinant NanH3, there was an increase in IMC infection compared to the positive, untreated control. When sialidase activity was inhibited, the increase was abrogated. Therefore, similar to *C. perfringens* NanI, NanH3 increased HIV infection when

target cells were treated with sialidase. Contrary to NanI, treatment of the virus and cells simultaneously with NanH3, apparently had no effect on infection. This could be because NanH3 and NanI have different substrate specificities and/or their enzyme activity is differentially affected by changes in the environment, such as pH.

Sialidases are very sensitive to pH changes and NanI activity was highest at pH 5.5. However, as it retained its activity at pH 7.2 [188], the approximate pH of mammalian cell culture, it was likely active during the HIV infection assay, albeit less efficiently than at pH 5.5. NanH3 pH sensitivity is unknown and thus its efficiency at cleaving sialic acids during the *in vitro* infection assay might have affected its apparent effect on IMC infection. This could have ramifications on its *in vivo* role in HIV acquisition because of pH changes in the FGT: the displacement of *Lactobacillus* spp., causes an increase in pH which might affect NanH3 activity.

There are many biochemical differences between bacterial sialidases which make comparison between species difficult. However, we have successfully cloned and isolated NanH3 from *G. piotii*, and for the first time, have shown that it is able to increase HIV infection *in vitro* through interaction with target cells.

Chapter 6

6.1 Discussion

Fifty two percent of women are BV-positive in sub-Saharan Africa, a region that also carries the highest burden of HIV infection [207], potentially because individuals with BV have a higher risk of HIV acquisition [15,46,171,208]. Multiple mechanisms have been suggested to explain the association, including how immune responses to BVAB result in high levels of proinflammatory cytokines within the FRT that are linked to HIV infection [105] . However, we hypothesised that, in addition to immune responses, it is also possible that certain BVAB directly enhance HIV infection through the action of virulence factors.

BVAB include, but are not limited to *Gardnerella vaginalis*, *Prevotella* sp., *Bacteroides* sp., *Peptostreptococcus* sp., *Mycoplasma hominis* and *Mobiluncus* spp. [209] [106]. *Gardnerella* is the main bacteria consistently associated with BV positive patients [210] although its association with HIV and FRT inflammation only trended to significance [105,211]. Furthermore, it is also known that *Gardnerella* is commonly found in the FRT of BV-negative women which suggests that either, the genus may not be important in initiating BV or it comprises of both pathogenic and non-pathogenic strains [92]. *Gardnerella* comprises of four species, *G. vaginalis, G. piotii, G. leopoldii*, and *G. Swidsinskii* as well as nine genome species [69] which could differ in virulence [92,144]. Differences in pathogenicity between strains may be due to variation in *Gardnerella* virulence factors: VLY expression and activity, biofilm formation capacity and sialidase activity. Therefore, this study investigated whether twenty *Gardnerella* strains isolated from women with Nugent-BV differ in expression and activity of VLY and sialidase, and whether differences were associated with biofilm formation and *in vitro* HIV infection.

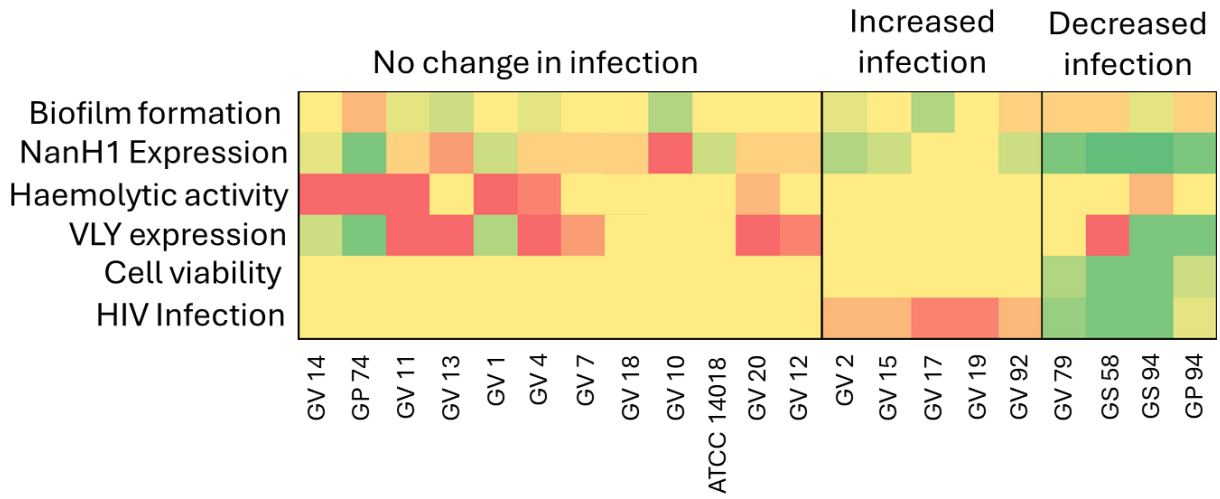


Figure 6.1 Overall comparison of *Gardnerella* strains. This is a graphical representation of Table 3.2 where the 20 strains were characterised based on their ability to impact HIV infection, cell viability, vly expression, haemolytic activity, nanH1 expression, and biofilm formation. The strains were grouped on their effect on HIV infection and cell viability. All measurements are relative to ATCC or another control as indicated in Table 3.2. Red represents a relative increase, yellow indicates no change relative to the control (and thus similar to each other) and green indicates a relative decrease. The intensity of the colour represents statistical differences with the lighter the colour, the less the effect and vice versa.

The culture medium and cell lysates from each strain were added to an HIV infection assay to determine the respective effect of VLY and NanH3 on IMC infection. Despite highly different phenotypes, the strains could be grouped according to their effect on HIV infection and cell viability with the majority not having an effect on either (Figure 6.1). Due to the cytotoxicity of four strains, HIV infection was reduced. And interestingly, this group comprised of the *G. swidsinskii* isolates and the *G. piotii* strains whereas the majority of the *G. vaginalis* strains did not affect infection nor viability of TZM-bl cells. The third group of strains increased HIV infection, did not induce cell death and all had similar VLY expression and activity. All of these strains were *G. vaginalis*, which could suggest that they were better able to enhance HIV infection. However, there was no overt consistent change in VLY expression nor activity that could explain why some strains affected HIV infection while others were cytotoxic. Although, the heat map suggested that the strains that did not have an effect on HIV infection were phenotypically diverse whereas those that either increased or decreased HIV infection had more homogenous characteristics (Figure 6.1). So, although no clear patterns emerged, it is possible that VLY activity could play a role in HIV infection.

Moreover, all comparisons of VLY expression and activity were relative to the ATCC 14018 strain to minimise the effect of inter-experimental error, so it is difficult to confirm the actual variation between strains. Despite the inconclusive relationship between VLY and increased IMC infection, we have shown that *Gardnerella* spp. secrete factors that can affect HIV infection. Furthermore, differences in cytotoxicity could also suggest that some strains might be better at disrupting the integrity of the protective epithelial cell layer of the FRT, allowing viruses to gain access to HIV permissive cells. Garcia*, et al.* [131] found variation in the sequence of *vly* isolated from different strains which should all be tested to determine variable impacts on HIV infection.

The way in which cell lysates were prepared affected whether strains reduced cell viability or not and preparations were optimised in an attempt to increase NanH3 activity. When cells were sonicated and centrifuged, the supernatants of seven strains reduced cell viability. All strains, irrespective of whether they were cytotoxic or not, could be grouped as increased, decreased, or had no effect on HIV infection relative to ATCC. Among these, were strains negative for NanH3, suggesting that other *Gardnerella* specific cell-associated factors were influencing HIV infection. When cells were disrupted by sonication and not centrifuged, the lysed cells increased HIV infection irrespective of whether they were positive or negative for NanH3. However, treatment with a sialidase inhibitor showed that NanH3 was responsible for the apparent increase in infection caused by the lysed cells of the two *G. piotii* strains. The complexity of inter-strain phenotype has been a challenge to identify correlations between changes in virulence determinants and HIV infection. However, we have shown that strains affect HIV infection differently which supports the hypothesis that some variants may be more pathogenic than others.

Correlation analysis indicated that there was no obvious relationship between differences in HIV infection, biofilm formation and expression and activity of VLY and NanH3. We hypothesised that the relationship was affected by the small sample size and masked by the high phenotypic diversity between strains. As there was evidence that NanH3 of GP 74 and GP 94 might enhance IMC infection, the virulence determinant was cloned, expressed, and purified. Both recombinant NanH3 and NanI from *C. perfringens* increased HIV infection when TZM-bl cells were treated with

the enzymes. However, HIV infection only increased when VLPs were treated with NanI and not NanH3. This suggested that the two sialidases might differ in substrate specificity, but without testing physiologically relevant substrates, we cannot confirm whether NanI cleaves sialic acids from the surface of viral particles more effectively than NanH3.

Caveats of the study include small sample size, unequal representation of *Gardnerella* spp., the isolation of multiple strains from a single individual and the absence of strains from Nugent-BV negative women. Differences between strains isolated from Nugent-BV negative and positive women might have been more pronounced, cancelling out the confounding effect of other bacteria-derived factors. The infection assay used to determine HIV infection, was conducting using TZM-bl cells. These HeLa cell derivatives are not a physiologically relevant cell type for HIV infection. However, a comparison between neutralisation assays conducted in peripheral blood mononuclear cells (PBMCs) and in TZM-bl cells showed that PBMCs may produce HIV-1 inhibitors which would affect our results [160]. Another assay which could have been utilised would be to test the effects of our samples on tissue explants which would mimic the FGT more suitably and was shown to exhibit variable results towards VLY based on the tissue type [155].

Despite all of these challenges, this study demonstrated that *Gardnerella* is a phenotypically complex and diverse bacteria which may easily be able to adjust expression of virulence factors based on its environment. In conclusion, sialidase is robustly capable of increasing HIV infection in an *in vitro* assay and therefore NanH3 and other sialidases in the FRT could play an important role in the increased susceptibility of BV-positive women to HIV infection.

6.2 Future work

This study has uncovered a number of questions pertaining to why *Gardnerella* spp. have such variable effects on HIV infection. This includes the regulation of VLY and NanH3 expression and activity. The cloning of VLY would confirm its role, or lack thereof, in HIV infection while NanH3 structural and functional analysis should be continued such as its sensitivity to pH and substrate specificity. A model system must be developed that includes tissue explants and mixed culture of *Gardnerella* strains for a more physiologically relevant environment. The effect of sialidase on mucin should be investigated as changes in mucin viscosity could facilitate HIV infection. Purification of NanH3 should be optimised to reduce the effect of contaminants. Finally, the molecular characterisation of the mechanism of how HIV infection is enhanced in BV-positive women must be supported by *in vivo* studies due to the intrinsic limitations of *in vitro* model systems.

References

1 Robinson, L. S., Schwebke, J., Lewis, W. G. & Lewis, A. L. Identification and characterization of NanH2 and NanH3, enzymes responsible for sialidase activity in the vaginal bacterium Gardnerella vaginalis. *J Biol Chem* **294**, 5230-5245, doi:10.1074/jbc.RA118.006221 (2019).

2 Carias, A. M. *et al.* Defining the interaction of HIV-1 with the mucosal barriers of the female reproductive tract. *J Virol* **87**, 11388-11400, doi:10.1128/JVI.01377-13 (2013).

3 Castro, J., Machado, D. & Cerca, N. Unveiling the role of Gardnerella vaginalis in polymicrobial Bacterial Vaginosis biofilms: the impact of other vaginal pathogens living as neighbors. *ISME J* **13**, 1306-1317, doi:10.1038/s41396-018-0337-0 (2019).

4 UNAIDS. *2021 UNAIDS Global AIDS Update — Confronting inequalities — Lessons for pandemic responses from 40 years of AIDS*, <https://www.unaids.org/en/resources/documents/2021/2021-global-aids-update> (2022).

5 Roser, M. & Ritchie, H. *HIV / AIDS A global epidemic and the leading cause of death in some countries*, <https://thepath.unaids.org/> (2014).

6 Durand, G., Bretelle, F. & Fenollar, F. in *Infectious Diseases* 498-504. e492 (Elsevier, 2017).

7 McKinnon, L. R. *et al.* The Evolving Facets of Bacterial Vaginosis: Implications for HIV Transmission. *AIDS Res Hum Retroviruses* **35**, 219-228, doi:10.1089/AID.2018.0304 (2019).

8 Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858, doi:10.1038/nprot.2015.053 (2015).

9 Wong, Y. P., Tan, G. C., Wong, K. K., Anushia, S. & Cheah, F. C. Gardnerella vaginalis in perinatology: An overview of the clinicopathological correlation. *The Malaysian journal of pathology* **40**, 267-286 (2018).

10 Kurukulasuriya, S. P., Patterson, M. H. & Hill, J. E. Slipped-Strand Mispairing in the Gene Encoding Sialidase NanH3 in Gardnerella spp. *Infect Immun* **89**, doi:10.1128/IAI.00583-20 (2021).

11 Wang, H. *et al.* Estimates of global, regional, and national incidence, prevalence, and mortality of HIV, 1980–2015: the Global Burden of Disease Study 2015. *The lancet HIV* **3**, e361-e387 (2016).

12 Dunkle, K. L. *et al.* Gender-based violence, relationship power, and risk of HIV infection in women attending antenatal clinics in South Africa. *Lancet* **363**, 1415-1421, doi:10.1016/S0140-6736(04)16098-4 (2004).

13 Kharsany, A. B. & Karim, Q. A. HIV Infection and AIDS in Sub-Saharan Africa: Current Status, Challenges and Opportunities. *Open AIDS J* **10**, 34-48, doi:10.2174/1874613601610010034 (2016).

14 Magadi, M. A. The disproportionate high risk of HIV infection among the urban poor in sub-Saharan Africa. *AIDS Behav* **17**, 1645-1654, doi:10.1007/s10461-012-0217-y (2013).

15 Taha, T. E. *et al.* Bacterial vaginosis and disturbances of vaginal flora: association with increased acquisition of HIV. *AIDS* **12**, 1699-1706, doi:10.1097/00002030-199813000-00019 (1998).

16 Barre-Sinoussi, F. *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**, 868-871 (1983).

17 Control, C. f. D. Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. *MMWR. Morbidity and mortality weekly report* **30**, 305 (1981).

18 Karim, Q. A. *et al.* Recruitment of high risk women for HIV prevention trials: baseline HIV prevalence and sexual behavior in the CAPRISA 004 tenofovir gel trial. *Trials* **12**, 67, doi:10.1186/1745-6215-12-67 (2011).

19 Abaasa, A. *et al.* Long-term consistent use of a vaginal microbicide gel among HIV-1 sero-discordant couples in a phase III clinical trial (MDP 301) in rural south-west Uganda. *Trials* **14**, 33, doi:10.1186/1745-6215-14-33 (2013).

20 UNAIDS. Global HIV statistics. (2023).

21 Coffin, J. M., Hughes, S. H. & Varmus, H. E. in *Retroviruses* (Cold Spring Harbor Laboratory Press, 1997).

22 Kowalski, M. *et al.* Functional regions of the envelope glycoprotein of human immunodeficiency virus type 1. *Science* **237**, 1351-1355, doi:10.1126/science.3629244 (1987).

23 Engelman, A. & Cherepanov, P. The structural biology of HIV-1: mechanistic and therapeutic insights. *Nat Rev Microbiol* **10**, 279-290, doi:10.1038/nrmicro2747 (2012).

24 Klatzmann, D. *et al.* T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* **312**, 767-768, doi:10.1038/312767a0 (1984).

25 Melikyan, G. B. *et al.* Evidence that the transition of HIV-1 gp41 into a six-helix bundle, not the bundle configuration, induces membrane fusion. *J Cell Biol* **151**, 413-423, doi:10.1083/jcb.151.2.413 (2000).

26 Liu, R. *et al.* Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* **86**, 367-377, doi:10.1016/s0092-8674(00)80110-5 (1996).

27 Pollakis, G. *et al.* N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization. *J Biol Chem* **276**, 13433-13441, doi:10.1074/jbc.M009779200 (2001).

28 Zhang, L., Huang, Y., He, T., Cao, Y. & Ho, D. D. HIV-1 subtype and second-receptor use. *Nature* **383**, 768-768 (1996).

29 Keele, B. F. *et al.* Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci U S A* **105**, 7552-7557, doi:10.1073/pnas.0802203105 (2008).

30 Wyatt, R. & Sodroski, J. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* **280**, 1884-1888, doi:10.1126/science.280.5371.1884 (1998).

31 Doores, K. J. The HIV glycan shield as a target for broadly neutralizing antibodies. *FEBS J* **282**, 4679-4691, doi:10.1111/febs.13530 (2015).

32 Anderson, D. J. *et al.* Targeting Trojan Horse leukocytes for HIV prevention. *AIDS* **24**, 163-187, doi:10.1097/QAD.0b013e32833424c8 (2010).

33 Zhong, P., Agosto, L. M., Munro, J. B. & Mothes, W. Cell-to-cell transmission of viruses. *Curr Opin Virol* **3**, 44-50, doi:10.1016/j.coviro.2012.11.004 (2013).

34 Balzarini, J. Targeting the glycans of glycoproteins: a novel paradigm for antiviral therapy. *Nat Rev Microbiol* **5**, 583-597, doi:10.1038/nrmicro1707 (2007).

35 Sun, J. *et al.* Syncytium formation and HIV-1 replication are both accentuated by purified influenza and virus-associated neuraminidase. *J Biol Chem* **277**, 9825-9833, doi:10.1074/jbc.M110764200 (2002).

36 Herzberg, M. C., Vacharaksa, A., Gebhard, K. H., Giacaman, R. A. & Ross, K. F. Plausibility of HIV-1 Infection of Oral Mucosal Epithelial Cells. *Adv Dent Res* **23**, 38-44, doi:10.1177/0022034511399283 (2011).

37 Tan, X. & Phillips, D. M. Cell-mediated infection of cervix derived epithelial cells with primary isolates of human immunodeficiency virus. *Arch Virol* **141**, 1177-1189, doi:10.1007/BF01718823 (1996).

38 Permanyer, M., Ballana, E. & Esté, J. A. Endocytosis of HIV: anything goes. *Trends in microbiology* **18**, 543-551 (2010).

39 Burgener, A., McGowan, I. & Klatt, N. R. HIV and mucosal barrier interactions: consequences for transmission and pathogenesis. *Current opinion in immunology* **36**, 22-30 (2015).

40 Boskey, E. R., Cone, R. A., Whaley, K. J. & Moench, T. R. Origins of vaginal acidity: high D/L lactate ratio is consistent with bacteria being the primary source. *Hum Reprod* **16**, 1809-1813 (2001).

41 Ravel, J. *et al.* Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* **108 Suppl 1**, 4680-4687, doi:10.1073/pnas.1002611107 (2011).

42 Munoz, A. *et al.* Correction to: Modeling the temporal dynamics of cervicovaginal microbiota identifies targets that may promote reproductive health. *Microbiome* **9**, 206, doi:10.1186/s40168-021-01171-1 (2021).

43 Mendes-Soares, H., Suzuki, H., Hickey, R. J. & Forney, L. J. Comparative functional genomics of Lactobacillus spp. reveals possible mechanisms for specialization of vaginal lactobacilli to their environment. *J Bacteriol* **196**, 1458-1470, doi:10.1128/JB.01439-13 (2014).

44 Cicenia, A. *et al.* Postbiotic activities of lactobacilli-derived factors. *J Clin Gastroenterol* **48 Suppl 1**, S18-22, doi:10.1097/MCG.0000000000000231 (2014).

45 Talarico, T. L. & Dobrogosz, W. J. Chemical characterization of an antimicrobial substance produced by Lactobacillus reuteri. *Antimicrob Agents Chemother* **33**, 674-679, doi:10.1128/AAC.33.5.674 (1989).

46 Gosmann, C. *et al.* Lactobacillus-Deficient Cervicovaginal Bacterial Communities Are Associated with Increased HIV Acquisition in Young South African Women. *Immunity* **46**, 29-37, doi:10.1016/j.immuni.2016.12.013 (2017).

47 Verstraelen, H. *et al.* Longitudinal analysis of the vaginal microflora in pregnancy suggests that L. crispatus promotes the stability of the normal vaginal microflora and that L. gasseri and/or L. iners are more conducive to the occurrence of abnormal vaginal microflora. *BMC Microbiol* **9**, 116, doi:10.1186/1471-2180-9-116 (2009).

48 Castro, J., Martins, A. P., Rodrigues, M. E. & Cerca, N. Lactobacillus crispatus represses vaginolysin expression by BV associated Gardnerella vaginalis and reduces cell cytotoxicity. *Anaerobe* **50**, 60-63, doi:10.1016/j.anaerobe.2018.01.014 (2018).

49 Petrova, M. I., Reid, G., Vaneechoutte, M. & Lebeer, S. Lactobacillus iners: Friend or Foe? *Trends Microbiol* **25**, 182-191, doi:10.1016/j.tim.2016.11.007 (2017).

50 Fettweis, J. M. *et al.* Differences in vaginal microbiome in African American women versus women of European ancestry. *Microbiology (Reading)* **160**, 2272-2282, doi:10.1099/mic.0.081034-0 (2014).

51 Borgdorff, H. *et al.* Lactobacillus-dominated cervicovaginal microbiota associated with reduced HIV/STI prevalence and genital HIV viral load in African women. *ISME J* **8**, 1781-1793, doi:10.1038/ismej.2014.26 (2014).

52 Schwebke, J. R., Muzny, C. A. & Josey, W. E. Role of Gardnerella vaginalis in the pathogenesis of bacterial vaginosis: a conceptual model. *J Infect Dis* **210**, 338-343, doi:10.1093/infdis/jiu089 (2014).

53 Swidsinski, A., Loening-Baucke, V., Swidsinski, S. & Verstraelen, H. Polymicrobial Gardnerella biofilm resists repeated intravaginal antiseptic treatment in a subset of women with bacterial vaginosis: a preliminary report. *Arch Gynecol Obstet* **291**, 605-609, doi:10.1007/s00404-014-3484-1 (2015).

54 Catlin, B. W. Gardnerella vaginalis: characteristics, clinical considerations, and controversies. *Clin Microbiol Rev* **5**, 213-237, doi:10.1128/CMR.5.3.213 (1992).

55 Nugent, R. P., Krohn, M. A. & Hillier, S. L. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *Journal of clinical microbiology* **29**, 297-301 (1991).

56 Dingens, A. S., Fairfortune, T. S., Reed, S. & Mitchell, C. Bacterial vaginosis and adverse outcomes among full-term infants: a cohort study. *BMC Pregnancy Childbirth* **16**, 278, doi:10.1186/s12884-016-1073-y (2016).

57 Sharma, H., Tal, R., Clark, N. A. & Segars, J. H. Microbiota and pelvic inflammatory disease. *Semin Reprod Med* **32**, 43-49, doi:10.1055/s-0033-1361822 (2014).

58 Kavoussi, S. K., Pearlman, M. D., Burke, W. M. & Lebovic, D. I. Endometrioma complicated by tubo-ovarian abscess in a woman with bacterial vaginosis. *Infect Dis Obstet Gynecol* **2006**, 84140, doi:10.1155/IDOG/2006/84140 (2006).

59 Allsworth, J. E. & Peipert, J. F. Severity of bacterial vaginosis and the risk of sexually transmitted infection. *Am J Obstet Gynecol* **205**, 113 e111-116, doi:10.1016/j.ajog.2011.02.060 (2011).

60 Wiesenfeld, H. C., Hillier, S. L., Krohn, M. A., Landers, D. V. & Sweet, R. L. Bacterial vaginosis is a strong predictor of Neisseria gonorrhoeae and Chlamydia trachomatis infection. *Clin Infect Dis* **36**, 663-668, doi:10.1086/367658 (2003).

61 Cherpes, T. L., Meyn, L. A., Krohn, M. A., Lurie, J. G. & Hillier, S. L. Association between acquisition of herpes simplex virus type 2 in women and bacterial vaginosis. *Clin Infect Dis* **37**, 319-325, doi:10.1086/375819 (2003).

62 Mirmonsef, P., Krass, L., Landay, A. & Spear, G. T. The role of bacterial vaginosis and trichomonas in HIV transmission across the female genital tract. *Curr HIV Res* **10**, 202-210 (2012).

63 Myer, L., Kuhn, L., Stein, Z. A., Wright, T. C., Jr. & Denny, L. Intravaginal practices, bacterial vaginosis, and women's susceptibility to HIV infection: epidemiological evidence and biological mechanisms. *Lancet Infect Dis* **5**, 786-794, doi:10.1016/S1473-3099(05)70298-X (2005).

64 Machado, A. & Cerca, N. Influence of Biofilm Formation by Gardnerella vaginalis and Other Anaerobes on Bacterial Vaginosis. *J Infect Dis* **212**, 1856-1861, doi:10.1093/infdis/jiv338 (2015).

65 Livengood III, C. H. Bacterial vaginosis: an overview for 2009. *Reviews in obstetrics and Gynecology* **2**, 28 (2009).

66 Patterson, J. L., Stull-Lane, A., Girerd, P. H. & Jefferson, K. K. Analysis of adherence, biofilm formation and cytotoxicity suggests a greater virulence potential of Gardnerella vaginalis relative to other bacterial-vaginosis-associated anaerobes. *Microbiology* **156**, 392-399, doi:10.1099/mic.0.034280-0 (2010).

67 Shipitsyna, E. *et al.* Composition of the vaginal microbiota in women of reproductive age--sensitive and specific molecular diagnosis of bacterial vaginosis is possible? *PLoS One* **8**, e60670, doi:10.1371/journal.pone.0060670 (2013).

68 Gelber, S. E., Aguilar, J. L., Lewis, K. L. & Ratner, A. J. Functional and phylogenetic characterization of Vaginolysin, the human-specific cytolysin from Gardnerella vaginalis. *J Bacteriol* **190**, 3896-3903, doi:10.1128/JB.01965-07 (2008).

69 Vaneechoutte, M. *et al.* Emended description of Gardnerella vaginalis and description of Gardnerella leopoldii sp. nov., Gardnerella piotii sp. nov. and Gardnerella swidsinskii sp.

nov., with delineation of 13 genomic species within the genus Gardnerella. *Int J Syst Evol Microbiol* **69**, 679-687, doi:10.1099/ijsem.0.003200 (2019).

70 Morrill, S., Gilbert, N. M. & Lewis, A. L. Gardnerella vaginalis as a cause of bacterial vaginosis: appraisal of the evidence from in vivo models. *Frontiers in cellular and infection microbiology* **10**, 168 (2020).

71 Aroutcheva, A. A., Simoes, J. A., Behbakht, K. & Faro, S. Gardnerella vaginalis isolated from patients with bacterial vaginosis and from patients with healthy vaginal ecosystems. *Clin Infect Dis* **33**, 1022-1027, doi:10.1086/323030 (2001).

72 van de Wijgert, J. H. *et al.* The vaginal microbiota: what have we learned after a decade of molecular characterization? *PLoS One* **9**, e105998, doi:10.1371/journal.pone.0105998 (2014).

73 Benito, R., Vazquez, J. A., Berron, S., Fenoll, A. & Saez-Neito, J. A. A modified scheme for biotyping Gardnerella vaginalis. *J Med Microbiol* **21**, 357-359, doi:10.1099/00222615-21-4-357 (1986).

74 Greenwood, J. & Pickett, M. Transfer of Haemophilus vaginalis Gardner and Dukes to a New Genus, Gardnerella: G. vaginalis (Gardner and Dukes) comb. nov. *International Journal of Systematic and Evolutionary Microbiology* **30**, 170-178 (1980).

75 Piot, P., Van Dyck, E., Totten, P. & Holmes, K. Identification of Gardnerella (Haemophilus) vaginalis. *Journal of clinical microbiology* **15**, 19-24 (1982).

76 Ingianni, A., Petruzzelli, S., Morandotti, G. & Pompei, R. Genotypic differentiation of Gardnerella vaginalis by amplified ribosomal DNA restriction analysis (ARDRA). *FEMS Immunol Med Microbiol* **18**, 61-66, doi:10.1111/j.1574-695X.1997.tb01028.x (1997).

77 Santiago, G. L. *et al.* Gardnerella vaginalis comprises three distinct genotypes of which only two produce sialidase. *Am J Obstet Gynecol* **204**, 450 e451-457, doi:10.1016/j.ajog.2010.12.061 (2011).

78 Paramel Jayaprakash, T., Schellenberg, J. J. & Hill, J. E. Resolution and characterization of distinct cpn60-based subgroups of Gardnerella vaginalis in the vaginal microbiota. *PLoS One* **7**, e43009, doi:10.1371/journal.pone.0043009 (2012).

79 Schellenberg, J. J., Patterson, M. H. & Hill, J. E. Gardnerella vaginalis diversity and ecology in relation to vaginal symptoms. *Res Microbiol* **168**, 837-844, doi:10.1016/j.resmic.2017.02.011 (2017).

80 Schellenberg, J. J. *et al.* Gardnerella vaginalis Subgroups Defined by cpn60 Sequencing and Sialidase Activity in Isolates from Canada, Belgium and Kenya. *PLoS One* **11**, e0146510, doi:10.1371/journal.pone.0146510 (2016).

81 Pleckaityte, M., Janulaitiene, M., Lasickiene, R. & Zvirbliene, A. Genetic and biochemical diversity of Gardnerella vaginalis strains isolated from women with bacterial vaginosis. *FEMS Immunol Med Microbiol* **65**, 69-77, doi:10.1111/j.1574-695X.2012.00940.x (2012).

82 Ahmed, A. *et al.* Comparative genomic analyses of 17 clinical isolates of Gardnerella vaginalis provide evidence of multiple genetically isolated clades consistent with subspeciation into genovars. *J Bacteriol* **194**, 3922-3937, doi:10.1128/JB.00056-12 (2012).

83 Martinez, J. L., Delgado-Iribarren, A. & Baquero, F. Mechanisms of iron acquisition and bacterial virulence. *FEMS Microbiol Rev* **6**, 45-56, doi:10.1111/j.1574-6968.1990.tb04085.x (1990).

84 Griffiths, E. Iron and bacterial virulence--a brief overview. *Biol Met* **4**, 7-13 (1991).

85 Jarosik, G. P., Land, C. B., Duhon, P., Chandler Jr, R. & Mercer, T. Acquisition of iron by Gardnerella vaginalis. *Infection and immunity* **66**, 5041-5047 (1998).

86 Yeoman, C. J. *et al.* Comparative genomics of Gardnerella vaginalis strains reveals substantial differences in metabolic and virulence potential. *PLoS One* **5**, e12411, doi:10.1371/journal.pone.0012411 (2010).

87 Jefferson, K. K. What drives bacteria to produce a biofilm? *FEMS Microbiol Lett* **236**, 163-173, doi:10.1016/j.femsle.2004.06.005 (2004).

88 Lewis, K. Riddle of biofilm resistance. *Antimicrob Agents Chemother* **45**, 999-1007, doi:10.1128/AAC.45.4.999-1007.2001 (2001).

89 Swidsinski, A. *et al.* Adherent biofilms in bacterial vaginosis. *Obstet Gynecol* **106**, 1013-1023, doi:10.1097/01.AOG.0000183594.45524.d2 (2005).

90 Swidsinski, A. *et al.* An adherent Gardnerella vaginalis biofilm persists on the vaginal epithelium after standard therapy with oral metronidazole. *Am J Obstet Gynecol* **198**, 97 e91-96, doi:10.1016/j.ajog.2007.06.039 (2008).

91 Machado, A., Jefferson, K. K. & Cerca, N. Interactions between Lactobacillus crispatus and bacterial vaginosis (BV)-associated bacterial species in initial attachment and biofilm formation. *Int J Mol Sci* **14**, 12004-12012, doi:10.3390/ijms140612004 (2013).

92 Harwich, M. D. *et al.* Drawing the line between commensal and pathogenic Gardnerella vaginalis through genome analysis and virulence studies. *BMC genomics* **11**, 375 (2010).

93 Briselden, A. M., Moncla, B. J., Stevens, C. E. & Hillier, S. L. Sialidases (neuraminidases) in bacterial vaginosis and bacterial vaginosis-associated microflora. *J Clin Microbiol* **30**, 663-666, doi:10.1128/jcm.30.3.663-666.1992 (1992).

94 Aldunate, M. *et al.* Vaginal concentrations of lactic acid potently inactivate HIV. *J Antimicrob Chemother* **68**, 2015-2025, doi:10.1093/jac/dkt156 (2013).

95 Agarwal, K. & Lewis, A. L. Vaginal sialoglycan foraging by Gardnerella vaginalis: mucus barriers as a meal for unwelcome guests? *Glycobiology* **31**, 667-680, doi:10.1093/glycob/cwab024 (2021).

96 Ferreira, C. S. T., Marconi, C., Parada, C., Ravel, J. & da Silva, M. G. Sialidase Activity in the Cervicovaginal Fluid Is Associated With Changes in Bacterial Components of Lactobacillus-Deprived Microbiota. *Front Cell Infect Microbiol* **11**, 813520, doi:10.3389/fcimb.2021.813520 (2021).

97 Lewis, W. G., Robinson, L. S., Gilbert, N. M., Perry, J. C. & Lewis, A. L. Degradation, foraging, and depletion of mucus sialoglycans by the vagina-adapted Actinobacterium Gardnerella vaginalis. *J Biol Chem* **288**, 12067-12079, doi:10.1074/jbc.M113.453654 (2013).

98 Lewis, W. G. *et al.* Hydrolysis of secreted sialoglycoprotein immunoglobulin A (IgA) in ex vivo and biochemical models of bacterial vaginosis. *J Biol Chem* **287**, 2079-2089, doi:10.1074/jbc.M111.278135 (2012).

99 Roselletti, E. *et al.* Apoptosis of vaginal epithelial cells in clinical samples from women with diagnosed bacterial vaginosis. *Scientific Reports* **10**, 1978 (2020).

100 Gardner, H. L. & Dukes, C. D. Haemophilus vaginalis vaginitis: a newly defined specific infection previously classified non-specific vaginitis. *Am J Obstet Gynecol* **69**, 962-976 (1955).

101 Randis, T. M. *et al.* Vaginolysin drives epithelial ultrastructural responses to Gardnerella vaginalis. *Infect Immun* **81**, 4544-4550, doi:10.1128/IAI.00627-13 (2013).

102 Huffman, D. L. *et al.* Mitogen-activated protein kinase pathways defend against bacterial pore-forming toxins. *Proc Natl Acad Sci U S A* **101**, 10995-11000, doi:10.1073/pnas.0404073101 (2004).

103 Cohen, C. R. *et al.* Bacterial vaginosis associated with increased risk of female-to-male HIV-1 transmission: a prospective cohort analysis among African couples. *PLoS Med* **9**, e1001251, doi:10.1371/journal.pmed.1001251 (2012).

104 Schmid, G., Markowitz, L., Joesoef, R. & Koumans, E. Bacterial vaginosis and HIV infection. *Sex Transm Infect* **76**, 3-4, doi:10.1136/sti.76.1.3 (2000).

105 Anahtar, M. N. *et al.* Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. *Immunity* **42**, 965-976, doi:10.1016/j.immuni.2015.04.019 (2015).

106 Spear, G. T., St John, E. & Zariffard, M. R. Bacterial vaginosis and human immunodeficiency virus infection. *AIDS Res Ther* **4**, 25, doi:10.1186/1742-6405-4-25 (2007).

107 Klebanoff, S. J. & Coombs, R. W. Viricidal effect of Lactobacillus acidophilus on human immunodeficiency virus type 1: possible role in heterosexual transmission. *J Exp Med* **174**, 289-292, doi:10.1084/jem.174.1.289 (1991).

108 Happel, A. U. *et al.* Exploring potential of vaginal Lactobacillus isolates from South African women for enhancing treatment for bacterial vaginosis. *PLoS Pathog* **16**, e1008559, doi:10.1371/journal.ppat.1008559 (2020).

109 Cohen, C. R. *et al.* Increased interleukin-10 in the the endocervical secretions of women with non-ulcerative sexually transmitted diseases: a mechanism for enhanced HIV-1 transmission? *AIDS* **13**, 327-332, doi:10.1097/00002030-199902250-00004 (1999).

110 Deman, J., Mareel, M. & Bruyneel, E. Effects of calcium and bound sialic acid on the viscosity of mucin. *Biochim Biophys Acta* **297**, 486-490, doi:10.1016/0304-4165(73)90095-0 (1973).

111 Olmsted, S. S., Meyn, L. A., Rohan, L. C. & Hillier, S. L. Glycosidase and proteinase activity of anaerobic gram-negative bacteria isolated from women with bacterial vaginosis. *Sex Transm Dis* **30**, 257-261, doi:10.1097/00007435-200303000-00016 (2003).

112 Sun, J., Barbeau, B., Sato, S. & Tremblay, M. J. Neuraminidase from a bacterial source enhances both HIV-1-mediated syncytium formation and the virus binding/entry process. *Virology* **284**, 26-36 (2001).

113 Wang, W. *et al.* A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization. *Retrovirology* **10**, 14, doi:10.1186/1742-4690-10-14 (2013).

114 Zou, Z. *et al.* Siglecs facilitate HIV-1 infection of macrophages through adhesion with viral sialic acids. *PLoS One* **6**, e24559, doi:10.1371/journal.pone.0024559 (2011).

115 Corfield, A. P., Higa, H., Paulson, J. C. & Schauer, R. The specificity of viral and bacterial sialidases for alpha(2-3)- and alpha(2-6)-linked sialic acids in glycoproteins. *Biochim Biophys Acta* **744**, 121-126, doi:10.1016/0167-4838(83)90080-8 (1983).

116 Raska, M. *et al.* Glycosylation patterns of HIV-1 gp120 depend on the type of expressing cells and affect antibody recognition. *Journal of Biological Chemistry* **285**, 20860-20869 (2010).

117 Mercer, J. & Helenius, A. Virus entry by macropinocytosis. *Nat Cell Biol* **11**, 510-520, doi:10.1038/ncb0509-510 (2009).

118 Gobeil, L. A., Lodge, R. & Tremblay, M. J. Macropinocytosis-like HIV-1 internalization in macrophages is CCR5 dependent and leads to efficient but delayed degradation in endosomal compartments. *J Virol* **87**, 735-745, doi:10.1128/JVI.01802-12 (2013).

119 Cohen, P. S. *et al.* The critical role of p38 MAP kinase in T cell HIV-1 replication. *Molecular Medicine* **3**, 339-346 (1997).

120 Kawamura, T., Kurtz, S. E., Blauvelt, A. & Shimada, S. The role of Langerhans cells in the sexual transmission of HIV. *J Dermatol Sci* **40**, 147-155, doi:10.1016/j.jdermsci.2005.08.009 (2005).

121 Evans, B. C. *et al.* Ex vivo red blood cell hemolysis assay for the evaluation of pH-responsive endosomolytic agents for cytosolic delivery of biomacromolecular drugs. *J Vis Exp*, e50166, doi:10.3791/50166 (2013).

122 Reed, L. J. & Muench, H. A simple method of estimating fifty per cent endpoints. *American journal of epidemiology* **27**, 493-497 (1938).

123 Castro, J. *et al.* Comparative transcriptomic analysis of Gardnerella vaginalis biofilms vs. planktonic cultures using RNA-seq. *NPJ Biofilms Microbiomes* **3**, 3, doi:10.1038/s41522-017-0012-7 (2017).

124 Wei, X. *et al.* Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob Agents Chemother* **46**, 1896-1905, doi:10.1128/AAC.46.6.1896-1905.2002 (2002).

125 Montefiori, D. C. Measuring HIV neutralization in a luciferase reporter gene assay. *Methods Mol Biol* **485**, 395-405, doi:10.1007/978-1-59745-170-3_26 (2009).

126 Lumngwena, E. N. *et al.* Selective transmission of some HIV-1 subtype C variants might depend on Envelope stimulating dendritic cells to secrete IL-10. *PLoS One* **15**, e0227533, doi:10.1371/journal.pone.0227533 (2020).

127 Ma, X. *et al.* Biofilm and pathogenic factor analysis of Gardnerella vaginalis associated with bacterial vaginosis in Northeast China. *Frontiers in Microbiology* **13**, 1033040 (2022).

128 Moncla, B., Braham, P. & Hillier, S. Sialidase (neuraminidase) activity among gram-negative anaerobic and capnophilic bacteria. *Journal of clinical microbiology* **28**, 422-425 (1990).

129 Hill, J. E., Albert, A. Y. K. & Group, V. R. Resolution and Cooccurrence Patterns of Gardnerella leopoldii, G. swidsinskii, G. piotii, and G. vaginalis within the Vaginal Microbiome. *Infect Immun* **87**, doi:10.1128/IAI.00532-19 (2019).

130 Qin, H. & Xiao, B. Research Progress on the Correlation Between Gardnerella Typing and Bacterial Vaginosis. *Front Cell Infect Microbiol* **12**, 858155, doi:10.3389/fcimb.2022.858155 (2022).

131 Garcia, E. M. *et al.* Sequence Comparison of Vaginolysin from Different Gardnerella Species. *Pathogens* **10**, doi:10.3390/pathogens10020086 (2021).

132 Castro, J. *et al.* Using an in-vitro biofilm model to assess the virulence potential of bacterial vaginosis or non-bacterial vaginosis Gardnerella vaginalis isolates. *Sci Rep* **5**, 11640, doi:10.1038/srep11640 (2015).

133 Babiychuk, E. B., Monastyrskaya, K., Potez, S. & Draeger, A. Blebbing confers resistance against cell lysis. *Cell Death Differ* **18**, 80-89, doi:10.1038/cdd.2010.81 (2011).

134 Brito, C., Cabanes, D., Sarmento Mesquita, F. & Sousa, S. Mechanisms protecting host cells against bacterial pore-forming toxins. *Cell Mol Life Sci* **76**, 1319-1339, doi:10.1007/s00018-018-2992-8 (2019).

135 Mercer, J., Schelhaas, M. & Helenius, A. Virus entry by endocytosis. *Annu Rev Biochem* **79**, 803-833, doi:10.1146/annurev-biochem-060208-104626 (2010).

136 Rottini, G. *et al.* Identification and partial characterization of a cytolytic toxin produced by Gardnerella vaginalis. *Infect Immun* **58**, 3751-3758, doi:10.1128/iai.58.11.3751-3758.1990 (1990).

137 Salamov, V. S. A. & Solovyevand, A. Automatic annotation of microbial genomes and metagenomic sequences. *Metagenomics and its applications in agriculture, biomedicine and environmental studies*, 61-78 (2011).

138 Kobayashi, M., Nagata, K. & Ishihama, A. Promoter selectivity of Escherichia coli RNA polymerase: effect of base substitutions in the promoter -35 region on promoter strength. *Nucleic Acids Res* **18**, 7367-7372, doi:10.1093/nar/18.24.7367 (1990).

139 Hooven, T. A., Randis, T. M., Hymes, S. R., Rampersaud, R. & Ratner, A. J. Retrocyclin inhibits Gardnerella vaginalis biofilm formation and toxin activity. *J Antimicrob Chemother* **67**, 2870-2872, doi:10.1093/jac/dks305 (2012).

140 Kudryashova, E., Seveau, S., Lu, W. & Kudryashov, D. S. Retrocyclins neutralize bacterial toxins by potentiating their unfolding. *Biochem J* **467**, 311-320, doi:10.1042/BJ20150049 (2015).

141 Hardy, L. *et al.* The presence of the putative Gardnerella vaginalis sialidase A gene in vaginal specimens is associated with bacterial vaginosis biofilm. *PLoS One* **12**, e0172522, doi:10.1371/journal.pone.0172522 (2017).

142 Sauer, K. The genomics and proteomics of biofilm formation. *Genome Biol* **4**, 219, doi:10.1186/gb-2003-4-6-219 (2003).

143 Gardner, M. J. & Altman, D. G. *Differences between means: type I and type II errors and power*, <https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/5-differences-between-means-type-i-an> (2023).

144 Janulaitiene, M. *et al.* Phenotypic characterization of Gardnerella vaginalis subgroups suggests differences in their virulence potential. *PLoS One* **13**, e0200625, doi:10.1371/journal.pone.0200625 (2018).

145 Turner, E., Sobel, J. D. & Akins, R. A. Prognosis of recurrent bacterial vaginosis based on longitudinal changes in abundance of Lactobacillus and specific species of Gardnerella. *PLoS One* **16**, e0256445, doi:10.1371/journal.pone.0256445 (2021).

146 Shvartsman, E., Hill, J. E., Sandstrom, P. & MacDonald, K. S. Gardnerella Revisited: Species Heterogeneity, Virulence Factors, Mucosal Immune Responses, and Contributions to Bacterial Vaginosis. *Infect Immun* **91**, e0039022, doi:10.1128/iai.00390-22 (2023).

147 Balzarini, J. & Van Damme, L. Microbicide drug candidates to prevent HIV infection. *Lancet* **369**, 787-797, doi:10.1016/S0140-6736(07)60202-5 (2007).

148 Bohr, L. L., Mortimer, T. D. & Pepperell, C. S. Lateral Gene Transfer Shapes Diversity of Gardnerella spp. *Front Cell Infect Microbiol* **10**, 293, doi:10.3389/fcimb.2020.00293 (2020).

149 Lucchini, S. *et al.* H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog* **2**, e81, doi:10.1371/journal.ppat.0020081 (2006).

150 Shishpal, P., Kasarpalkar, N., Singh, D. & Bhor, V. M. Characterization of Gardnerella vaginalis membrane vesicles reveals a role in inducing cytotoxicity in vaginal epithelial cells. *Anaerobe* **61**, 102090, doi:10.1016/j.anaerobe.2019.102090 (2020).

151 Deacon, K., Mistry, P., Chernoff, J., Blank, J. L. & Patel, R. p38 Mitogen-activated protein kinase mediates cell death and p21-activated kinase mediates cell survival during

chemotherapeutic drug-induced mitotic arrest. *Mol Biol Cell* **14**, 2071-2087, doi:10.1091/mbc.e02-10-0653 (2003).

152 Owji, H., Nezafat, N., Negahdaripour, M., Hajiebrahimi, A. & Ghasemi, Y. A comprehensive review of signal peptides: Structure, roles, and applications. *Eur J Cell Biol* **97**, 422-441, doi:10.1016/j.ejcb.2018.06.003 (2018).

153 Al-Faris, L., Al-Humood, S., Behbehani, F. & Sallam, H. Altered Expression Pattern of CD55 and CD59 on Red Blood Cells in Anemia of Chronic Kidney Disease. *Med Princ Pract* **26**, 516-522, doi:10.1159/000481823 (2017).

154 Ragaliauskas, T. *et al.* Inerolysin and vaginolysin, the cytolysins implicated in vaginal dysbiosis, differently impair molecular integrity of phospholipid membranes. *Sci Rep* **9**, 10606, doi:10.1038/s41598-019-47043-5 (2019).

155 Garcia, E. M., Kraskauskiene, V., Koblinski, J. E. & Jefferson, K. K. Interaction of Gardnerella vaginalis and Vaginolysin with the Apical versus Basolateral Face of a Three-Dimensional Model of Vaginal Epithelium. *Infect Immun* **87**, doi:10.1128/IAI.00646-18 (2019).

156 Varki, A. & Gagneux, P. Multifarious roles of sialic acids in immunity. *Ann N Y Acad Sci* **1253**, 16-36, doi:10.1111/j.1749-6632.2012.06517.x (2012).

157 Wiggins, R., Hicks, S., Soothill, P., Millar, M. & Corfield, A. Mucinases and sialidases: their role in the pathogenesis of sexually transmitted infections in the female genital tract. *Sexually transmitted infections* **77**, 402-408 (2001).

158 Guttman, M., Kahn, M., Garcia, N. K., Hu, S. L. & Lee, K. K. Solution structure, conformational dynamics, and CD4-induced activation in full-length, glycosylated, monomeric HIV gp120. *J Virol* **86**, 8750-8764, doi:10.1128/JVI.07224-11 (2012).

159 Wilhelm, D., Behnken, H. N. & Meyer, B. Glycosylation assists binding of HIV protein gp120 to human CD4 receptor. *Chembiochem* **13**, 524-527, doi:10.1002/cbic.201100740 (2012).

160 Geonnotti, A. R. *et al.* Differential inhibition of human immunodeficiency virus type 1 in peripheral blood mononuclear cells and TZM-bl cells by endotoxin-mediated chemokine and gamma interferon production. *AIDS research and human retroviruses* **26**, 279-291 (2010).

161 Louder, M. K. *et al.* HIV-1 envelope pseudotyped viral vectors and infectious molecular clones expressing the same envelope glycoprotein have a similar neutralization phenotype, but culture in peripheral blood mononuclear cells is associated with decreased neutralization sensitivity. *Virology* **339**, 226-238, doi:10.1016/j.virol.2005.06.003 (2005).

162 Kim, H. J., Schweiker, S., Powell, K. & Levonis, S. An efficient and robust HPLC method to determine the sialylation levels of human epithelial cells. *PLoS One* **17**, e0257178, doi:10.1371/journal.pone.0257178 (2022).

163 Bulavaite, A., Maier, T. & Pleckaityte, M. Discrimination of Gardnerella Species by Combining MALDI-TOF Protein Profile, Chaperonin cpn60 Sequences, and Phenotypic Characteristics. *Pathogens* **10**, doi:10.3390/pathogens10030277 (2021).

164 Severgnini, M. *et al.* Gardnerella vaginalis clades in pregnancy: New insights into the interactions with the vaginal microbiome. *PLoS One* **17**, e0269590, doi:10.1371/journal.pone.0269590 (2022).

165 Grewal, K., MacIntyre, D. A. & Bennett, P. R. The reproductive tract microbiota in pregnancy. *Bioscience Reports* **41**, BSR20203908 (2021).

166 Provine, N. M., Puryear, W. B., Wu, X., Overbaugh, J. & Haigwood, N. L. The infectious molecular clone and pseudotyped virus models of human immunodeficiency virus type 1 exhibit significant differences in virion composition with only moderate differences in infectivity and inhibition sensitivity. *J Virol* **83**, 9002-9007, doi:10.1128/JVI.00423-09 (2009).

167 Tikole, S. & Sankararamakrishnan, R. A survey of mRNA sequences with a non-AUG start codon in RefSeq database. *J Biomol Struct Dyn* **24**, 33-42, doi:10.1080/07391102.2006.10507096 (2006).

168 Orr, M. W., Mao, Y., Storz, G. & Qian, S. B. Alternative ORFs and small ORFs: shedding light on the dark proteome. *Nucleic Acids Res* **48**, 1029-1042, doi:10.1093/nar/gkz734 (2020).

169 Hecht, A. *et al.* Measurements of translation initiation from all 64 codons in E. coli. *Nucleic Acids Res* **45**, 3615-3626, doi:10.1093/nar/gkx070 (2017).

170 Varki, A. Sialic acids in human health and disease. *Trends Mol Med* **14**, 351-360, doi:10.1016/j.molmed.2008.06.002 (2008).

171 Atashili, J., Poole, C., Ndumbe, P. M., Adimora, A. A. & Smith, J. S. Bacterial vaginosis and HIV acquisition: a meta-analysis of published studies. *AIDS* **22**, 1493-1501, doi:10.1097/QAD.0b013e3283021a37 (2008).

172 Mitchell, C. & Marrazzo, J. Bacterial vaginosis and the cervicovaginal immune response. *Am J Reprod Immunol* **71**, 555-563, doi:10.1111/aji.12264 (2014).

173 Thurman, A. R. & Doncel, G. F. Innate immunity and inflammatory response to Trichomonas vaginalis and bacterial vaginosis: relationship to HIV acquisition. *Am J Reprod Immunol* **65**, 89-98, doi:10.1111/j.1600-0897.2010.00902.x (2011).

174 Hoang, T. *et al.* The cervicovaginal mucus barrier to HIV-1 is diminished in bacterial vaginosis. *PLoS Pathog* **16**, e1008236, doi:10.1371/journal.ppat.1008236 (2020).

175 Francis, A. C. *et al.* HIV-1 replication complexes accumulate in nuclear speckles and integrate into speckle-associated genomic domains. *Nat Commun* **11**, 3505, doi:10.1038/s41467-020-17256-8 (2020).

176 Schweighardt, B. *et al.* R5 human immunodeficiency virus type 1 (HIV-1) replicates more efficiently in primary CD4+ T-cell cultures than X4 HIV-1. *J Virol* **78**, 9164-9173, doi:10.1128/JVI.78.17.9164-9173.2004 (2004).

177 Li, Z. *et al.* In silico insights into the receptor-mediated endocytosis of virus-like nanoparticles. *Chemical Physics Letters* **790**, 139360 (2022).

178 Hirst, G. K. The quantitative determination of influenza virus and antibodies by means of red cell agglutination. *The Journal of experimental medicine* **75**, 49 (1942).

179 Gottschalk, A. & Bhargava, A. in *The enzymes* Vol. 5 321-342 (Elsevier, 1971).

180 Soong, G. *et al.* Bacterial neuraminidase facilitates mucosal infection by participating in biofilm production. *J Clin Invest* **116**, 2297-2305, doi:10.1172/JCI27920 (2006).

181 Gaskell, A., Crennell, S. & Taylor, G. The three domains of a bacterial sialidase: a β-propeller, an immunoglobulin module and a galactose-binding jelly-roll. *Structure* **3**, 1197-1205 (1995).

182 Wang, Y. H. Sialidases From Clostridium perfringens and Their Inhibitors. *Front Cell Infect Microbiol* **9**, 462, doi:10.3389/fcimb.2019.00462 (2019).

183 Kim, S., Oh, D. B., Kang, H. A. & Kwon, O. Features and applications of bacterial sialidases. *Appl Microbiol Biotechnol* **91**, 1-15, doi:10.1007/s00253-011-3307-2 (2011).

184 Chang, Y. C., Uchiyama, S., Varki, A. & Nizet, V. Leukocyte inflammatory responses provoked by pneumococcal sialidase. *mBio* **3**, doi:10.1128/mBio.00220-11 (2012).

185 King, S. J. Pneumococcal modification of host sugars: a major contributor to colonization of the human airway? *Mol Oral Microbiol* **25**, 15-24, doi:10.1111/j.2041-1014.2009.00564.x (2010).

186 McCombs, J. E. & Kohler, J. J. Pneumococcal Neuraminidase Substrates Identified through Comparative Proteomics Enabled by Chemoselective Labeling. *Bioconjug Chem* **27**, 1013-1022, doi:10.1021/acs.bioconjchem.6b00050 (2016).

187 Shtyrya, Y., Mochalova, L. & Bovin, N. Influenza virus neuraminidase: structure and function. *Acta Naturae (англоязычная версия)* **1**, 26-32 (2009).

188 Li, J. & McClane, B. A. The Sialidases of Clostridium perfringens type D strain CN3718 differ in their properties and sensitivities to inhibitors. *Appl Environ Microbiol* **80**, 1701-1709, doi:10.1128/AEM.03440-13 (2014).

189 Vimr, E. R. Microbial sialidases: does bigger always mean better? *Trends Microbiol* **2**, 271-277, doi:10.1016/0966-842x(94)90003-5 (1994).

190 Mizan, S., Henk, A., Stallings, A., Maier, M. & Lee, M. D. Cloning and characterization of sialidases with 2-6' and 2-3' sialyl lactose specificity from Pasteurella multocida. *J Bacteriol* **182**, 6874-6883, doi:10.1128/JB.182.24.6874-6883.2000 (2000).

191 Kruse, S., Kleineidam, R. G., Roggentin, P. & Schauer, R. Expression and purification of a recombinant "small" sialidase from Clostridium perfringens A99. *Protein Expr Purif* **7**, 415-422, doi:10.1006/prep.1996.0062 (1996).

192 Traving, C., Schauer, R. & Roggentin, P. Gene structure of the ‘large’sialidase isoenzyme from Clostridium perfringens A99 and its relationship with other clostridial nanH proteins. *Glycoconjugate journal* **11**, 141-151 (1994).

193 Belinky, F., Rogozin, I. B. & Koonin, E. V. Selection on start codons in prokaryotes and potential compensatory nucleotide substitutions. *Sci Rep* **7**, 12422, doi:10.1038/s41598-017-12619-6 (2017).

194 Gasteiger, E. *et al.* ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* **31**, 3784-3788, doi:10.1093/nar/gkg563 (2003).

195 Theoret, J. R. *et al.* Native or Proteolytically Activated NanI Sialidase Enhances the Binding and Cytotoxic Activity of Clostridium perfringens Enterotoxin and Beta Toxin. *Infect Immun* **86**, doi:10.1128/IAI.00730-17 (2018).

196 Roggentin, T., Kleineidam, R. G., Schauer, R. & Roggentin, P. Effects of site-specific mutations on the enzymatic properties of a sialidase from Clostridium perfringens. *Glycoconj J* **9**, 235-240, doi:10.1007/BF00731135 (1992).

197 Roggentin, P. *et al.* Conserved sequences in bacterial and viral sialidases. *Glycoconjugate journal* **6**, 349-353 (1989).

198 Newstead, S., Watson, J. N., Knoll, T. L., Bennet, A. J. & Taylor, G. Structure and mechanism of action of an inverting mutant sialidase. *Biochemistry* **44**, 9117-9122, doi:10.1021/bi050517t (2005).

199 Newstead, S. L. *et al.* The structure of Clostridium perfringens NanI sialidase and its catalytic intermediates. *J Biol Chem* **283**, 9080-9088, doi:10.1074/jbc.M710247200 (2008).

200 Taylor, G. Sialidases: structures, biological significance and therapeutic potential. *Curr Opin Struct Biol* **6**, 830-837, doi:10.1016/s0959-440x(96)80014-5 (1996).

201 Bolanos-Garcia, V. M. & Davies, O. R. Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography. *Biochim Biophys Acta* **1760**, 1304-1313, doi:10.1016/j.bbagen.2006.03.027 (2006).

202 Pinske, C., Sargent, F. & Sawers, R. G. SlyD-dependent nickel delivery limits maturation of [NiFe]-hydrogenases in late-stationary phase Escherichia coli cells. *Metallomics* **7**, 683-690, doi:10.1039/c5mt00019j (2015).

203 Sakurada, K., Ohta, T. & Hasegawa, M. Cloning, expression, and characterization of the Micromonospora viridifaciens neuraminidase gene in Streptomyces lividans. *Journal of bacteriology* **174**, 6896-6903 (1992).

204 Seo, J.-W. & Lee, K.-J. Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. *BMB Reports* **37**, 35-44 (2004).

205 Valguarnera, E., Kinsella, R. L. & Feldman, M. F. Sugar and Spice Make Bacteria Not Nice: Protein Glycosylation and Its Influence in Pathogenesis. *J Mol Biol* **428**, 3206-3220, doi:10.1016/j.jmb.2016.04.013 (2016).

206 Aisaka, K., Igarashi, A. & Uwajima, T. Purification, crystallization, and characterization of neuraminidase from Micromonospora viridifaciens. *Agricultural and biological chemistry* **55**, 997-1004 (1991).

207 Kenyon, C., Colebunders, R. & Crucitti, T. The global epidemiology of bacterial vaginosis: a systematic review. *Am J Obstet Gynecol* **209**, 505-523, doi:10.1016/j.ajog.2013.05.006 (2013).

208 Sewankambo, N. *et al.* HIV-1 infection associated with abnormal vaginal flora morphology and bacterial vaginosis. *Lancet* **350**, 546-550, doi:10.1016/s0140-6736(97)01063-5 (1997).

209 Fredricks, D. N., Fiedler, T. L. & Marrazzo, J. M. Molecular identification of bacteria associated with bacterial vaginosis. *N Engl J Med* **353**, 1899-1911, doi:10.1056/NEJMoa043802 (2005).

210 Srinivasan, S. *et al.* Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. *PLoS One* **7**, e37818, doi:10.1371/journal.pone.0037818 (2012).

211 McClelland, R. S. *et al.* Evaluation of the association between the concentrations of key vaginal bacteria and the increased risk of HIV acquisition in African women from five cohorts: a nested case-control study. *Lancet Infect Dis* **18**, 554-564, doi:10.1016/S1473-3099(18)30058-6 (2018).

Appendix I – Supplementary data

Table S4.1: Distribution of cytosines in whole genome sequencing reads of GP 74

Number of Cs	Number of reads	Percentage
C_8	0	0
C_9	9	10
C_{10}	69	78
C_{11}	9	10
C_{12}	1	0.01

Table S4.2: Distribution of cytosines in whole genome sequencing reads of GP 94

Number of Cs	Number of reads	Percentage
C_8	3	3
C_9	6	7
C_{10}	67	74
C_{11}	14	16
C_{12}	0	0

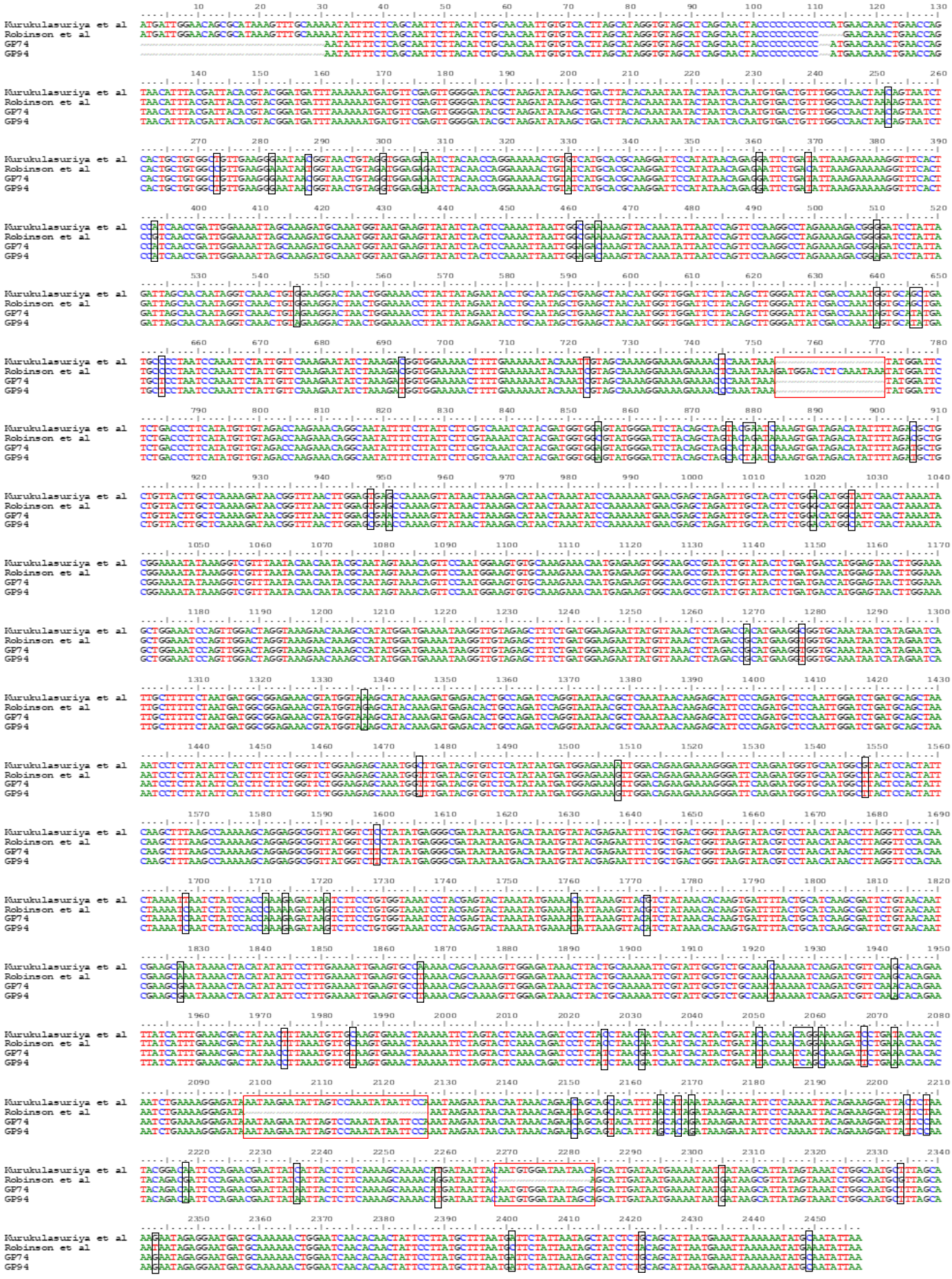


Figure S4.1: Alignment of GP 74 and GP 94 to NanH3 from Kurukulasuriya, *et al.* [10] and Robinson, *et al.* [1]. Mutations are outlined by a black box and insertions/deletions are outlined in red.

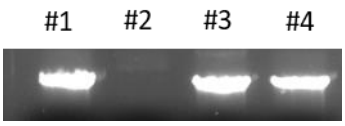


Figure S5.1: PCR of *nanh3* cloned from GP 74. Three positive colonies and one negative colony was miniprep'd and PCR amplified using cloning nanh3 cloning primers.

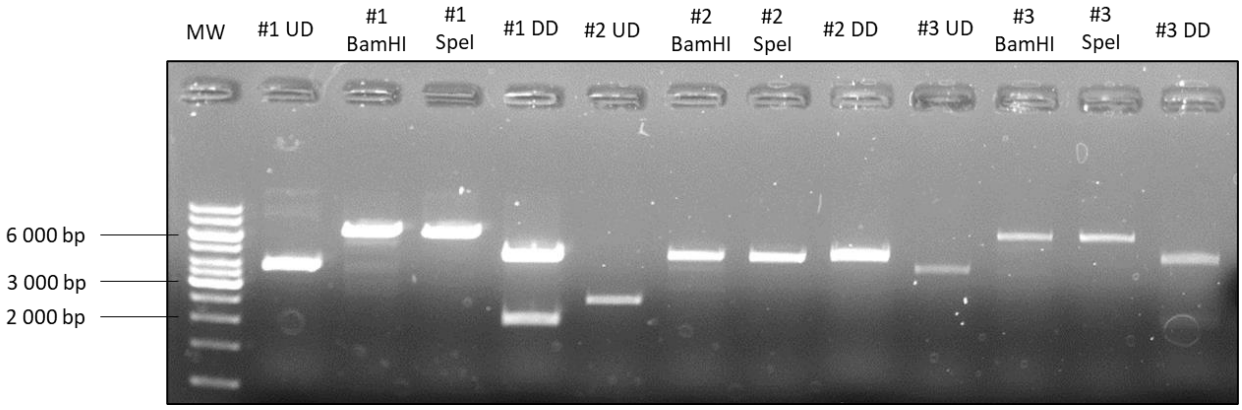


Figure S5.2: Restriction enzyme digest of miniprep'd clones. Three colonies were selected and digested with BamHI and SpeI. *nanh3* is approximately 2 000 bp. The pProEx-HTa vector is approximately 4 700 bp. The double digests for # 1 and # 3 release a 2 000 bp fragment, corresponding to the size of *nanh3* while # 2 lacks this band. MW = molecular weight marker, UD = undigested, BamHI = single digest with BamHI, SpeI = single digest with SpeI, DD = double digest with BamHI and SpeI.

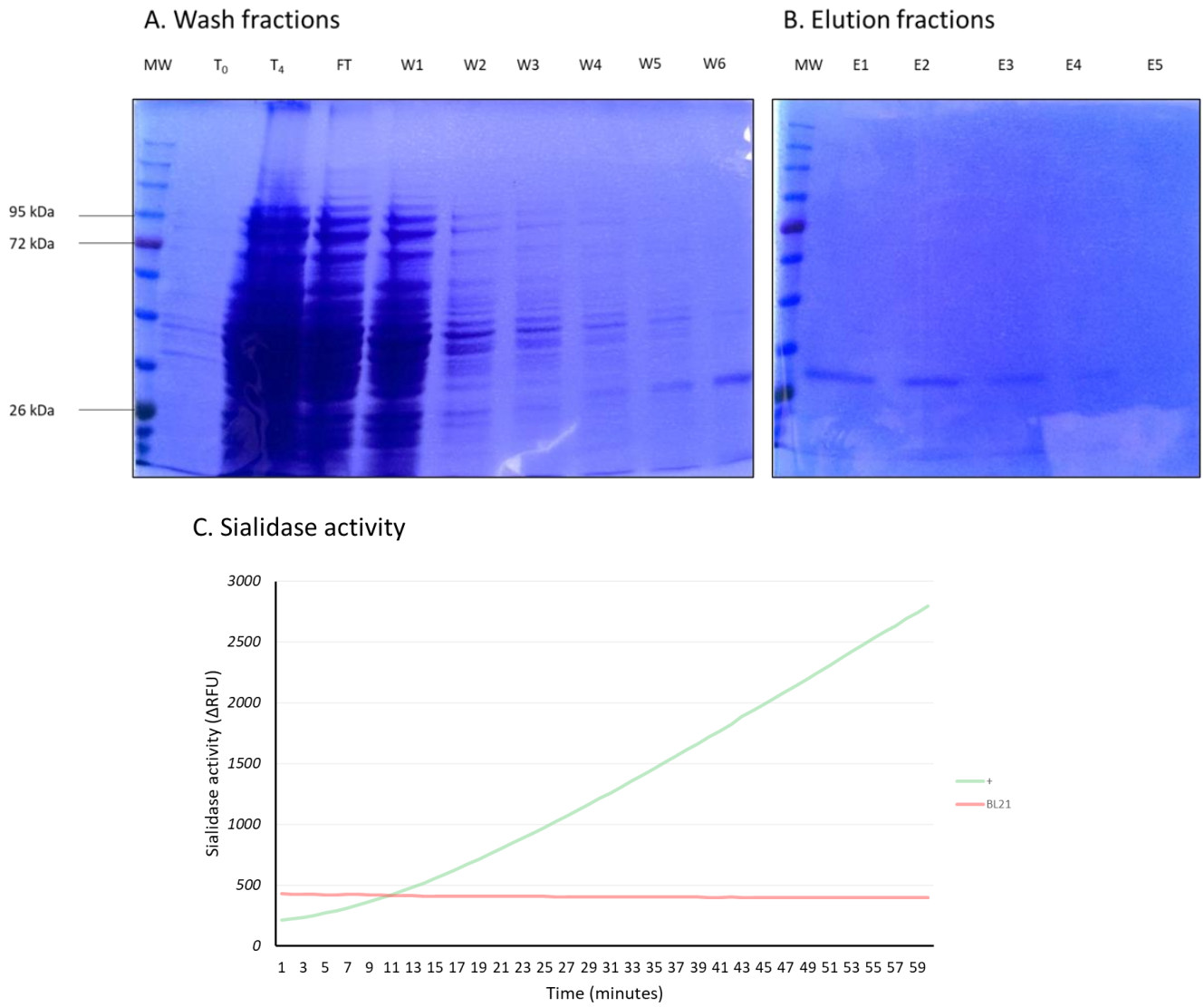


Figure S5.3: Untransformed BL21 column purified samples. A) Coomassie stain of column purified BL21 cells. T_0 – uninduced T_4 – IPTG induced, FT – flow through, W1 – W6 – washes 1 – 6. B) Coomassie stain of eluted BL21 fractions. E1 – E5 – elutions 1 – 5. The Coomassies show that the same 26 kDa band is present in the BL21 untransformed samples and eluted fractions, suggesting that it is a BL21-specific protein. C) Sialidase assay conducted on BL21 eluted fraction. "+" – 10 U of *C. perfringens* sialidase. BL21 sample shows no sialidase activity.

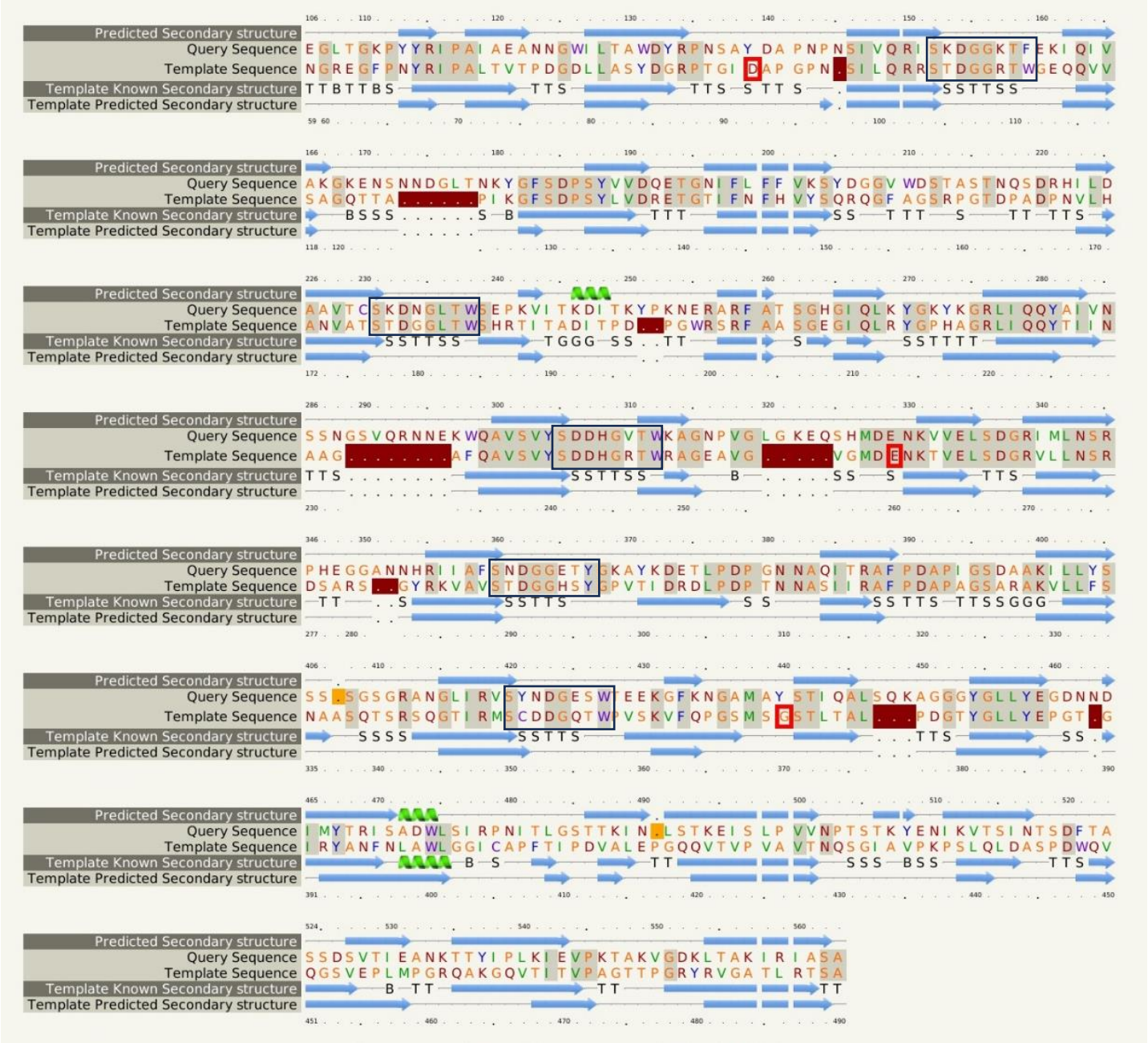


Figure S5.4. Alignment of GP 74 with *M. viridifaciens* sialidase. The protein sequence of GP 74 (query sequence) was submitted to the protein homology/analogy recognition engine (PHYRE V2.0) [8] and the sialidase of *M. viridifaciens*, with the closest structural identity and coverage, was used as a template to determine the predicted structure of NanH3. The Asp-boxes are indicated by black boxes.

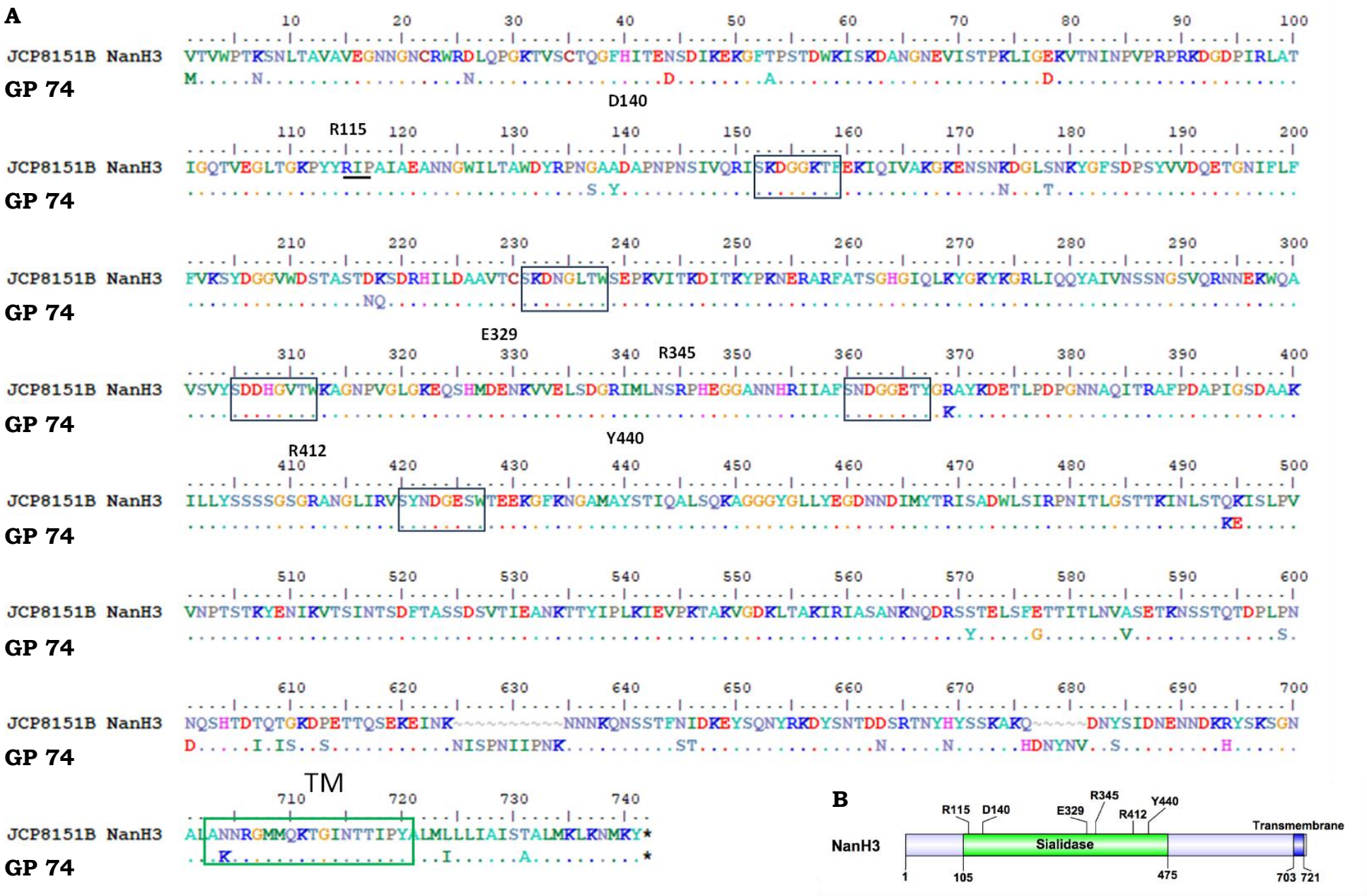


Figure S5.5.: Alignment of JCP8151B NanH3 with GP 74 NanH3 protein sequences and schematic of NanH3 protein. A) Protein sequence comparison of GP 74 NanH3 with JCP8151B NanH3 shows conservation of RIP domain, ASP boxes (outlined in black) and active sites. Transmembrane domain outlined in green with one amino acid change. B) Schematic diagram of the NanH3 protein from Robinson, *et al.* [1].

Appendix II – Solutions

New York City (NYC) III culture medium (1L)

4 g Hepes

15 g Proteose peptone No. 3

5 g NaCl

875 ml dH2O

5 g glucose

1.9 g yeast extract

100 ml heat inactivated horse serum

New York City (NYC) III agar plates (1L)

1 L NYC II culture medium

15 g agar

Luria broth (1L)

10 g Tryptone

5 g Yeast extract

10 g NaCl

H_2O up to 1 L

Luria agar (1L)

1 L Luria broth

15 g agar

6 x DNA loading dye (10 ml)

0.025 g Bromophenol Blue

0.025 g Xylene Cyanol

6 mL 50% glycerol

H_2O up to 4 ml

10 x TBE (1L)

108 g Tris

55 g Boric acid

9.3 g EDTA

Adjust pH to 8.5 with HCl

H_2O up to 1 L

Wash buffers (1L)

50 mM NaH_2PO_4

300 mM NaCl

20, 30 or 50 mM imidazole

Adjust to pH 8.0 with NaOH

H_2O up to 1 L

Equilibration buffer (1L)

50 mM NaH_2PO_4

300 mM NaCl

10 mM imidazole

Adjust to pH 8.0 with NaOH

H_2O up to 1 L

5 % stacking gel (5 ml)

1 ml 30 % Bis-acrylamide

0.625 ml 1 M Tris-Cl (pH 6.8)

25 µl 20 % SDS

30 µl 20 % Ammonium persulphate

10 µl TEMED

H_2O up to 5 ml

8 % resolving gel (7.5 ml)

2 ml Bis-acrylamide

2.8 ml 1 M Tris-Cl (pH 8.8)

37.5 µl 20 % SDS

40 µl 20 % SDS

20 µl TEMED

H_2O up to 7.5 ml

5 x protein loading dye (10 ml)

1 ml 1 M Tris-Cl (pH 6.8)

2.5 ml 20 % SDS

4 ml 50 % glycerol

0.01 g Bromophenol blue

H_2O up to 10 ml

Transfer buffer (1L)

3.025 g Tris-Cl

14.4 g glycine

200 ml methanol

H_2O up to 1L

10 x running buffer (1L)

30.3 g Tris-Cl

144 g glycine

10 g SDS

Adjust to pH 8.3 with HCl

H_2O up to 1L

Coomassie stain (100 ml)

50 ml methanol

10 ml acetic acid

1 g Coomassie Brilliant Blue

H_2O up to 100 ml

Destaining solution (500 ml)

125 ml ethanol

50 ml acetic acid

H_2O up to 500 ml

10 x TBS (1L)

117 g NaCl

60.6 g Tris-Cl

Adjust to pH 7.4 with HCl

H_2O up to 1 L

10 x TBS-T (1L)

1 L 10 x TBS

1 ml Tween-20

Blocking buffer (50 ml)

2.5 g skim milk powder

50 ml TBS-T