



UNIVERSITY OF STELLENBOSCH

**“Development of *Pichia pastoris* as a  
production system for HPV16 L1 virus-like  
particles as component to a subunit  
vaccine”**



Presented by

**Lara Kotzé**

Thesis submitted in partial fulfillment of the requirements for the degree  
of

**Masters of Science in Engineering  
(Chemical Engineering)**

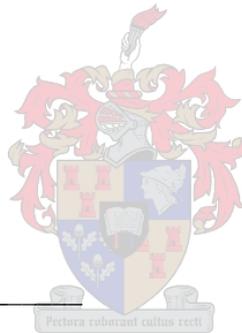
*in the Department of Process Engineering  
at the University of Stellenbosch*

Supervised by:  
Dr J.F. Görgens and Prof. W. H. van Zyl  
March 2007

## DECLARATION

I , the undersigned hereby declare that the work contained in this document is of my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

The experiments in this thesis constitute work carried out by the candidate unless otherwise stated and complies with the stipulations set out for the degree of Masters of Science in Process Engineering, by the University of Stellenbosch.



L. Kotzé: \_\_\_\_\_

Date: March 2007

L. Kotzé  
Department of Process Engineering  
University of Stellenbosch;  
Student Number: 13373765  
Stellenbosch 7600,  
South Africa

## **I ACKNOWLEDGEMENTS**

I would like to acknowledge the following people for contributing greatly to this study and my personal development:

Prof W.H van Zyl, for the use of the facilities in his laboratory and valuable insight into my project. Thank you for instilling in your students the value of being balanced individuals and for recognizing the learning opportunity presented by every obstacle.

Dr J.F Görgens, for immense support, guidance and a contagious enthusiasm for research. I could not have asked for better supervisors.

Dr Riaan den Haan, for most interesting conversations and invaluable discussions. Thank you for the wit that brightened many a dull day.

Laboratory colleagues (Microbiology and Biochemistry), for imparting advice and techniques necessary for this study.

My parents, sister and friends, for encouragement throughout and without whom none of this would have been possible.

My grandparents who never ceased praying for the small things.

The NRF and the Department of Process Engineering, for financial support during this study.

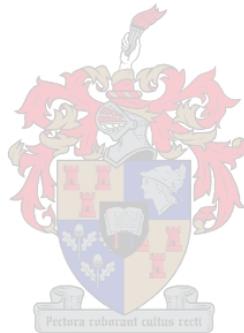
Lastly, unto my Father in Heaven, the creator of all things, the glory and praise forever.

## II SYNOPSIS

Human papillomavirus (HPV) is a sexually transmitted virus and known precursor to cervical cancer, the second most lethal cancer in females across the world. Two virus-like particle (VLP) vaccines exist that provide immunity against the main serotypes of the disease and are produced in *Saccharomyces cerevisiae* (*S. cerevisiae*) and baculovirus infected insect cells. *Pichia pastoris* (*P. pastoris*) was chosen as an alternative expression system for HPV VLP production based on its history as prolific heterologous protein producer that circumvent many of the problems associated with aforementioned expression systems. The strongly inducible *AOX* promoter allows three-phase fermentations (1.3 L bioreactors) in which high cell densities ( $>100\text{g}_{\text{CDW}}\cdot\text{L}^{-1}$ ) are obtained prior to induction with methanol. During the induction phase the dissolved oxygen concentration may be used to control addition of methanol. It is also possible to use predetermined methanol feed rates and to adjust the amount of additional oxygen sparged to maintain a constant dissolved oxygen level. The effects of these control strategies, different gene constructs and multiple gene integrations were quantified through monomer-, VLP- and mRNA production levels.

Increased biomass concentrations in the 20% dissolved oxygen control strategy led to the highest volumetric VLP concentration ( $68.53\ \mu\text{g}\cdot\text{L}^{-1}$ ). VLPs were located intracellularly in both the cytoplasm and membranes of the yeast cells. Despite lower codon adaptation of the h-L1 gene expressed in the X33[h-L1] strain it still had higher volumetric VLP concentrations under 40% dissolved oxygen control than the X33[Syn-L1] and X33[SA-L1] strain containing the SA-L1 and Syn-L1 genes. This was ascribed to the possible presence of rare codons in the Syn-hL1 and SA-L1 genes and a lower A+T content in the h-L1 gene. Multiple gene integrations of the h-L1 gene had a negative effect on VLP production and this conclusion was supported by lower mRNA concentrations indicating lower transcriptional efficiency. Increased methanol induction efficiency in the DO control strategies was indicated by higher specific L1 monomer levels. Decreased VLP to monomer ratios in the DO control strategies indicated that a bottleneck existed in the assembly process due to increased L1 monomer concentrations.

Due to the hydrophobic region on the L1 protein, these proteins associated with the membranes within the yeast cells especially when efficient assembly to VLPs did not occur. HPV16 L1 VLP concentrations obtained in *P. pastoris* in this study are comparable to the study by Li et al., (2003), but much lower than expression levels obtained in baculovirus infected insect cells. Based on the expression levels of HBsAg VLPs obtained in *P. pastoris*, this system, with the necessary recommended optimisation, has the capacity for increased HPV VLP production ability.

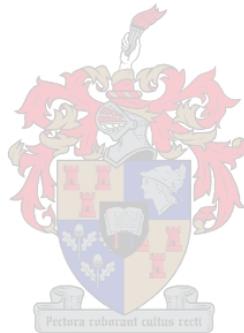


## OPSOMMING

Menslike papillomavirus is 'n seksueel oordraagbare virus wat wel bekend is as 'n voorloper van servikale kanker in vrouens; die tweede grootste oorsaak van dood agv kanker in vrouens reg oor die wêreld. Huidiglik bestaan daar twee tipes pseudovirale partikel (VLPs) entostowwe wat infeksie met menslike papillomavirus voorkom. Die entostowwe word in *Saccharomyces cerevisiae* (*S. cerevisiae*) en baculo-virus geïnfekteerde insek selle geproduseer. *Pichia pastoris* (*P. pastoris*) is gekies as alternatiewe uitdrukking sisteem vir papillomavirus pseudovirale partikel produksie as gevolg van die uitstekende resultate wat al voorheen met die sisteem verkry is, asook die feit dat dit hindernisse, wat ondervind word met bogenoemde sisteme, omseil. Die kragtige induseerbare AOX promotor baan die weg vir drie fase fermentasies (1.3 L bioreaktors) waarin hoë digtheid biomassa ( $>100\text{g}_{\text{CDW}}\cdot\text{L}^{-1}$ ) verkry word voor induksie met methanol geskied. Gedurende die induksie fase kan die opgeloste suurstof konsentrasie van die fermentasie kultuur gebruik word om die metanol voer te beheer. Dit is ook moontlik om vooraf vasgestelde voertempos te gebruik om met behulp van ekstra suurstof 'n konstante opgeloste suurstofvlak te handhaaf. Die invloed van die voorafgaande beheer strategieë, asook die effek van drie verskillende geen tipes en meervoudige geen integrasies word bepaal deur middel van monomeer-, VLP- en mRNA produksie vlakke.

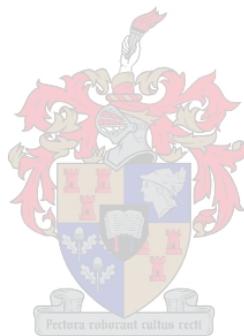
Verhoogde biomassa konsentrasies in die 20% opgeloste suurstof beheer fermentasies het die hoogste volumetriese pseudovirale partikel konsentrasie ( $68.53\ \mu\text{g}\cdot\text{L}^{-1}$ ) tot gevolg gehad. Die pseudovirale partikels was intrasellulêr geleë in beide die sitoplasma en membrane van die gis selle. Ten spyte van laer kodon aanpassingswaardes vir die h-L1 geen, het die X33[h-L1] ras steeds met hoër volumetriese pseudovirale partikel konsentrasies beter gevaar onder 'n 40% opgeloste suurstof konsentrasie as die Syn-L1 en SA-L1 geen bevattende rasse. Hierdie tendens word toegeskryf aan die moontlike teenwoordigheid van raar kodons in die Syn-L1 en SA-L1 gene en laer konsentrasies A+T areas in die h-L1 geen. Meervoudige geen integrasies het nie 'n verhoging in pseudovirale partikel produksie vlakke tot gevolg gehad nie en hierdie gevolgtrekking word ondersteun deur laer mRNA vlakke wat laer transkripsie effektiwiteit impliseer.

Verhoogde induksie effektiwiteit onder opgeloste suurstof beheer strategieë het hoër L1 monomeer konsentrasies tot gevolg gehad. Verlaagde pseudovirale partikel tot monomeer verhoudings het aangedui dat daar 'n bottelnek ondervind word met samestelling van die pseudovirale partikel weens hoër monomeer konsentrasies. As gevolg van die hidrofobiese areas eie aan die L1 monomeer het die L1 proteïen met die membrane van die gis selle geassosieer veral wanneer die samestellings van die pseudovirale partikels problematies was. Die HPV16 L1 pseudovirale partikel konsentrasies verkry in die *P. pastoris* sisteem in hierdie studie, is vergelykbaar met 'n studie gepubliseer deur Li *et al.*,(2003) maar wel steeds laer as produksie vlakke in baculovirus geïnfekteerde insek selle. Op grond van die uitdrukkingsvlakke wat veral met HBsAg pseudovirale partikels in *P. pastoris* bereik is, het hierdie sisteem die potensiaal om met die nodige verbeterings verhoogde HPV pseudovirale partikel produksie vlakke te lewer.



### **III TABLE OF CONTENTS**

I ACKNOWLEDGEMENTS_____	<u>I</u>
II SYNOPSIS_____	<u>II</u>
III TABLE OF CONTENTS_____	<u>VII</u>
IV LIST OF ABBREVIATIONS_____	<u>X</u>
V LIST OF FIGURES_____	<u>XII</u>
VI LIST OF TABLES_____	<u>XIV</u>

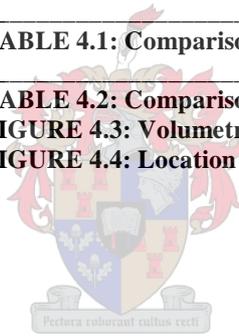


## **TABLE OF CONTENTS**

<b>1. INTRODUCTION</b>	<b>1</b>
1.1 Background	2
1.2 Aim of project	3
1.3 Implications of study	3
<b>2. LITERATURE REVIEW</b>	<b>4</b>
2.1 Introduction	4
2.2 Prevalence of Cervical cancer	4
2.3. Prevalence of Human Papillomavirus	5
2.4 The structure of the Human Papillomavirus	7
2.5 Vaccines	9
2.6 VLP production systems	14
2.7 Microbial hosts	15
2.7.1 <i>Pichia pastoris</i>	16
2.7.2 Expression levels	17
2.8 Factors affecting cultivation of <i>P.pastoris</i>	19
2.8.1 Methanol metabolism	20
2.8.2 Alternative promoters	22
2.8.3 Methanol-utilizing phenotypes	23
2.8.4 Multiple gene integrations	24
2.8.5 The role of codon bias on heterologous protein expression	26
2.8.6 Reactor requirements	28
2.8.6.1 Feeding strategies	29
2.8.6.1.1 Fed-batch feeding strategy	29
2.8.6.2 Induction with methanol	30
2.8.6.2.1 Constant, linear or exponential feed rates.	30
2.8.6.2.2 Dissolved oxygen control	31
2.8.6.2.3 Constant methanol concentration by methanol sensor and control loop.	33
2.8.7 Nutrients	37
2.8.8 Proteolysis and the temperature effect	38
<b>3. EXPERIMENTAL PROCEDURES</b>	<b>40</b>
3.1 Materials and Methods	40
3.1.1 Strains and plasmids	40
3.1.2 Media and culture conditions	40
3.1.3 Plasmid construction	41
3.1.4 Transformation	41

3.1.5 Shake flask cultivation	42
3.1.6 Bioreactor cultivations	43
3.1.6.1 Glycerol batch phase	44
3.1.6.2 Glycerol fed-batch phase	44
3.1.6.3 Methanol fed-batch phase	44
3.1.6.3.1 Methanol feeding for adaptation phase	44
3.1.6.3.2 Methanol feeding with DO control	45
3.1.6.3.3 Dissolved oxygen measurement	45
3.1.7 Analytical methods	46
3.1.7.1 Measurements of optical density and cell dry weight	46
3.1.7.2 Preparation of cell lysates	46
3.1.7.3 Determination of HPV monomer and VLP levels	46
3.1.7.4 Total protein determination	47
3.1.7.5 Analysis of m-RNA	47
3.1.7.5.1 Total RNA isolation	47
3.1.7.5.2 Slotblot	47
3.1.8 Plate $\beta$ -Galactosidase activity assays	48
<b>3.2 Experimental philosophy</b>	<b>48</b>
3.2.1 Type of cultivation and mode of operation	48
3.2.2 Process factors influencing product yield and the effect on experimental development	49
3.2.3 Experimental Design	50
3.2.4 Experimental analysis	51
<b>4. RESULTS</b>	<b>54</b>
4.1 Introduction	54
4.2 Typical cultivation analysis	54
4.3 Comparison of the effect of different strain constructs and gene dosage	55
4.4 Comparison of the effect of methanol addition strategy on the X33[h-L1] strain	58
4.5 Volumetric production levels	60
4.6 Assembly efficiency and location of the VLPs	61
4.7 Shake flask experiments	62
4.8 $\beta$ -galactosidase plate assays	63
<b>5. DISCUSSION</b>	<b>64</b>
5.1 Introduction	64
5.2 Comparison of strains with different gene constructs	64
5.2.1 Codon adaptation, G+C content and the effect of Mut phenotype	64
5.2.2 The effect of multiple gene integrations	66
5.3 The influence of methanol addition control strategy	67
5.4 Location and assembly efficiency of VLPs	68

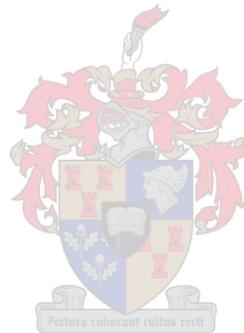
5.5 Shake flasks	70
5.6 $\beta$ -galactosidase plate assays	71
5.7 Benchmarking with alternative recombinant expression systems	71
<b>6. CONCLUSIONS</b>	<b>73</b>
<b>7. RECOMMENDATIONS AND FUTURE PROSPECTS</b>	<b>75</b>
7.1 Molecular modifications	75
7.2 Process modifications	75
<b>8. REFERENCES</b>	<b>77</b>
<b>9. APPENDIX</b>	<b>85</b>
9.1 ANOVA	85
9.1.1 ANOVA Statistical analysis of <b>TABLE 4.1: Comparison of strains and gene dosage effects</b>	85
9.1.2 ANOVA Statistical analysis of <b>TABLE 4.2: Comparison of different control strategies</b>	86
9.1.3 ANOVA Statistical analysis of <b>FIGURE 4.3: Volumetric production levels</b>	87
9.1.4 ANOVA Statistical analysis of <b>FIGURE 4.4: Location of VLPs</b>	87



#### IV LIST OF ABBREVIATIONS

Abbreviation	Full text
AOX	Alcohol Oxidase
BMGY	Buffered Minimal Gliserol medium
BMMH	Buffered Minimal Methanol medium
BSA	Bovine Serum Albumin
CAI	Codon Adaptation Index
CANSA	Cancer Association of Southern Africa
CHO	Chinese Hamster Ovary cells
CIN	Cervical Intraepithelial Neoplasia
CMF	Constant Methanol Feed rate
DNA	Dioxyribonucleic acid
DO	Dissolved Oxygen
EIP	Elastase Inhibiting Peptide
ELISA	Enzyme Linked Immunosorbent Assays
FDA	Food and Drug Administration, USA Government
FLD	Formaldehyde Dehydrogenase promoter
GAP	Glyceraldehydes-3-phospate dehydrogenase promoter
GMP	Good Manufacturing Practice
GSK	GlaxoSmithKline
HBsAg	Hepatitis B surface antigen
HBV	Hepatitis B virus
HIV	Human Immunodeficiency Virus
HPLC	High Performance Liquid Chromatography
HPV	Human Papillomavirus
IARC	International Agency for Research on Cancer
mRNA	messenger Ribonucleic acid
Mut	Methanol Utilisation Type
NBS	New Brunswick Scientific
OEM	Oxygen Enriching Membrane

OTR	Oxygen Transfer Rate
PMF	Predetermined Methanol Feed rate
PVA	Polivinyalcohol
VLP(s)	Virus-like Particle(s)
vvm	volume of inlet gas flow per volume of culture per minute
WHO	World Health Organisation

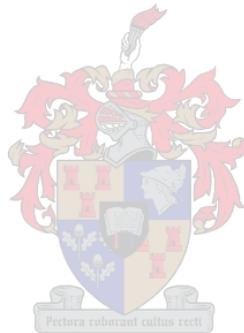


## V LIST OF FIGURES

<b>FIGURE 2.1:</b> Incidence of cervix uteri cancer (Parkin, 2001). Legend indicates percentage incidence of cervical cancer across the world. ....	<b>5</b>
<b>FIGURE 2.2:</b> (A) A model of the papillomavirus capsid. The pentameres contain 5 molecules of L1. One molecule of L2 protein fits into the central dimple. (B) Papillomavirus particles containing DNA. (C) HPV <sup>16</sup> L1 virus-like particles expressed in baculovirus. Morphologically similar to particles seen in (B). (Stanley <i>et al.</i> , 2006).....	<b>7</b>
<b>FIGURE 2.3:</b> Side view (A) and internal structure (B) of HPV L1 pentameric capsomer. (Chen <i>et al.</i> , 2000).....	<b>8</b>
<b>FIGURE 2.4:</b> Hydrophobicity of the HPV16 L1 protein.....	<b>9</b>
<b>FIGURE 2.5:</b> Methanol metabolism in <i>P. pastoris</i> (Houard <i>et al.</i> , 2002).....	<b>21</b>
<b>FIGURE 2.6:</b> Schematic representation of fermentation strategy.....	<b>30</b>
<b>FIGURE 3.1:</b> Southern blot showing four integrated copies of the h-L1 gene (Lane 4).....	<b>42</b>
<b>FIGURE 3.2:</b> Plasmid construction of pPICZ-A plasmids containing the three alleles of the L1 genes.....	<b>42</b>
<b>FIGURE 3.3:</b> Experimental setup of two 1.3 L NBS bioreactors.....	<b>49</b>
<b>FIGURE 3.4:</b> Experimental design strategy including both shake flask and bioreactor cultivations.....	<b>51</b>
<b>FIGURE 4.1 :</b> Typical fermentation analysis of the X33[h-L1] strain under 40%DO control. CO <sub>2</sub> (%) [■], Methanol addition (mL) [◆], Methane (%) [×], HPV16 VLP Production [□], Cell dry weight (g <sub>DCW</sub> .L <sup>-1</sup> ) [○].....	<b>55</b>
<b>FIGURE 4.2:</b> mRNA slot blots comparing mRNA concentration between X33[h-L1] and X33[multi-h-L1]. Top: Bands 1-3 are a triplicate of the X33[h-L1] 40%DO control strategy mRNA followed by bands 4-6 of X33[multi-h-L1] mRNA under the same control strategy. Bottom: Comparison of <i>ACT1</i> concentrations for the corresponding X33[h-L1] and X33[multi-h-L1] strain mRNA. Band number 7 is the positive control for the h-L1 mRNA at the top and <i>ACT1</i> at the bottom.....	<b>58</b>
<b>FIGURE 4.3:</b> Comparison of total volumetric HPV16 L1 VLP concentrations of different strains and cultivation strategies.....	<b>61</b>

**FIGURE 4.4:** Specific HPV16 L1 VLP concentration in cytoplasmic and membrane fraction. (▣)Cytoplasmic fraction, (▣)Membrane fraction.....**62**

**FIGURE 4.5:**  $\beta$ -galactosidase plate assays (blue zones indicate induction occurring and  $\beta$ -galactosidase being produced).....**63**



## VI LIST OF TABLES

<b>TABLE 2.1:</b> Classification of HPV types by cervical oncogenicity (Munoz <i>et al.</i> , 2003).....	<b>6</b>
<b>TABLE 2.2:</b> Percentage HPV16 prevalence in Africa.....	<b>6</b>
<b>TABLE 2.3:</b> Summary of vaccine types and advantages and disadvantages related to them ( Adapted from Kolls and Sherris, 2000).....	<b>13</b>
<b>TABLE 2.4:</b> HPV VLP expression systems.....	<b>15</b>
<b>TABLE 2.5:</b> Intracellular heterologous proteins expressed in <i>P. pastoris</i> .....	<b>18</b>
<b>TABLE 2.6:</b> Extracellular heterologous proteins expressed in <i>P. pastoris</i> .....	<b>19</b>
<b>TABLE 2.7:</b> The advantages and disadvantages of the <i>AOX</i> promoter system (Macauley-Patrick <i>et al.</i> , 2005).....	<b>21</b>
<b>TABLE 2.8:</b> Successful increase in expression with increased gene dosage.....	<b>24</b>
<b>TABLE 2.9:</b> Improvement of expression levels through codon optimisation adapted from Gustafsson <i>et al.</i> , (2004).....	<b>27</b>
<b>TABLE 3.1:</b> Strain constructs with the pPICZ-A plasmid .....	<b>40</b>
<b>TABLE 3.2:</b> Parameters varied in cultivations.....	<b>52</b>
<b>TABLE 4.1:</b> Comparison of the effect of different strain constructs and gene dosage under 40%DO control.....	<b>57</b>
<b>TABLE 4.2:</b> Comparison of the effect of methanol addition strategy on the X33[h-L1] strain.....	<b>60</b>
<b>TABLE 9.1.1:</b> ANOVA Statistical analysis of TABLE 4.1: Comparison of strains and gene dosage effects.....	<b>85</b>
<b>TABLE 9.1.2:</b> ANOVA Statistical analysis of TABLE 4.2: Comparison of different control strategy.....	<b>86</b>
<b>TABLE 9.1.3:</b> ANOVA Statistical analysis of FIGURE 4.3: Volumetric production levels.....	<b>87</b>
<b>TABLE 9.1.4:</b> ANOVA Statistical analysis of FIGURE 4.4: Location of VLPs.....	<b>87</b>

## 1. Introduction

Human papillomavirus (HPV) is a sexually transmitted virus that affects mainly the female population and is one of the leading causes of the development of cervical cancer (IARC, 1999). Cervical cancer is most prevalent in less developed countries (Parkin, 2001) and although good screening programs should be sufficient in identifying and preventing cervical cancer, these programs are not readily available in poorer countries. The only viable solution to preventing HPV and therefore cervical carcinoma in poorer countries is the development of a cost effective vaccine against HPV that is widely accessible to less developed countries.

The virus, which causes lesions that progress to cervical carcinoma, cannot be propagated in tissue culture and therefore no attenuated or live virus vaccines are available. Recombinant gene technology has provided the opportunity to express virus-like particles in various recombinant expression systems. Both prokaryotic and eukaryotic systems have been used in this way. They are bacterial, insect and mammalian cell cultures, yeasts and filamentous fungal systems as well as transgenic plants.

Currently Merck and GSK have vaccines on the market to combat 4 serotypes of HPV. Unfortunately the costs of these vaccines are prohibitive towards general distribution to developing countries as one dose of the vaccine (of which three are necessary to induce sufficient protection) can cost up to US\$120 (Merck, 2006).

The development of a recombinant system that can produce a virus-like particle vaccine at reduced cost and in larger volumes will go a long way towards increasing availability of a HPV vaccine. Many of the limitations of the current production systems might be overcome by the use of a different system. These limitations might include glycosylation of the VLP, costly substrates and rich media as well as long production times and low yields.

## 1.1 Area of focus

In this study *Pichia pastoris* (*P. pastoris*) is investigated as a possible heterologous host for the production of HPV16 L1 proteins that self-assemble into virus-like particles (VLPs). These VLPs induce immunity to the virus when introduced to the human system in the form of a vaccine. *P. pastoris* is the subject of this research firstly due to ease of genetic manipulation and transformation that is possible in this system (Invitrogen). Secondly, high density cell cultures can be obtained through relatively simple and cost efficient fed-batch fermentation on a defined medium (Stratton *et al.*, 1998). Thirdly, expression of the protein in question is under tight control of the strong *AOX* promoter that is induced with methanol (Cregg *et al.*, 1987). *P. pastoris*, *Hansenula polymorpha* (*H. polymorpha*) and *Candida boidinii* (*C. boidinii*) are three yeasts that can assimilate methanol as its only carbon source and have shown great potential as heterologous hosts (Anthony, 1982).

Although HPV16 L1 VLPs have been produced in Chinese hamster ovary cells, insect cells, *Saccharomyces cerevisiae* (*S. cerevisiae*), *Schizosaccharomyces pombe* (*S. pombe*) and potato and tobacco plants, the only instance of HPV being produced in *P. pastoris* was HPV6 (Li *et al.*, 2003). Research into HPV16 L1 VLPs is therefore exploratory and it still needs to be determined whether expression is possible in *P. pastoris* and whether these HPV16 L1 monomers and VLPs can correctly assemble to induce the correct immunological response.

If expression and assembly does take place, the question remains whether expression levels obtainable in *P. pastoris* are high enough to justify further investigation into *P. pastoris* as a HPV16 L1 VLP producer. The main competition for this expression system at the moment is insect cells with baculovirus vectors, *S. cerevisiae* and expression of HPV16 L1 VLPs in transgenic plants. It will have to be proven that expression levels obtained with *P. pastoris* are indeed sufficient to be a competitor in this field.

## 1.2 Objectives of project

The objectives and aims of this project serve to direct the investigation of *P. pastoris* as a microbial host for HPV16 L1 production toward answering the research questions as mentioned above in the following manner:

- i) Testing the hypothesis that *P. pastoris* can express HPV16 L1 monomers and VLPs of three genes expressed in the same host strain.
- ii) Establish benchmark production levels through an in-depth literature study.
- iii) Cultivate an understanding of the *P. pastoris* expression system to facilitate adequate investigation of VLP production and in particular the effect of codon bias and gene copy number on HPV expression levels.
- iv) To optimize expression (if at all possible) through different methanol feed control strategies and compare the efficiency of multiple L1 gene integrations.
- v) Study and application of analytic techniques (ELISA, mRNA detection) that facilitate interpretation of data gathered through fermentations.

## 1.3 Implications of study

The main focus of this research is to develop *P. pastoris* as an expression platform for HPV16 VLP vaccine. This will make a significant contribution to the development of vaccine production abilities in Southern Africa. In addition, the Western Cape has become a hub for the development of vaccines.

As insect cell and mammalian cell lines have expensive media requirements and difficulty in scale-up, an alternative eukaryotic expression system such as *P. pastoris* might lead to lowering production cost and enable increased availability of vaccines to poorer countries. Producing a vaccine developed with a South African serotype of the virus might significantly increase efficacy of the vaccine and lead to better immunogenicity.

Furthermore, with *P. pastoris* firmly established as a vaccine producer other vaccine candidates can be expressed in this system, while maintaining the knowledge base of cultivation methods. This study should lead to a better understanding of the host

expression system and further breakthroughs in vaccine technology in the prevention of communicable diseases.

## **2. Literature Review**

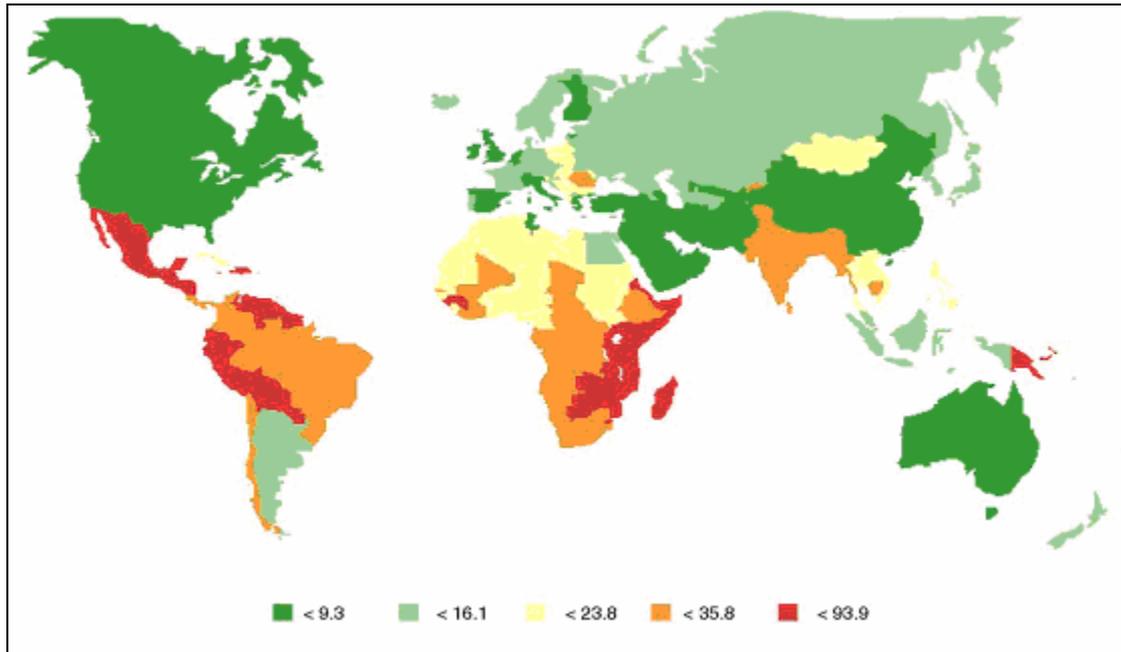
### **2.1 Introduction**

The literature review will aim to bring insight into HPV as a sexually transmitted virus and the causative link to cervical cancer in females across the world. The merits of current vaccine strategies are discussed and the state of HPV vaccines. Microbial systems and their applications in VLP production is evaluated with particular emphasis on *P. pastoris* as a heterologous host. Finally factors influencing cultivations and possible bioreactor control strategies are considered.

### **2.2. Prevalence of Cervical cancer**

Studies undertaken across the world have irrevocably shown an etiological relationship between human papillomavirus (HPV) and cervical cancer (IARC, 1999). In 1995 the World Health Organization (WHO) officially classified HPV as a human carcinogen (IARC, 1995). The etiological relationship is even more visible when one learns that the epidemiology of this cancer closely follows that of sexually transmitted diseases (Lowy *et al.*, 1994).

Cervical cancer is the second largest cause of shortened life span in females across the world (Parkin, 2001), and the most common cancer in black females in South Africa (Kay *et al.*, 2003). The WHO reports that it is the cause of 288 000 deaths yearly, while 510 000 cases are reported each year. Eighty percent of the cases originate in developing countries (68 000 in Africa, 77 000 in Latin America and 245 000 in south and south-east Asia). Incidence studies conducted by Parkin (2001) have shown that the highest incidence rates of cervical cancer occur in less developed countries as shown in FIG 2.1 and that this was due to the lack of screening programs for cervical cancer. Parkin (2001) also cited HPV as the most important cause of cervical cancer.



**FIGURE 2.1: Incidence of cervix uteri cancer (Parkin, 2001). Legend indicates percentage incidence of cervical cancer across the world.**

### 2.3. Prevalence of Human Papillomavirus

In the South-African study by Kay *et al.*, (2003) 82% of cervical cancer biopsies and 56.6% of cervical intraepithelial neoplasia biopsies showed evidence of HPV16 infection. This indicates that HPV16 is the most prevalent HPV type in South Africa and also accounts for 50% of global HPV infection causing cervical cancer (Bosch *et al.*, 1995).

To date 120 HPV serotypes have been identified with 18 of these being classified as high risk types associated with cervical cancer (De Villiers *et al.*, 2004). High risk HPV DNA has been associated with 99% of uterine cervix cancer (Wallboomers *et al.*, 1999). TABLE 2.1 classifies some of the serotypes into high-risk, probable high-risk and low risk groups.

**TABLE 2.1:** Classification of HPV types by cervical oncogenicity  
(Munoz *et al.*, 2003)

Classification of HPV types	
Risk Classification	HPV types
High-risk	16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, 82
Probable high-risk	26, 53, 66
Low risk	6, 11, 40, 42, 43, 44, 54, 61, 70, 72, 81, CP6108
Undetermined risk	34, 57, 83

Roden *et al.*, (1997) postulated that papillomaviruses adapted into different genotypes due to evolutionary pressure and to escape neutralisation. Due to this, HPV shows little or no cross reactivity between genotypes and therefore an effective vaccine would incorporate some of the main HPV types prevalent to a region (Kay *et al.*, 2003). In African countries HPV16 infection is higher than any other types of HPV (TABLE 2.2)., closely followed by HPV 18

**TABLE 2.2:** Percentage HPV16 prevalence in Africa

HPV16 prevalence rates in Africa		
Country	%	Reference
Algeria	33.3	Bosch <i>et al.</i> , 1995
Uganda	53.5	Bosch <i>et al.</i> , 1996
Morocco	67.7	Chauki <i>et al.</i> , 2000
Mozambique	13.0	Castellsague <i>et al.</i> , 2001

Worldwide, 57.6% of cervical cancers can be attributed to HPV16, 14.1 % to HPV 18 and 5.4 % to HPV 45 (WHO, 1999).

An aggravating factor in developing countries is that immuno-compromised individuals, such as the HIV-infected, are more prone to infection, show faster progression from lesions to cancer, are more difficult to treat and have a high rate of cancer recurring (Levi *et al.*, 2004; Baay *et al.*, 2004). One can presume that in South Africa and many under developed countries where HIV prevalence is high, HPV may increase the mortality rates due to cervical cancer.

If vaccination against HPV high risk types can be implemented, approximately 30% of infections with high risk types may be avoided, as well as 40-50% of cytological abnormalities and 50-60% of high-grade cervical intraepithelial neoplasia (CIN) within 5 to 10 years (Kaufmann and Schneider, 2006).

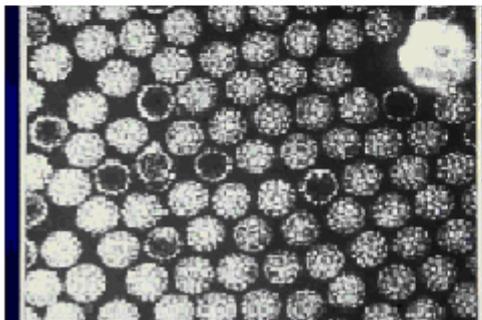
## 2.4 The structure of the Human Papillomavirus

The Human papillomavirus consist of approximately 55 nm (diameter) non-enveloped virions with icosahedral symmetry (FIG 2.3). It contains a naked protein capsid containing a ~8000 base double-stranded circular DNA genome.

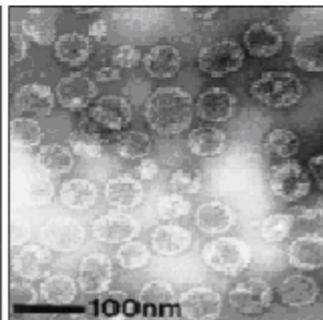
A. MODEL OF THE VIRUS-LIKE CAPSID



B. VIRUS PARTICLES

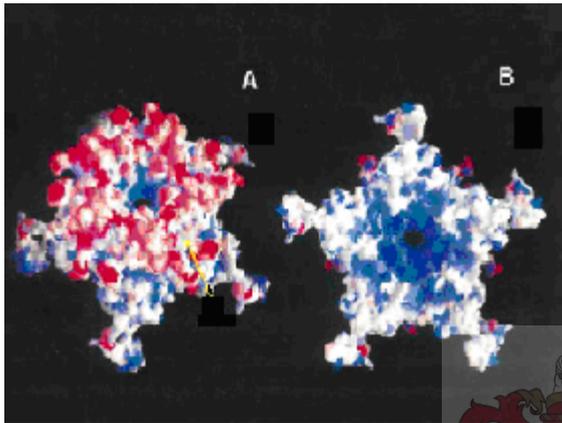


C. VIRUS-LIKE PARTICLES



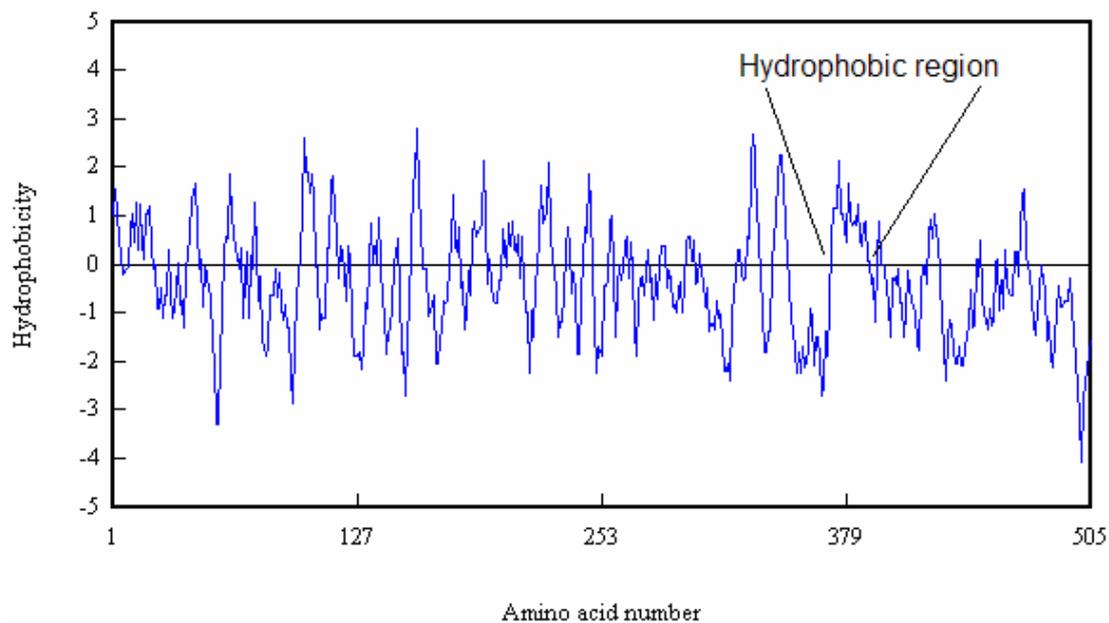
**FIGURE 2.2:** (A) A model of the papillomavirus capsid. The pentamers contain 5 molecules of L1. One molecule of L2 protein fits into the central dimple. (B) Papillomavirus particles containing DNA. (C) HPV16 L1 virus-like particles expressed in baculovirus. Morphologically similar to particles seen in (B). (Stanley *et al.*, 2006)

The capsid is made up of 72 pentameric capsomers (FIG 2.2) of the major L1 capsid protein (~55 kDa) arranged on an icosahedral lattice with between 30 and 72 copies of the minor L2 protein (Finnen *et al.*, 2003). The L2 protein is mainly an internal protein and less abundant than L1 proteins (Stanley *et al.*, 2006). They occur in a 1 to 30 ratio to the L1 protein (Chen *et al.*, 2000). The HPV L1 protein is known to self-assemble without the presence of the L2 protein. Studies have shown that co-expression with the L2 protein increases the VLP yield by about four fold in baculovirus systems (Kirnbauer *et al.*, 1993).



**FIGURE 2.3: Side view (A) and internal structure (B) of HPV L1 pentameric capsomer. (Chen *et al.*, 2000)**

The nature of the heterologous protein being expressed in the recombinant host is of great importance. Not only does it give insight into the mechanism of self-assembly of the pentameres and VLPs, but it impacts on the affinity of the protein for the cytoplasm or the membrane. With regards to this, the hydrophobicity of the HPV16 L1 protein was investigated. FIG 2.4 was obtained by inserting the protein sequence into DNAMAN. As indicated on the graph, there is a large region that is hydrophobic.



**FIGURE 2.4: Hydrophobicity of the HPV16 L1 protein**

Chen and co-workers (2000) established that L1 pentamer-pentamer contacts come from small laterally projecting domains that have a highly hydrophobic bond. Finnen *et al.*, (2003) found that there is a region on the L2 protein that is hydrophobic to facilitate adhesion of the L2 protein to the L1 pentamer. Therefore both the major and the minor protein have areas of hydrophobicity to facilitate not only L1-L1 interaction and assembly, but also L1-L2 interaction.

## 2.5 Vaccines

The only viable option for preventing the progression of HPV lesions to carcinomas is the development of safe, cost effective prophylactic vaccines (Goldie, 2002). It is suggested that a vaccine specifically targeting HPV type 16 would appreciatively lessen the occurrence of cervical cancer in South-Africa (Kay *et al.*, 2003). If one considers the impact of the hepatitis B vaccine program on the incidence of hepatocellular carcinomas, one can postulate that a HPV vaccine might follow a similar trend in preventing anogenital carcinomas (Steller, 2003). Huang and Lin (2000) published results indicating that since 1984 when a hepatitis B vaccination program was introduced in Taiwan the prevalence of childhood hepatitis was

decreased by 90% and mortality decreased by 80%. Hepatocellular carcinoma incidence has decreased four fold. To effect a similar decrease on cervical cancer incidence some vaccines have been developed with varying production yields and efficacies, and are in different stages of human phase trials. Up to the year 2005, HPV vaccine candidates were limited to proof-of-concept studies.

On 8 June 2006 a quadrivalent human papillomavirus L1 virus-like particles (types 6, 11, 16, and 18) (Gardasil<sup>TM</sup>) developed by Merck Pharmaceuticals (Villa *et al.*, 2005) (IAVI Report, 2005) was approved by the FDA ([http://www.merck.com/newsroom/press\\_releases/product/2006\\_0608.html](http://www.merck.com/newsroom/press_releases/product/2006_0608.html)). This is the first approved HPV vaccine to reach the market. Their VLP vaccine is yeast-based (*S. cerevisiae*). They have also submitted an application to the European Medicines Agency. A single dose of Gardasil<sup>TM</sup> costs US\$120. Three doses are necessary to complete the vaccination. Currently the cost is prohibitive to implementation in the developing world.

GlaxoSmithKline Biologicals in Rixsart, Belgium produced a bivalent HPV16 and HPV 18 (Cervarix<sup>TM</sup>) based on recombinant insect cell (SF9)/baculovirus technology. Phase III results were expected by the end of 2005 and an application to the European regulatory authorities for licensure was to be submitted by the first half of 2006 (IAVI Report, 2005).

A vaccine against HPV would work by stimulating a humoral or antibody mediated response against the virus at genital mucosal surfaces and inducing neutralizing antibodies that recognize and can inactivate the virus before it infects the host cells. Vaccines can be categorized into two categories, depending on the function they must perform. Prophylactic vaccines must induce antibody response to an infection while a therapeutic vaccine will stimulate cell-mediated immunity. This distinction is necessary because once the virus is incorporated into the host cell; the antibodies from an antibody mediated immunological attack cannot affect the virus anymore. A therapeutic vaccine would prevent a low grade disease from progressing, as well as causing existing lesions to regress. This reaction would control the spread of metastatic cancer and prevent recurrence of cervical cancer after appropriate treatment (Kolls and Sherris, 2000).

In general HPV cannot be propagated in tissue cultures and therefore it is not possible to develop attenuated or inactivated live virus vaccines. Genetic engineering provides a solution to this problem in the form of recombinant subunit vaccines. These subunit vaccines do not contain the viral genes, cannot induce HPV and are therefore safer and have fewer side effects.

Researchers are investigating different approaches to HPV vaccine development through recombinant genetic engineering and TABLE 2.3 is a summary of the vaccine producing technologies with benefits and disadvantages connected to them. The five types of vaccines will be discussed briefly:

- *Recombinant live vector vaccines*

A harmless host virus or bacterium is genetically engineered to produce a HPV antigen. The immune system responds to both the host organism and the HPV antigen (Kolls and Sherris, 2000).

- *Protein/ peptide vaccines*

An organism, such as yeast, is genetically modified to produce a HPV protein or peptide. After this antigen is purified, it is combined with an adjuvant that helps trigger the immune system (Kolls and Sherris, 2000).

- *DNA vaccines*

HPV genetic material is inserted into bacterial plasmids. When these circular DNA structures are used in a vaccine, the DNA is expressed in human cells that then produce an HPV antigen (Kolls and Sherris, 2000).

- *Edible vaccines*

Plants are genetically engineered to express HPV antigens in fruits and vegetables. Eating the foods leads to immunization in the gastrointestinal tract (Warzecha *et al.*, 2003, Santi *et al.*, 2006).

- *Virus like particles (VLPs) vaccines*

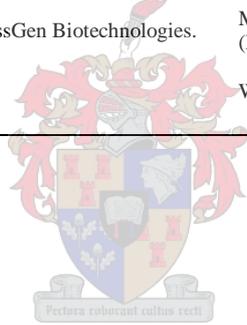
Cultured cells are genetically engineered to produce HPV capsid proteins that self assemble into empty shells resembling virus particles (Zhou *et al.*, 1991; Hagensee *et al.*, 1993; Vassileva *et al.*, 2001; Deml *et al.*, 2005).

VLP production systems have many advantages over other systems; for instance, VLP vaccines cannot multiply in the mammal, as could be the case in live or attenuated viruses. The VLP consists of the capsid proteins and does not contain the viral DNA. The VLPs are morphologically indistinguishable from the virus (Giroglou *et al.*, 2001) and therefore VLPs induce high titres of neutralizing antibodies due to the identical correct folding of monomeric proteins that induce a high epitope concentration (Grgacic and Anderson, 2006). VLPs are also easily purified, which reduces the downstream processing costs (Pattenden *et al.*, 2005).



**TABLE 2.3:** Summary of vaccine types and advantages and disadvantages related to them. (Adapted from Kolls and Sherris, 2000)

Type of Vaccine	<i>Recombinant live vector</i>	<i>Protein/peptide</i>	<i>Naked DNA</i>	<i>Edible</i>
<b>Advantage</b>	Induces a strong immune response with fewer injections and gives both antibody-and-cell mediated immunity.	Safe, ease of production and cost efficiency.	Induces cell-and-antibody mediated immunity. Simplify the production of multivalent vaccines. Less expensive, no need for a cold chain. Long shelf life.	Simple, inexpensive way to mass produce vaccines. No cold chain required. Easier to supply foodstuffs that require no infection prevention measures or trained medical staff
<b>Disadvantage</b>	Not safe to use for immuno compromised individuals. As some of the vectors are already used for other vaccines, widespread immunity to vector already exists and will have no effect as a HPV vaccine	Does not generate strong T-cell response. Difficulty in epitope isolation.	May induce cell mutations, disrupt cellular genes. Might induce anti-DNA antibodies. Might not elicit mucosal immune responses	May have to ingest large volumes to induce immunity ( if live vaccine- some part of it may degrade – exact dosage unknown).
<b>Current research</b>	Cantab Pharmaceuticals, Transgene S.A., Johns Hopkins School of Medicine, Wyeth-Lederle Vaccines & Pediatrics, CANSA with UCT, Wistar.	Cantab Pharmaceuticals, StressGen Biotechnologies.	Wyeth-Lederle Vaccines collaborating with Apollon, inc. (Malvern, USA), Merck and Vical, Inc. (Emeryville, USA), Wistar Institute.	CANSA and UCT (SA).



## 2.6 VLP production systems

Discovery of the intrinsic ability of viral coat proteins to self assemble after recombinant expression has provided scientists with a pure source of large macromolecular assemblies with unique properties. These VLPs can be isolated directly from expression in eukaryotic cells or assembled *ex vivo* from coat protein subunits purified from any recombinant hosts (Pattenden *et al.*, 2005). Recent advances in technology allow the inclusion of small particles in the capsid structure to use VLPs as a delivery system (Garcea and Gissman, 2004), making them vessels for gene and drug delivery.

Additionally, VLPs are well sized ( $\pm 40$  nm) for uptake in dendritic cells. Therefore it is less likely that adjuvants are necessary. Targeting the dendritic cells is crucial in eliciting an efficient induction of broadly neutralizing antibodies. VLPs are more immunogenic than subunit vaccines, as they can stimulate both humoral and cell-mediated immune responses (Grgacic and Anderson, 2006).

In the past VLPs have been successfully expressed through baculovirus in insect cells, Vaccinia virus in Chinese hamster ovary cells (CHOs), *Salmonella typhimurium* (Nardelli-Haeffliger *et al.*, 1997; Baud *et al.*, 2004), transgenic plants (Varsani *et al.*, 2003; Liu *et al.*, 2005; Santi *et al.*, 2006), as well as yeast systems such as *S. cerevisiae* (Neeper *et al.*, 1996) and *Schizosaccharomyces pombe* (Sasagawa *et al.*, 2005). Certain systems show advantages above others and varying production levels have been reached.

Much research has been done on VLPs as vaccines (Grgacic and Anderson, 2006) and the main drawbacks to date are affordability and vaccine stability. It should be noted that other vaccines are in less developed stages and would need more time for development and preclinical evaluation than VLP production systems. The effect of vaccines will only be seen in the next twenty years, as HPV lesions take time to progress to carcinomas.

VLPs have been successfully produced for HPV types (Lowy *et al.*, 1998). The L1 major capsid protein self-assembles into naked icosahedrons when expressed in

microbial or cellular expression systems engineered by recombinant techniques (Schiller and Lowy, 2001). In TABLE 2.4 a short summary is given of some of the expression systems used to express HPV VLPs (refer to Stanley *et al.*, (2006) and Santi *et al.*, (2006) for review articles). Already in 1992, Kirnbauer and co-workers reported that HPV L1 VLPs had successfully assembled in insect cells at high levels.

**TABLE 2.4:** HPV VLP expression systems

Human papilloma virus-like particles		
VLP	Expression system	Reference
HPV16 L1 protein	<i>Schizosaccharomyces pombe</i>	Sasagawa <i>et al.</i> , 1995
HPV16 L1 protein	Insect cells	Rose <i>et al.</i> , 1993
HPV11 L1 protein	<i>Saccharomyces cerevisiae</i>	Neeper <i>et al.</i> , 1996
HPV11 L1 protein	<i>Escherichia coli</i>	Li <i>et al.</i> , 1997
HPV16 L1 protein	<i>Salmonella typhimurium</i>	Nardelli-Haefliger <i>et al.</i> , 1997
HPV11 L1 protein	Potato plant	Warzecha <i>et al.</i> , 2003
HPV16&18 E6 and E7	Vaccinia virus	Boursnell <i>et al.</i> , 1996
HPV16 L1 protein	Insect cells	Kirnbauer <i>et al.</i> , 1992
HPV16 L1/E7 chimera	Insect cells	Greenstone <i>et al.</i> , 1998
HPV16 L1 and L2 protein	SemLiki Forest Virus	Heino <i>et al.</i> , 1995
HPV6 L1 protein	<i>Pichia pastoris</i>	Li <i>et al.</i> , 2003
HPV16 E6 and E6	Chinese hamster ovary cells	Zheng <i>et al.</i> , 2002

## 2.7 Microbial hosts

Microbial systems such as yeasts show great potential as VLP expression hosts. Yeasts are unicellular eukaryotic microorganisms in which the ease of genetic manipulation is combined with the ability to perform eukaryotic processing steps such as folding of proteins. In 1982 Valenzuela and co-workers expressed recombinant HBsAg VLPs in *S. cerevisiae*. The first commercial recombinant vaccine (HBV) was produced by this system in 1986. Much is known about this system and it has been a basis of knowledge in genetics, molecular biology and physiology. Methylotrophic yeast systems like *Hansenula polymorpha*, *Candida boidinii* and *Pichia pastoris* offer considerable improvements on difficulties often encountered in *S. cerevisiae*-based production systems, namely mitotic instability of recombinant strains, undesirable hyperglycosylation, obstacles in adapting production schemes to an industrial scale and low production levels. Studies over the last decade have conclusively shown that the methylotrophs are proficient heterologous protein producers

(Gellisen and Hollenberg, 1997). Furthermore, *P. pastoris* shows great potential as VLP expression host.

### **2.7.1 *Pichia pastoris***

Yeast expression systems have been used to produce numerous eukaryotic proteins and since 1984 many foreign proteins have been expressed in the methylotrophic yeast *P. pastoris* (Lin Cereghino and Cregg, 2000). *P. pastoris* is one of 12 different yeast species comprising of different genera capable of metabolizing methanol. The other genera are *Candida*, *Hansenula* and *Torulopsis*.

The Phillips Petroleum Company contracted Salk Institute Biotechnology/Industrial Associates Incorporated to develop *P. pastoris* as an organism for heterologous protein expression. Specifically, a strain was genetically engineered to produce single cell proteins on a cost efficient minimal defined medium (Wegner, 1983). In 1993 Phillips Petroleum Co. sold the system patent to Research Corporation Technologies who is still the current patent holder. Phillips Petroleum Co. also licensed Invitrogen Corporation to sell components of the system. This combines a well documented protocol for working on *P. pastoris* with a very efficient customer help service.

This eukaryote is capable of many functions carried out in higher eukaryotic cells such as glycosylation, folding and disulphide bond formation. *P. pastoris* is similar to the well-known *S. cerevisiae* system in that the techniques used in molecular genetic manipulation are simple and established. It can also produce foreign heterologous proteins at high levels both intracellularly and extracellularly (Cereghino and Cregg, 1999). One of the main advantages of this system is its availability in a commercial kit, simplifying cloning and transformation.

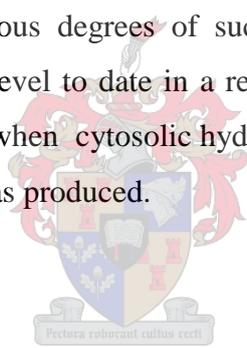
Media for *P. pastoris* cultures are economical and well defined, making it ideal for large scale production of heterologous proteins. As the most rewarding section in terms of cost saving in the development process is the choice of the correct organism and optimal environmental conditions, it is assuring to know that this system is less expensive than insect and mammalian tissue culture cell systems and that in general *P. pastoris* is known for high expression levels. The correct choice of expression

system surpasses the effect of improving reactor design and its operation in later stages (Pattenden *et al.*, 2005).

Additional factors that make *P. pastoris* a popular choice for foreign protein expression is the ability to stably integrate expression plasmids at specific sites in the genome in single or multicopy and the presence of the strongly regulating alcohol oxidase (*AOX*) promoter (Cereghino *et al.*, 2002). This promoter is strongly repressed in cultures grown on glucose and induced over a 1000 fold in methanol medium. This makes it possible to repress foreign protein expression until sufficient biomass has accumulated.

### 2.7.2 Expression levels

Research has been done on the expression of both intracellular and extracellular heterologous proteins with various degrees of success (TABLE 2.5) The highest intracellular protein expression level to date in a recombinant *P. pastoris* strain was 22 g/L (Hasslacher *et al.*, 1997) when cytosolic hydroxynitrile lyase from the tropical rubber tree *Hevea brasiliensis* was produced.



**TABLE 2.5:** Intracellular heterologous proteins expressed in *P. pastoris*

<b>Intracellular complex protein yields</b>			
Type of protein	Type of fermentation	Expression levels	Reference
		[mg/L]	
Envelope protein Dengue type 2 virus chimera with HBsAg (VLP)	Shake flasks	2.6	Bisht <i>et al.</i> , 2001
Anti-p185 <sup>HER-2</sup> scFv	Shake flasks	70	Gurkan <i>et al.</i> , 2004
Tetanus toxin fragment C (14 gene copies)	Shake flasks	466	Clare <i>et al.</i> , 1991
	Fed-batch Fermentation	12 000	
<i>Plasmodium vivax</i> Apical membrane antigen (AMA <sup>-1</sup> )	Chemostat	>50	Kocken <i>et al.</i> , 1999
Botulinum neurotoxin serotype A (BoNT/A(Hc))	Fed-batch fermentation	260	Zhang <i>et al.</i> , 2000
Hepatitis B Surface antigen (HBsAg) (VLP)	Fed-batch fermentation	400	Cregg <i>et al.</i> , 1987
Hepatitis B Surface antigen (HBsAg) (VLP)	Fed-batch fermentation	1000	Chauhan <i>et al.</i> , 1999

It is important to note that to our knowledge Li and Co-workers (2003) were the only group to have expressed HPV6 L1 proteins in *P. pastoris*. Therefore the main benchmark for HPV L1 VLP production in *P. pastoris* is 125 µg from 1-litre cultures. It is however unclear whether this amount of VLPs was per liter of culture or from a concentrated broth of multiple fermentations.

The Hepatitis B surface antigen VLPs produced in *P. pastoris* is another close benchmark for HPV VLP production in *P. pastoris*. Cregg and co-workers (1987) obtained 2.3 mg/L in shake flasks and 0.4 g/L HBsAg in fed-batch cultivation with a Mut<sup>s</sup> strain. Chauhan *et al.*, (1999) reported levels in excess of 1 g/L (0.5 g/L soluble protein) in a Mut<sup>-</sup> strain. However, some genes do not express any measurable amounts of protein. This is often due to yeast transcriptional terminators resulting in truncated mRNA (Romanos, 1995). The AT-rich regions in certain genes are the

cause of this phenomenon and can be solved by gene synthesis as was the case for tetanus toxin gene in *S. cerevisiae* (Romanos *et al.*, 1992)

**TABLE 2.6:** Extracellular heterologous proteins expressed in *P. pastoris*

Secreted extracellular complex protein yields			
Type of protein	Type of fermentation	Expression levels [mg/L]	Reference
Fusion protein	fed-batch fermentation	1500	Jahic <i>et al.</i> , 2003a
Elastase inhibiting peptide	fed-batch fermentation	846	Lee <i>et al.</i> , 2003a
Chimeric alpha amylase (Amy1A)	fed-batch fermentation	340	Lee <i>et al.</i> , 2003b
Angiostatin	fed-batch fermentation	191	Xie <i>et al.</i> , 2005
scFv	Shake flasks	25	Shi <i>et al.</i> , 2003
Merozite surface protein ( MSP1) (Subunit vaccine)	fed-batch fermentation	500	Brady <i>et al.</i> , 2001

Although high secretion levels for extracellular proteins are possible (TABLE 2.6), for use as a vaccine, it is necessary that VLPs are not glycosylated or structurally different from the virus structure as this will influence the immunogenicity of the vaccine (Cereghino *et al.*, 2002). Glycosylation occurs in yeasts and to a lesser degree in filamentous fungi as a post translational modification.

## 2.8 Factors affecting cultivation of *P. pastoris*

The development of a Human Papillomavirus virus-like particle production system in *Pichia pastoris* requires the optimisation of a fermentation strategy. The design of a defined controlled high cell density cultivation strategy consists of two parts:

- (i) The generation of biomass (utilizing available carbon sources)
- (ii) Induction of a foreign protein production through the tight regulation of the *AOX* promoter.

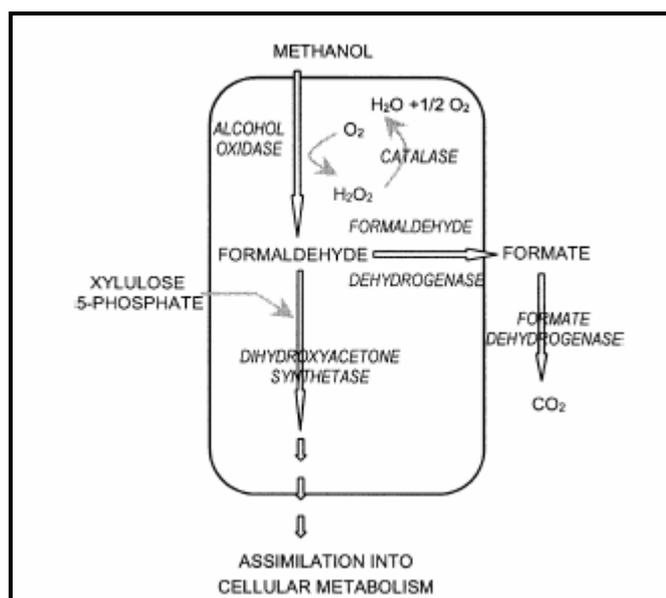
The yield of foreign protein from the system depends on growth conditions, the cultivation process, operating parameters and also the character of foreign protein

itself: its protein stability, cell toxicity and sensitivity to proteases. The methanol concentration and its effect on dissolved oxygen in the fermentation broth have an important effect on the production capacity of recombinant *P. pastoris* (Stratton and Meagher, 1998). *P. pastoris* is a poor fermenter, which is an advantage considering that in fermentation with *S. cerevisiae*, ethanol builds up rapidly and is toxic to the cell, limiting not only growth, but also foreign protein expression. Respiratory growth is favoured by *P. pastoris* and it can be cultured at very high densities.

The following sections discuss factors affecting the expression of heterologous protein. These factors include gene copy number and methanol utilizing phenotype, the role of codon bias, possible feeding strategies, medium composition (specifically non-repressing carbon sources) and methods to prevent protease inhibition of recombinant protein production.

### **2.8.1 Methanol metabolism**

Heterologous protein production is induced by methanol and as such the methylotrophs have very specific enzyme pathways connected to growth on methanol (FIG 2.5). The first step in the methanol utilization pathway is catabolized by alcohol oxidase. In this step methanol is converted to formaldehyde and hydrogen peroxide (Sahm, 1977; Anthony, 1982). Oxidation, facilitated by the *AOX* enzyme, takes place inside a highly specialized organelle called the peroxisome. This organism deftly encapsulates the toxic hydrogen peroxide sequestering it away from the rest of the cell. The formaldehyde can either be utilized to generate biomass or be oxidized by two dehydrogenase reactions to carbon dioxide (Gellissen and Hollenberg, 1997). *AOX* has a poor affinity for oxygen and therefore is present in vast numbers inside the peroxisome. The genome of *P. pastoris* contains two copies of the *AOX* gene. *AOX1* regulates 85% of the oxidase activity in the cell and *AOX2* that regulates the other 15%. *AOX1* is the promoter used to drive foreign protein expression.



**FIGURE 2.5:** Methanol metabolism in *P. pastoris* (Houard *et al.*, 2002)

The advantages and disadvantages of the AOX promoter system are shown in TABLE 2.7.

**TABLE 2.7:** The advantages and disadvantages of the AOX promoter system. (Macauley-Patrick *et al.*, 2005)

<i>Advantages</i>	<i>Disadvantages</i>
Transcription of foreign protein is tightly regulated and controlled by a repression/derepression mechanism.	Monitoring methanol during a process is often difficult, due to the unreliability of on-line probes and the complications of measuring off-line
High levels of foreign proteins can be expressed, even if they are toxic to the cell.	Methanol is a fire hazard, therefore storing of large quantities for fermentations are undesirable.
The repression of transcription by the initial carbon source ensures that good cell growth is obtained before the gene product is over expressed.	Methanol is mainly derived from petrochemical sources, which may be unsuitable for use in the production of certain food products and additives.
Induction of transcription is easily achieved by the addition of methanol.	Two carbon sources are required, with a switching over from one to the other at a precise time point.

In shake flasks AOX makes up 5% of the soluble protein. This percentage is even more increased in fermenter cultures where it is present in excess of 30% when methanol is fed at a limiting rate (Cregg *et al.*, 1987). The expression of AOX is controlled at transcriptional level and is very similar to *S. cerevisiae* mechanism where there exists a repression/derepression and induction mechanism.

### 2.8.2 Alternative promoters

The glutathione-dependent formaldehyde dehydrogenase (*FLDI*) promoter is strongly and independently induced by either methanol as sole carbon source with ammonium as nitrogen source, or methylamine as sole nitrogen source with glucose as carbon source. This facilitates induction of foreign gene expression without the use of methanol. *FLDI* oxidizes formaldehyde to formate and a second enzyme, formate dehydrogenase, breaks it down to CO<sub>2</sub> to supply energy for growth and foreign gene expression in the form of NADH (Sunga and Cregg, 2004).

Shen *et al.* (1998) found that  $\beta$ -lactamase was secreted under the *FLDI* promoter at levels comparable to those under *AOX1*, allowing the use of this promoter as an alternative to methanol induction. Studies by Cos and co-workers (2005) in batch and fed-batch cultures compared *pAOX* and *pFLD* and found similar productivity in both systems. Similarly Resina *et al.* (2004) showed that *Rhizopus oryzae* lipase expression levels were positively influenced by co-induction with methanol and methylamine under both promoters in shake flasks and 1 L bioreactor batch cultures. Making use of different fermentation strategies and using methylamine and sorbitol as nitrogen and carbon sources, respectively, Resina *et al.*, (2005) was able to further increase *Rhizopus oryzae* lipase expression levels significantly more than with an *AOX*-based system producing the same enzyme.

The glyceraldehyde-3-phosphate dehydrogenase (*GAP*) promoter is also a methanol free alternative to *AOX*-based systems. Under the *AOX* promoter, expression of the gene of interest only occurs some time after the initial phase of biomass growth. By using the constitutive *GAP* promoter, expression and growth occurs concurrently in medium containing glucose. Using this *GAP*-based system Vassileva *et al.* (2001) was able to show that HBsAg production compared favourably to the *AOX*-based system. Similarly Aloulou *et al.* (2005) was able to constitutively express human pancreatic lipase-related protein1 in *P. pastoris*, also showing that expression of the protein occurred simultaneously with growth, reaching levels of protein expression of 100-120 mg/L at 40 h.

Both of these promoters are alternatives to the *AOX*-based system, and circumvent the use of methanol as an inducer. This solves problems of methanol toxicity, longer extended fermentation times and storage hazards. Even the use of combinations of the promoters as in the case of enhancing the expression of human granulocyte macrophage colony-stimulating factor in *P. pastoris* (Wu *et al.*, 2003) has shown improved expression while still decreasing the amount of methanol used.

### 2.8.3 Methanol-utilizing phenotypes

Three methanol-utilizing phenotypes (*Mut*) of *P. pastoris* exist and they are the *Mut*<sup>+</sup>, *Mut*<sup>-</sup> and *Mut*<sup>s</sup> strains. The *Mut*<sup>+</sup> strain contains both *AOX1* and *AOX2* and is very sensitive to methanol feeding, grows on methanol at the wild-type rate and requires high feeding rates of methanol in large-scale fermentations. *Mut*<sup>-</sup> is deficient in both the *AOX* promoters and is unable to grow on methanol. This is an advantageous characteristic for certain recombinant products that require low growth rates. In the *Mut*<sup>s</sup> strain only the *AOX2* gene is intact and growth and protein production is then dependent on the weaker *AOX2* for methanol metabolism. This causes slower growth and slower methanol utilisation. The *Mut*<sup>s</sup> strain is used for intracellular expression of heterologous proteins. They have lower levels of *AOX* protein, giving rise to higher specific yields of heterologous protein and the expressed protein can be more readily purified (Sreekrishna *et al.*, 1997). As growth is slow on methanol, longer fermentations which facilitate folding and assembly of proteins is favoured (Cregg *et al.*, 1987). Since as *Mut*<sup>s</sup> strains are not as sensitive to residual methanol in the cultivation medium as *Mut*<sup>+</sup> strains, hence process scale-up is easier (Stratton *et al.*, 1998). Advantages of expressing genes of interest in a *Mut*<sup>-</sup> strain are higher protein yields and a 35-fold reduction in the amount of methanol needed for induction (Chiruvolu *et al.*, 1997). This is important because storage of methanol in explosion proof facilities is expensive and a substantial safety hazard.

Gregg *et al.*, (1987) reported that *Mut*<sup>+</sup> strains produced lower levels of HBsAg compared to strains in which *AOX1* was deleted. Overall, *Mut*<sup>s</sup> and *Mut*<sup>-</sup> have been known to produce VLPs at higher levels than *Mut*<sup>+</sup> (Cregg *et al.*, 1987). Clare *et al.* (1991) argued that it was not so much the *Mut* phenotype that influenced the expression of tetanus toxin fragment C, but the combination of this factor with the

number of integrated gene copies. They showed that in both Mut<sup>s</sup> and Mut<sup>+</sup> single gene copies gave similar production levels. Upon increasing the gene copy number to 9 they found that fragment C levels in Mut<sup>+</sup> strains were 20% lower than for Mut<sup>s</sup>.

#### 2.8.4 Multiple gene integrations

It would seem that with the extremely high expression levels of alcohol oxidase produced from the native *AOX1* gene, a single copy *AOX1*-promoter expression vector would sufficiently induce foreign gene expression. In the case of hepatitis B surface antigen (Cregg *et al.*, 1987a) and  $\beta$ -galactosidase (Cregg *et al.*, 1987b) this proved true, but subsequent studies have shown that an increase in gene copy number could further increase expression levels. TABLE 2.8 shows a summary of increased expression with multi copy recombinant strains.

**TABLE 2.8:** Successful increase in expression with increased gene dosage

Product	Gene copies	Increase in production compared to single copy transformants	References
Murine epidermal growth factor	19	13 fold	Clare <i>et al.</i> , 1991b
Tetanus toxin fragment	14	6 fold	Clare <i>et al.</i> , 1991a
Tumour necrosis factor	>200	200 fold	Sreekrishna <i>et al.</i> , 1989
Aprotinin	5	7.fold	Thill <i>et al.</i> , 1990
Insulin-like growth factor	6	5.fold	Brierly <i>et al.</i> , 1994
HBsAg	8	9.fold	Vassileva <i>et al.</i> , 2001

Two methods of generating multi-copy strains have evolved: i) Identifying multi copy strains that exist naturally at a frequency of a few percent within transformed cell populations through screening methods or, ii) introducing multiple expression cassette

copies into a single vector before transformation. Both methods have given rise to stable strains within the bioreactor environment (Cregg *et al.*, 1993).

It is commonly believed that vector copy number is a determining factor affecting protein yield in intracellular foreign protein production (Vassileva *et al.*, 2001). Studies done on expression of intracellular tetanus toxin fragment C showed that although Mut phenotype and the site of integration had a minor effect on protein levels, increased copy number had a marked effect (Clare *et al.*, 1991). In their study they were able to increase foreign protein production from 3.7 to 8.3% as a percentage of total cell protein to 27% (approximately 12 g/L of fragment C) by increasing the gene copy number to 14. A marked increase (2.5 to 10-fold) was also noticed in scale up from shake flasks to bioreactors. This was due to increased aeration and high cell density possible in the bioreactor. In the production of *Rhizopus oryzae* lipase (ROL) Cos and co-workers (2005) found that an increase in gene copy number resulted in a marked decrease in extracellular expression. They also found that Mut phenotype, in this case Mut<sup>s</sup> gave higher production in the single copy strain than Mut<sup>+</sup>.

As transcript level is a limiting factor in foreign gene expression (Romanos, 1995), maximizing gene copy number is a suitable solution to the problem. Not only does it increase foreign gene expression, but other factors such as protein instability and mRNA secondary structure can be overcome. As to the question of why heterologous genes cannot be expressed as abundantly as *AOX* there are many arguments. It is possible that although mRNA levels of the heterologous protein may be equal or even in excess of *AOX* mRNA present, the translation efficiency of the foreign protein is inhibited. The codon-bias of the heterologous gene may be less favourable than *AOX*. Obviously many of these factors may still have an effect as it depends on the stability of the protein being expressed. This we are only able to determine empirically.

In the case of Hepatitis B surface antigen (HBsAg) an increase in HBsAg gene copy number is accompanied by an increase in steady state mRNA levels. Overall expression of HBsAg increases with copy number by similar amounts (Vassileva *et al.*, 2001). In all transformants the majority of proteins occurred in the particulate form.

Hohenblum and co-workers (2004) conducted a study on stress responses in *Pichia*. During this study on human trypsinogen production, they found that an increase of gene copy number above 2 copies resulted in a decrease in production. This was possibly due to cellular stress response to unfolded proteins. They also conclude that the improvement of expression by increasing the copy number could constitute a major bottleneck for optimisation of recombinant protein secretion. It seems that the effect of gene copy number on production levels is most unpredictable (Sreekrishna *et al.*, 1997). It would be wise to study the production level as a function of gene dosage.

### **2.8.5 The role of codon bias on heterologous protein expression**

All genes consist of nucleotide triplets called codons. The 61 nucleotide triplets encode for 20 amino acids and three codons serve to terminate translation. Different combinations of nucleic acids are able to encode for the same proteins. Certain organisms are biased towards using certain codons. When inserting a certain gene into an organism, it is possible that the specific bias of the organism does not correlate with the gene and that this might have a significant impact on heterologous protein expression (Gustafsson *et al.*, 2004).

A codon adaptation index (CAI) helps us to visualise the relationship/correlation between the codon bias of a gene and its expression level. The index can specify the degree of preference for a specific codon, but it cannot indicate the nature of the preference. Therefore a high CAI towards a certain codon does not automatically mean a gene will be expressed well in an organism. Preferred codons correlate well with a profusion of transfer (tRNA) in the cell content. This in turn optimises the translational capacity of the organism.

When rare codons occur in a gene it is more unlikely that heterologous proteins will be expressed at reasonable levels. To remedy the situation the rare codons in the gene are targeted and modified to more directly resemble the codons used by the host. This is done by sequential site-directed mutagenesis or by resynthesis of the entire gene (Gustafsson *et al.*, 2004). As can be seen from TABLE 2.9 adapted from Gustafsson

*et al.*, (2004) reported a vast improvement in expression by codon optimisation for both HPV in human hosts and other proteins in *P. pastoris*.

**TABLE 2.9:** Improvement of expression levels through codon optimisation (adapted from Gustafsson *et al.*, (2004)).

Gene origin	Protein	Host	Improvement	Reference
HPV	L1	<i>H. sapiens</i>	1x 10 <sup>4</sup> – 1x 10 <sup>5</sup> fold	Leder <i>et al.</i> , 2001
HPV	E5	Mammalian	6 -9 fold	Disbrow <i>et al.</i> , 2003
HPV	E7	Mammalian	20-100 fold	Cid-Arregui <i>et al.</i> , 2003
<i>Corynebacterium diphtheriae</i>	Diphtheria toxin	<i>P. pastoris</i>	0 vs. 10mg L <sup>-1</sup>	Woo <i>et al.</i> , 2002
<i>Actina equina</i>	Equistatin	<i>P. pastoris</i>	2- fold	Outchkourov <i>et al.</i> , 2002
<i>H. sapiens</i>	Glucocerebrosidase	<i>P. pastoris</i>	8- and 10- fold	Sinclair and Choy, 2002
<i>Plasmodium</i>	F2 domain of EBA175	<i>P. pastoris</i>	4- and 9-fold	Yadava and Ockenhouse, 2003

In a study to optimise expression of scFV in *P. pastoris* Hu and coworkers (2006) found that gene copy number had no significant effect on expression. They did, however, find that codon optimisation and lowering of the G+C content of the synthetic gene from 56% to 43% had a significant effect on production and in some cases improved protein yield up to 5 times.

Another example of increasing heterologous gene expression is found in a study by Sinclair and Choy (2002). Their attempt to produce luciferase had limited success despite high levels of transcription. Results suggested that translational inefficiencies were limiting production and that synonymous codon usage bias differences were a translational barrier. They compared a codon optimised construct to another construct with an altered G+C content and found that the latter was the major contributor toward increasing translational efficiency in *P. pastoris*. Therefore, although codon bias has a significant effect on heterologous protein expression, it is not the only factor involved and should be seen as a contributing factor to insufficient production levels.

### 2.8.6 Reactor requirements

Successful foreign protein expression in shake flask cultures of *P. pastoris* have been documented, but are typically low in comparison with expression levels in fermenter cultures (Romanos, 1995). Reasons for this are that transcription levels for *AOX1* is 3 to 5 times greater in cultures that have been fed methanol at growth limiting rates than for cultures cultivated in excess methanol (Romanos, 1995). Even for intracellularly expressed proteins yields of product as a percentage of total cellular proteins are significantly higher in fermenter cultures. Another plausible explanation is that methanol metabolism has a high oxygen demand that can only be met in a fermenter where dissolved oxygen levels can be closely regulated. The *AOX* enzyme pathway has high oxygen requirements not only for cellular electron transport, but alcohol oxygenase requires molecular oxygen as a substrate. Therefore oxygen limitation has a negative effect on foreign gene expression (Lin Cereghino and Cregg, 2000). It is also true that it is only possible to culture *P. pastoris* to the high cell densities it is known for in bioreactors.

The most important fermenter characteristics for high density cultures, such as *Pichia*, are the bioreactors maximum oxygen transfer rate (OTR) and maximum heat transfer rate of the bioreactor. In dense robust cultures, the fermenter must:

1. **Incorporate oxygen** at a high rate from the sparged gas into the dissolved oxygen needed for metabolism. Additionally, OTR depends on agitation-motor power and impeller design.
2. **Dissipate the heat of metabolism and agitation** without allowing culture temperature to rise above the growth optimum. Good temperature control depends on cooling system design and coolant temperature. Of course factors such as substrate concentration and metabolite build-up can also be limiting, but these are often more controllable than inherent physical limitations of the fermenter.

### 2.8.6.1 Feeding strategies

The *P. pastoris* expression system has been used extensively for the production of heterologous proteins. This is due to the tight regulation offered by the alcohol oxidase (AOX) promoter, which is induced over 1000 fold by methanol (Lin Cereghino *et al.*, 2002). It is known that *P. pastoris* can assimilate methanol, but cannot tolerate high methanol concentrations. The oxidized products of methanol, formaldehyde and hydrogen peroxide, accumulate inside the cells and are toxic to the cells.

#### 2.8.6.1.1 Fed-batch feeding strategy

In order to obtain high-density growth and low residual levels of methanol, a fed batch strategy is employed. The strategy usually involves three phases (also illustrated in FIG 8.6.1).

(i) *Glycerol batch phase*

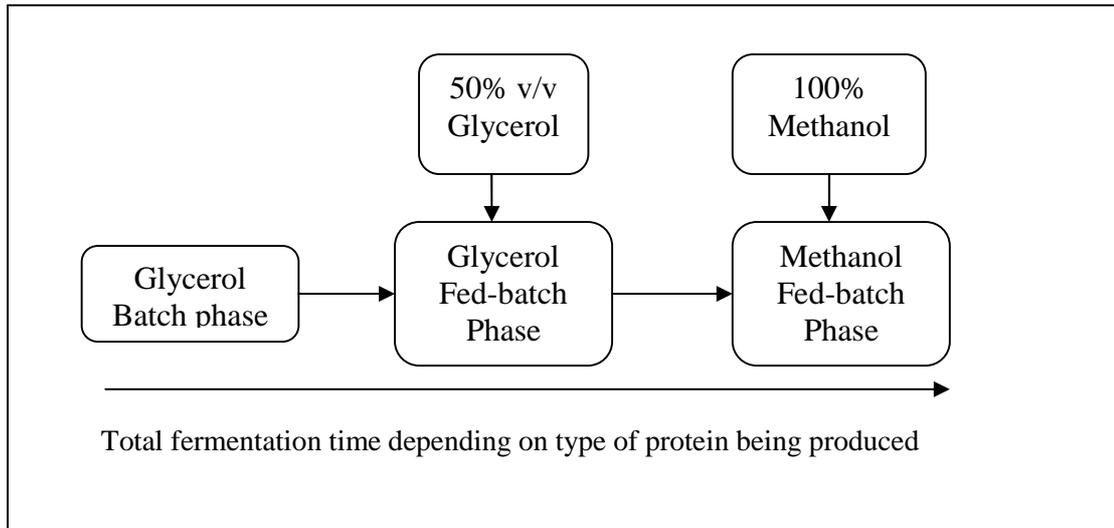
A batch culture of the engineered strain is placed in a simple non fermentable but repressing carbon source, such as glycerol, to accumulate biomass.

(ii) *Glycerol fed batch phase*

After all glycerol is depleted from the batch phase, glycerol is fed to the culture at growth limiting rate to further increase biomass concentration and to prepare cells for methanol induction (derepressing). It also facilitates the consumption of metabolites (ethanol and acetate) that may be inhibitory to AOX1 induction (Zhang *et al.*, 2000).

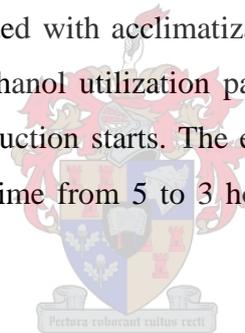
(iii) *Methanol induction phase*

The induction phase is initiated upon completion of the glycerol fed-batch phase by adding methanol to the culture at a slow rate. This facilitates the acclimation of the culture to methanol as carbon source and initiates the synthesis of recombinant protein. From here the methanol feed rate is adjusted upwards until the desired growth rate is achieved (Lin Cereghino *et al.*, 2002).



**FIGURE 2.6: Schematic representation of fermentation strategy**

Some researchers have added a mixed glycerol-methanol feeding phase after the glycerol fed batch phase (Zhang *et al.*, 2000. Brady *et al.*, 2001.). This phase is included to reduce time associated with acclimatization of the cells to the methanol induction (derepressing the methanol utilization pathway), and to ensure a reduced glycerol concentration when induction starts. The employment of this technique has reduced the initial acclimation time from 5 to 3 hours in some cases (Zhang *et al.*, 2000).



### 2.8.6.2 Induction with methanol

The methanol feeding strategy is one of the most important factors in heterologous protein production in *Pichia* as it is directly influenced by biochemical reactions associated with cell growth. The main problem that investigators face is being able to meet the large oxygen transfer rates required as a consequence of the high cell densities achieved. There are two different empirical feeding strategies suggested by Invitrogen in “*Pichia* fermentation process guidelines” that are commonly applied to attempt to overcome this challenge:

#### 2.8.6.2.1 Constant, linear or exponential feed rates.

Feed rates can be constant, linear or exponential, are not related to the metabolism of the organism and do not exceed  $11 \text{ mL.L}^{-1}.\text{h}^{-1}$ . In fact, the amount of methanol present

determines growth of the microorganism. If the methanol concentration supplied is below that necessary to maintain cell growth at  $\mu_{\max}$ , then the actual methanol concentration is zero and its supply is limited (Chauhan *et al.*, 1999; Trinh *et al.*, 2002). The methanol feeding rate is increased and periodically the feed pump is switched off and a dissolved oxygen (DO) spike performed to see whether growth is limited by methanol. A spike is defined as the time it takes for the DO level to increase by 10%. If this time exceeds 1 minute, then the feeding rate is decreased by a set amount (Stratton *et al.*, 1998). While pre-programmed feed rates may give reasonably good results, tighter control of the carbon source and therefore better results, is possible with DO control.

#### **2.8.6.2.2 Dissolved oxygen control**

In this method dissolved oxygen levels are used as an indication of carbon source availability and as such are linked to the methanol feeding scheme. A cascade is selected in the controller to maintain DO at set-point through automatic adjustment of agitation speed and aeration rate. The agitation cascade increases agitation speed with increasing oxygen demand and switches over to pure oxygen sparging when the maximum agitation speed is reached. This method is used to a great extent in *Pichia* fermentations. It is based on the fact that as soon as the carbon source is depleted, the DO concentration starts to increase. Control is maintained in the system, by adhering to the predetermined set-point whilst addition of methanol is taking place. Lee *et al.* (2003) adapted this control by using an oxygen enriching membrane (OEM) instead of the usual cascaded control of agitation and aeration rate. Because zero methanol concentration is not necessarily desirable for maximum protein production, set points of the controllers could be altered to maintain methanol concentrations at desirable levels.

Two approaches can be followed, where in one case the DO is set at a required set-point and methanol is fed at predetermined rates. The system employs agitation and air sparging rates to keep the DO at set-point (Stratton *et al.*, 1998; Ikeda *et al.*, 2004). The risk here is that the culture could outgrow the predetermined feeding rate and then be subject to carbon source limitation to such an extent that heterologous protein

production is negatively influenced if not halted altogether. To avoid this scenario the DO can be controlled by methanol addition in the DO-stat method (Lee *et al.*, 2003). When the DO concentration increases, methanol feeding is switched on and vice versa. The DO-stat method has the advantage of sustaining high cell density cultures in active cellular states for extended time periods and minimizing the accumulation of incompletely reduced substrates. One drawback is that biomass increases constantly and regulation of DO and therefore methanol feeding rates are decreasing over the fermentation run. The decreased feeding rates toward the end of the fermentation do not necessarily induce optimum methanol concentrations for production of heterologous proteins.

Lee and co-workers found that increasing the lower DO set point from 10% to 30% and more frequent methanol additions resulted in significantly enhanced elastase inhibiting peptide (EIP) expression, indicating that higher DO tension stimulates the methanol utilization pathway. The drawbacks of the DO method are:

1. May be complicated to implement.
2. May repeatedly expose cells to non inducing levels of methanol (designed to keep residual levels very low so as not to be toxic to cells)

The first two protocols are designed to keep residual methanol concentrations at nearly zero. This may be optimum for growth rates, but not necessarily for recombinant protein production. A point in case is investigations by Zhang *et al.* (2000) into two methanol fed-batch fermentation strategies. The one culture was grown at methanol-limited feeding rates and the other culture was grown with methanol-excess feeding. From these two strategies an unrestricted growth model was developed that accurately described the system. It should be noted that this growth model is specific to the recombinant protein being produced and therefore not generally applicable. Nonetheless the experiment indicated that a significant difference in production levels was experienced between these two extremes. They successfully employed through the growth model a feeding scheme that coupled growth rate to production levels.

Another method employs off-line gas chromatography to monitor methanol concentration in the culture medium. Results show that in this specific system, a 2.5 fold increase in recombinant protein production was obtained in comparison with the DO method (Minning *et al.*, 2001). Although in this instance tighter control of the carbon source was possible, chromatography or HPLC analysis is relatively expensive and difficult to implement online.

#### **2.8.6.2.3 Constant methanol concentration by methanol sensor and control loop.**

In this control strategy the microorganism is controlling the amount of methanol added as a response to its own metabolism. Methanol supply is practically unlimited and methanol concentration can have any value below 4 g/L where toxicity becomes a problem (Trinh *et al.*, 2002).

Certain advantages of online monitoring of methanol concentration make it a better option than DO control. With constant methanol control it is possible to:

1. Obtain more reproducible results.
2. Employ more sophisticated fermentation in order to methodically approach research in this area.
3. Help avoid safety hazards during large scale fermentations.
4. Contribute to good manufacturing practices (GMP) by the availability of on-line data on the concentration of toxic substrate methanol during the production of therapeutically relevant proteins.

Guarna and co-workers (1997) were some of the first researchers to employ a methanol sensor and control system to specifically measure methanol concentrations on-line and be able to adjust them to a certain set point value. The sensor system was an organic solvent vapour detector TGS 822 (SnO<sub>2</sub>) obtained from Figaro (Minoh, Osaka). Methanol concentration values obtained with this sensor correlated well with those obtained from an off-line HPLC analysis.

The voltage output of the sensor is connected to a personal computer where it is compared to a manually entered set-point in order to generate an actuating signal to drive the methanol feed pump. A low methanol consumption rate compared to the lowest practical feeding rate indicated the use of a simple “on/off “pulse feed control. After a pulse, a wait state was entered for 5-7 min (longer than normal response time of probe) before the concentration was reevaluated. Non-linearity evident in the sensor response function results from a combination of two effects: an artifact of the electronic measuring circuit that employs a simple voltage divider configuration to measure conductance changes in the detector element, yielding a hyperbolic relationship between output voltage and conductance and a composite effect results from physical chemistry and the physical electronic principles of semi-conducting metal-oxide gas sensors.

Although these experiments by Guarna and co-workers (1997) were conducted in shake-flasks to produce the N-lobe of human transferrin, a five fold increase in volumetric productivity over the conventional fed-batch method was obtained. This shows that a residual level of methanol might actually enhance heterologous protein production.

Hellwig *et al.* (2000) employed the Frings Alkosense Metha probe and a control system (Acetomat II signal processor) to express a single chain antibody fragment (scFv4813). It operates similarly to the Raven sensor, and has a silicone membrane. Studies by them showed that methanol control of this manner increased production. They also concluded that in *P. pastoris* fermentation, when DO controlled or in a limited methanol feed control strategy, carbon source availability may be the limiting factor for expression.

Cunha *et al.* (2004) did experiments to optimize the methanol feeding strategy and cell concentration at the induction point integrated with the dissolved oxygen control, thus allowing the use of the total available power input in the fermenter to achieve highest titre and product quality. They also used the Figaro organic vapour metal oxide sensor to maintain constant methanol concentration. They found that the methanol uptake rate had a significant effect on the specific production rate and that uptake rates below 0.026 g/gcells/h gave rise to a linear increase in specific

production rate. The optimisation resulted in a ten fold increase in the dimeric form of scFv.

Katakura *et al.* (1998) expressed Human  $\beta 2$  Glycoprotein I Domain V in *P. pastoris* and used mixed methanol and glycerol feeding to determine the effect on specific production and compared this with a base case methanol feed. This was the first study on constant methanol concentration on bench-top fermentation scale. A methanol control system with semiconductor gas sensor and relay was used, similar to Guarna *et al.* (1997). Once again the methanol measurement was taken from the fermenter headspace or exhaust gas.

When characterising the probe performance, it was found that the fermenter temperature drift affected the methanol measurements as well as the airflow rate to the fermenter. These problems could be rectified by ensuring tight control of these variables at constant values. Sensor values correlated well with off-line measurements.

Studies on the effect of methanol on specific rates revealed a decrease in specific growth rate as well as consumption rate of methanol as methanol concentration increased. Interestingly enough, specific production rates increased. This shows that at high methanol concentration cell growth is markedly inhibited, consumption is inhibited slightly, but production is increased. Inhibition must then be due to toxic effects of methanol on DNA replication and membrane synthesis, because metabolism and protein biosynthesis were not significantly influenced.

Katakura *et al.* (1998) hypothesized that the increase in productivity could be due to an increase in protein gene transcription, an increase in permeability across cell membrane for protein or an increase in energy supply for protein production.

The following are suggestions for the elimination of these reasons

1. AOX activity decreases with methanol concentrations higher than 20 g/L and that specific methanol consumption decreased, thus an increase in gene copy number could not be responsible for an increase in production levels.
2. It is known that solvents permeabilise cell membranes. Thus if the larger protein was present in the broth, smaller molecules (metabolites) would also

have been present in the broth and would cause serious damage to cell. As this did not occur, an increase in permeability could not be responsible for an increase in production.

3. Energy for cell growth and protein production obtained from methanol oxidation and growth and production compete for energy in the production phase. When cell growth is inhibited by high methanol concentrations, production is benefited. This is supported by the fact that specific production rates were higher in the transition phase than in the production phase. Glycerol was also fed in this phase and therefore more energy was available for production.

The main obstacle to high production would then seem to be a lack of a sufficient energy from the carbon source. When the methanol concentration is increased, inhibition occurs. To remedy this situation mixed feeding with glycerol was investigated as a simple way of increasing energy supply to cells by increasing the carbon source concentration. Specific rates were constant during the early production phase and the specific methanol consumption rate remained the same compared to the straight methanol feeding. The specific growth rate increased by approximately 20% with mixed feeding and specific production levels increased 2.3 fold. This indicated that foreign protein expression was not repressed by glycerol (Katakura *et al.*, 1998).

The expression of recombinant laccase was investigated by Hong *et al.* (2002). They used a methanol sensor similar to the one used by Guarna *et al.* (1997), the main difference being the point of obtaining the methanol sample. The methanol sensor and control system was obtained from Raven Biotech and consisted of a probe with gas-permeable silicone rubber tubing that was inserted into the fermentation broth. The control system was also provided by Raven Biotech.

The fermentation strategy employed by Hong *et al.* (2002) included the batch and fed-batch glycerol phases and aimed to find the best temperature and constant methanol concentration for the induction phase. They concluded that a 0.5% constant methanol concentration gave the best laccase activity, but not necessarily the best growth rate. An induction temperature of 30°C was more favourable towards induction than 20°C. Methanol inhibition was only noticed at methanol concentrations higher than 2%.

Any of these cultivation methods could be employed as a starting point for process optimisation, keeping in mind the availability of analytic equipment as well as the limited applicability of some of the methods to certain types of proteins being expressed.

### 2.8.7 Nutrients

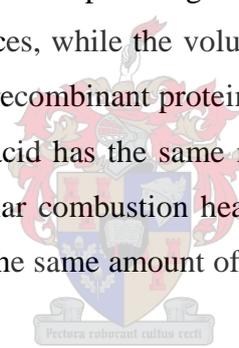
Much time, effort and capital is spent on the setting up of a fermentor run, therefore it is important to obtain the maximum growth and heterologous expression from the carbon source used. The methanol feeding strategy in the induction phase is dependent on the Mut strain and it influences the amount of growth and expression of foreign proteins. The *AOX* promoter induces foreign gene expression, and almost all the energy from the carbon source is directed toward foreign gene expression and very little towards cell maintenance and growth. Most carbon sources in mixed feed with methanol repress the *AOX* promoter, effectively decreasing foreign protein expression. It is therefore necessary to identify a non-repressing carbon source that can support cell growth and maintenance during the induction phase.

Besides the fact that glycerol is repressing to the *AOX* promoter, growth on glycerol causes the formation of metabolic substances such as ethanol and acetate, which are also repressing. Iman and Meagher (2001) found that ethanol and acetate was preferentially used, repressing the methanol utilization and expression of foreign protein during the transition phase. This is due to a high glycerol feed rate and careful control of glycerol feed rate is necessary when using the Mut<sup>-</sup> strain.

Inan *et al.* (2001) investigated the use of different non-repressive carbon sources on a Mut<sup>-</sup> strain. They compared different sources in terms of their ability to support growth and expression of  $\beta$ -galactosidase. They found that glucose, glycerol, ethanol and acetate repressed production of  $\beta$ -galactosidase, while growth and expression were supported on alanine, mannitol, sorbitol or trehalose. Alanine delivered the best results when fed in a mixed feed with methanol.

Sreekrishna *et al.*(1997) reported that sorbitol and alanine increased the recovery of expressed protein when used in a mixed feed batch fermentation to accomplish several cycles of Matrix metalloproteinases (MMP-2) production. Sorbitol was used as the main carbon source and methanol was added to induce expression. In another experiment by Chauhan *et al.*, (1999) found that the addition of 0.1% alanine or casamino acids increased the yield of HbsAg particles. Sorbitol at this concentration had an inhibitory effect, as had addition of the other two components at higher concentrations.

The use of lactic acid by Xie *et al.*, (2005) as a main carbon source for the production of angiostatin increased the yield 1.7- 2.5 fold. Accumulation of lactic acid in the fermentation broth did not repress expression. In comparison with the 191 mg/L angiostatin produced on the lactic acid strategy, 141 mg/L angiostatin was produced on sorbitol. Interestingly enough the specific growth rate on lactic acid was much lower than on other carbon sources, while the volumetric angiostatin production was still high. This means that more recombinant protein was produced on fewer cells in a shorter period. Although lactic acid has the same molar carbon number as glycerol, the degree of reduction and molar combustion heat is lower. This means that more lactic acid is consumed to form the same amount of biomass and a lower cell yield on lactic acid was obtained.



### **2.8.8 Proteolysis and the temperature effect**

Proteolysis occurs in most cells after a certain amount of time has passed in the fermentation. It occurs due to the release of extracellular proteases and mainly affects extracellular protein production and is closely associated with pH and temperature. Many studies have been done to establish a link between these factors. The effect of pH in the production of an intracellular protein is not as great as with an extracellular product (subject to degradation by proteases). Certain pH ranges may however be more optimum for growth of the host cells and, indirectly, recombinant protein production. *Pichia* fermentation process guidelines from Invitrogen suggest the best pH for cell growth is in the range of 5-6.

Another factor that plays an integral part in cell degradation and death is temperature. While the optimal growth temperature for *P. pastoris* is between 28°C and 30°C, it has been proved that lower temperatures might be conducive to higher recombinant protein production. During the expression of herring antifreeze protein in *P. pastoris* (Li *et al.*, 2001) a reduced temperature of 23°C improved productivity without decreasing cell growth. A much debated point is the relationship between temperature and proteolysis. Curvers *et al.* (2001) suggested reduction in proteolytic activity due to decrease in the specific activity of the proteases, while Jahic *et al.* (2003b) proposed that three factors may contribute to a decrease in proteolysis *i.e.* i) reduced proteolysis due to lower temperature, ii) reduced proteolysis due to lower cell death and protease release to the medium, iii) increased synthesis rate due to higher AOX activity at lower temperatures. The third reason could indicate that there exists an optimum temperature that increases AOX activity and therefore expression of recombinant protein.

It was suggested by Hong *et al.* (2002) that the temperature be reduced during the induction phase from 30°C to 20°C. Although the expression of laccase in their experiments was extracellular, they found that a decrease in temperature reduced the folding problems of the protein by decreasing the growth rate and allowing more time for protein folding. In the case of self-assembling virus-like particles allowing time for folding may not be of great use as proteins self-assemble and have till the end of the fermentation to do so.

It can be seen from the literature reviewed that most recombinant proteins expressed in *Pichia* is secreted. For intracellular proteins the pH may have little effect and a decrease in temperature may at most increase the viability of cells.

### 3. Experimental procedures

#### 3.1 Materials and Methods

##### 3.1.1 Strains and plasmids

The wild type strain *Pichia pastoris* X-33 (Mut<sup>-</sup>), the lab strain KM71H (*AOX1::ARG4 HIS4* Mut<sup>s</sup>) and the plasmid pPICZ-A were purchased from Invitrogen as part of the EasySelect™ *Pichia* expression kit. GS115/pPICZ/lacZ strain was used as a positive control for induction and was obtained from Invitrogen (Carlsbad, CA, USA). The *Escherichia coli* strain DH5α was used for plasmid preparations. Strain constructs are shown in TABLE 3.1. All strains were prepared by Dr. Riaan den Haan (Department of Microbiology, University of Stellenbosch).

**TABLE 3.1:** Strain constructs with the pPICZ-A plasmid

Strain	Gene	Abbreviation	Reference
<i>P. pastoris</i> X-33	h-L1	X33[h-L1]	This study
<i>P. pastoris</i> X-33	h-L1	X33[multi-h-L1]*	This study
<i>P. pastoris</i> X-33	Syn-L1	X33[Syn-L1]	This study
<i>P. pastoris</i> X-33	SA-L1	X33[SA-L1]	This study
<i>P. pastoris</i> GS115	lacZ	GS115[lacZ]	Invitrogen

\* This strain contains 4 copies of the h-L1 gene, while others contain a single copy

##### 3.1.2 Media and culture conditions

For *E. coli* transformation, a low salt LB (10 g.L<sup>-1</sup> tryptone, 5 g.L<sup>-1</sup> NaCl, 5 g.L<sup>-1</sup> yeast extract, pH 7.5) containing 25 µg.mL<sup>-1</sup> zeocin (Invitrogen) was used. Cells were routinely cultured at 37°C on a rotating wheel or on an orbital shaker at 100 rpm. Yeast transformants were plated on YEPD (10 g.L<sup>-1</sup> yeast extract, 20 g.L<sup>-1</sup> peptone, 20 g.L<sup>-1</sup> dextrose, 20 g.L<sup>-1</sup> agar) containing 100 µg.mL<sup>-1</sup> zeocin with the pH adjusted to 7.5 with NaOH and grown for 2-10 days at 30°C. To determine the Mut phenotype of transformants, cells were cultivated overnight in 5 mL liquid YEPD at 30°C. One microlitre of the cell suspensions were plated on minimal dextrose (13.4 g.L<sup>-1</sup> yeast nitrogen base, 20 g.L<sup>-1</sup> dextrose, 4 x 10<sup>-4</sup> g.L<sup>-1</sup> biotin, 20 g.L<sup>-1</sup> agar) and minimal

methanol (13.4 g.L<sup>-1</sup> yeast nitrogen base, 5 g.L<sup>-1</sup> methanol, 4 x 10<sup>-4</sup> g.L<sup>-1</sup> biotin, 20 g.L<sup>-1</sup> agar) media and grown for two days at 30°C.

### 3.1.3 Plasmid construction

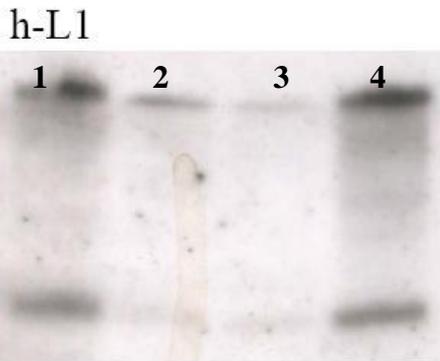
Standard procedures for DNA manipulation were followed as outlined by Sambrook *et al.* (1989). Three plasmids, pGEM SA-L1, pGEM syn-L1 and pCRscript h-L1 were used for the cloning experiments. All three L1 genes are derived from HPV type 16 encoding the 506 aa L1- protein. The gene variants syn-L1 and h-L1 were codon-optimized for expression in potato plants (*Solanum tuberosum*) and human cell lines respectively. All L1 genes were kindly provided by Prof. Ed Rybicki (Molecular and Biosciences Department, UCT) and all three L1 gene variants were cloned into a pPICZ-A plasmid, an *E. coli-P. pastoris* shuttle vector under the transcriptional control of the *PAOX1* (FIG 3.1). Both the h-L1 and syn-L1 genes were inserted into the *EcoRI*-site of pPICZ-A. SA-L1 was inserted into the *NotI*-site of pPICZ-A. Further restriction digestions with other restriction enzymes confirmed the correct orientation of these genes in the plasmids.

### 3.1.4 Transformation

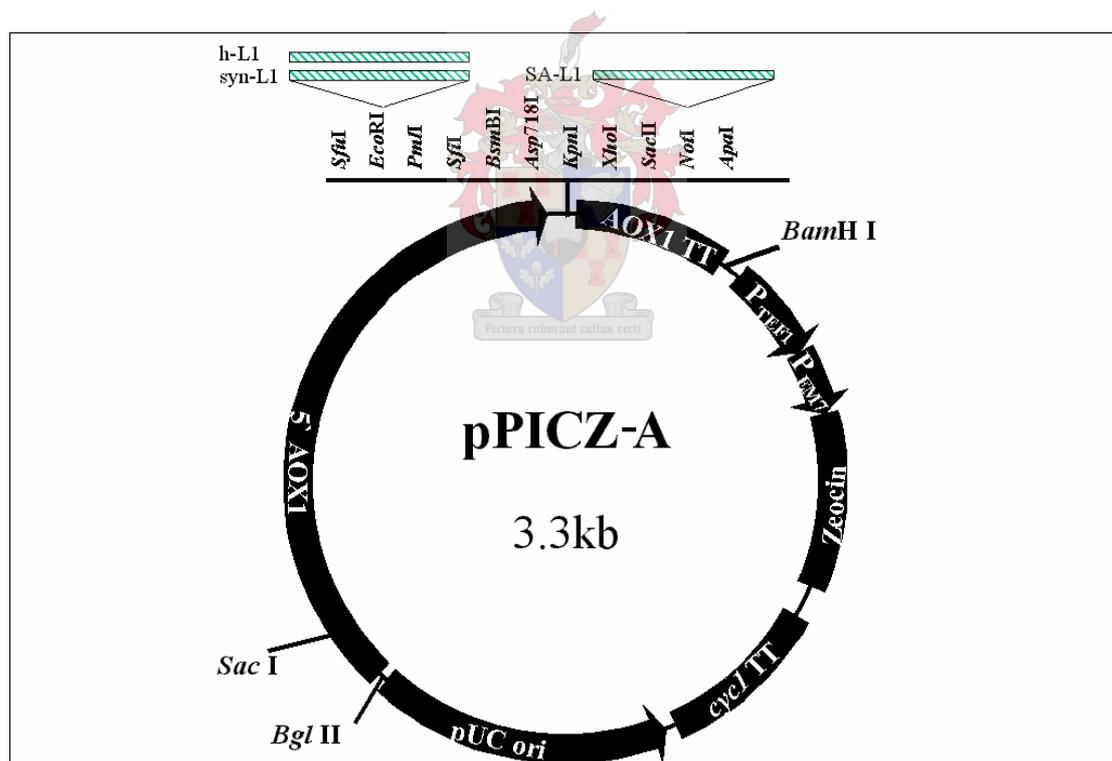
The *Pichia pastoris* host strain used in all subsequent experiments was the X-33 wildtype strain. Three recombinant strains of *P. pastoris* X-33 recombinants were constructed containing the three HPV L1 gene variants (FIG 3.1). Furthermore a multi-copy recombinant with 4 copies of the h-L1 gene was also used. The expression of the h-L1, Syn-L1 and SA-L1 genes are under the control of both *AOX* promoters (*Mut*<sup>+</sup>) which is induced with methanol.

The pPICZ-A plasmids containing the respective L1 gene inserts were linearized with *SacI* (Roche) and transformed into X-33 and KM71H using the standard yeast electroporation protocol provided by the *Pichia* expression kit manual (Invitrogen). For the selection of multiple copy integrants the zeocin concentration used was 500, 1000 and 2000 µg.mL<sup>-1</sup>. Only *P. pastoris* X-33 was transformed with pPICZ-A h-L1. Identification of the presence of the integrated L1 gene was confirmed with polymerase chain reactions (PCR) using primers complementary to internal sequences in all the gene variants. The Southern blot in FIG confirmed four integrated copies of the h-L1 gene. Only one of the screened transformants had multiple gene integration

due to difficulty in integration and was used in this study. For PCR, *TaqI* DNA polymerase (New England Biolabs) was used as recommended by the manufacturers for the Perkin Elmer GeneAmp® PCR system 2400.



**FIGURE 3.1: Southern blot showing four integrated copies of the h-L1 gene (Lane 4).**



**FIGURE 3.2: Plasmid construction of pPICZ-A plasmids containing the three alleles of the L1 genes.**

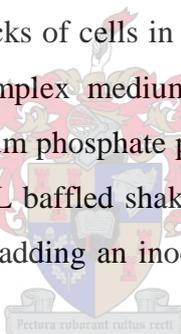
### 3.1.5 Shake flask cultivation

All shake flask experiments were carried out in 250mL baffled shake flasks with a 50 mL working volume. Cultures were inoculated from single colonies on YEPD

plates and incubated in Buffered Minimal Glycerol Medium (BMGH: 100 mM potassium phosphate pH 6.0, 13.4 g.L<sup>-1</sup> YNB, 4 x10<sup>-4</sup> g.L<sup>-1</sup> biotin, 10 g.L<sup>-1</sup> glycerol) at 30°C on a rotary shaker at a constant speed of 100 rpm. When culture optical density of OD<sub>600</sub> 2-6 was reached, cells were centrifuged at 1500 x g at 4°C for 5 minutes and resuspended in 10 g.L<sup>-1</sup> minimal methanol medium (BMMH; 100 mM potassium phosphate pH 6.0, 13.4 g.L<sup>-1</sup> YNB, 4 x10<sup>-4</sup> g.L<sup>-1</sup> biotin, 10 g.L<sup>-1</sup> methanol). Methanol (final concentration 0.5%, 1%, 2%, 3%) was added every 24 h during the induction phase. Cultures were sampled every 24 h and 2 mL per shake flask was removed for analysis.

### 3.1.6 Bioreactor cultivations

Cultivations were carried out in a 1.3 L New Brunswick Bioreactor with a working volume of 1 L at 30°C and a maximum agitation of 900 rpm with a Rushton turbine impeller. The initial culture volume was 500 mL. Pre-cultures for fermentations were inoculated from 1 mL glycerol stocks of cells in the early stationary phase of growth into 50 mL buffered glycerol complex medium (BMGY; 10 g.L<sup>-1</sup> yeast extract, 20 g.L<sup>-1</sup> peptone, 100 mM potassium phosphate pH 6.0, 13.4 g.L<sup>-1</sup> YNB, 4 x10<sup>-4</sup> g.L<sup>-1</sup> biotin, 10 g.L<sup>-1</sup> glycerol) in 250 mL baffled shake flasks. Cultures were grown to an OD<sub>600</sub> 3 (for approx. 11 h) before adding an inoculum of 40 mL to reach a working volume of 500 mL.



The filtered sterile air (0.2 µm absolute filters, resterilizable) was passed through a rotameter to control the aeration rate at 1 vvm (volume of gas inlet flow rate per volume of culture per minute). In cultivations where the culture dissolved oxygen (DO) concentration required control independent from the agitation and aeration rate, NBS gas mixers were used. DO concentrations were analysed by an online NBS polarographic electrode. An online pH electrode (Mettler-Toledo, USA) measured the pH of the culture broth. Ammonium hydroxide (28%) (NH<sub>4</sub>OH, R&S Enterprise) solution was used to maintain a pH of 5 in the bioreactor. Ammonium hydroxide was the main nitrogen source and did not cause a build-up of salts. In the exhaust gas of bioreactor cultures, the concentration of O<sub>2</sub> and CO<sub>2</sub> were determined by infrared gas analysis systems (INNOVA AirTech Instruments 1301 Gas Analyser). Foam formation was suppressed by the addition of antifoam reagent (Antifoam Y-30 Emulsion, Sigma), with the aid of a level sensor. A foam trap was installed as an extra precaution.

### 3.1.6.1 Glycerol batch phase

A defined basal salts medium was used as fermentation medium and consisted of the following: 26.7 mL.L<sup>-1</sup> H<sub>3</sub>PO<sub>4</sub>; 0.93 g.L<sup>-1</sup> CaCl<sub>2</sub>; 18.2 g.L<sup>-1</sup> K<sub>2</sub>SO<sub>4</sub>; 14.9 g.L<sup>-1</sup> MgSO<sub>4</sub>.7H<sub>2</sub>O; 4.13g.L<sup>-1</sup> KOH; 40 g.L<sup>-1</sup> Glycerol. A trace salts solution (PTM1) containing: 6 g.L<sup>-1</sup> CuSO<sub>4</sub>.5H<sub>2</sub>O; 0.08 g.L<sup>-1</sup> KI; 3 g.L<sup>-1</sup> MnSO<sub>4</sub>.H<sub>2</sub>O; 0.2 g.L<sup>-1</sup> Na<sub>2</sub>MoO<sub>4</sub>.2H<sub>2</sub>O; 0.02 g.L<sup>-1</sup> H<sub>3</sub>BO<sub>3</sub>; 0.5 g.L<sup>-1</sup> CoCl<sub>2</sub>; 20 g.L<sup>-1</sup> ZnCl<sub>2</sub>; 65 g.L<sup>-1</sup> Fe<sub>2</sub>SO<sub>4</sub>.7H<sub>2</sub>O; 0.2 g.L<sup>-1</sup> Biotin; 5 mL.L<sup>-1</sup> H<sub>2</sub>SO<sub>4</sub> was filter sterilized, stored at 4°C and added to the basal medium at 4.4 mL.L<sup>-1</sup> medium once the fermentor had cooled down to 30°C after being autoclaved. Automatic DO control was set at 40% and control was cascaded to agitation and by the amount of oxygen sparged through the gas mixer once the maximum agitation rate of 900 rpm has been reached. A sharp rise in dissolved oxygen concentration occurred 20-24 h after inoculation and indicated the depletion of the glycerol carbon source.

### 3.1.6.2 Glycerol fed-batch phase

During the glycerol fed-batch phase a 50% (w/v) glycerol solution with 12 mL.L<sup>-1</sup> trace salts solution was fed to the fermenter at 18.3 mL.L<sup>-1</sup>.h<sup>-1</sup> initial volume for 5 h or until wet cell weight of approximately 200 g.L<sup>-1</sup> was reached. At this stage automatic DO control was still set at 40% and controlled by the amount of oxygen sparged through the gas mixer with a constant maximum agitation rate of 900 rpm. DO spikes were carried out during this phase to ensure that no glycerol accumulation occurred. A DO spike consisted of shutting off the glycerol feed and timing how long it took for the DO to increase by 10%. A satisfactory DO spike was observed under 1 minute and indicated growth limitation based on the available carbon source.

### 3.1.6.3 Methanol fed-batch phase

#### 3.1.6.1.1 Methanol feeding for adaptation phase (Stratton *et al.*, 1998).

The glycerol feed was stopped when sufficient biomass levels were reached and the DO rose to approximately 100%. For the methanol fed-batch stage, a solution of 100% methanol and 12 mL.L<sup>-1</sup> trace salts solution was fed to the bioreactor at predetermined feeding rates. The first stage was initiated to adapt the culture to growth on methanol and flow rates were typically 3.6 mL.L<sup>-1</sup>.h<sup>-1</sup> initial volume. During this adaptation period (2-5 h) DO steadily decreased to approximately 40%. Eventually the DO stabilized (approx. 2 h) and increased again, this meant that the

culture was starved for methanol. In the second stage the methanol feed rate was increased to  $7.2 \text{ mL.L}^{-1}.\text{h}^{-1}$  initial volume. A DO spike was performed after 2 h and the feed rate was maintained until the DO spike time was less than 30 s. When the spike time was less than 30 s, the methanol flow rate was increased by  $1 \text{ mL.L}^{-1}.\text{h}^{-1}$  initial volume at 1h intervals until the optimal feed rate of  $11\text{-}12 \text{ mL.L}^{-1}.\text{h}^{-1}$  initial volume was reached. If at any time the DO spike time exceeded 1 min, the methanol flow rate was decreased by  $1\text{-}2 \text{ mL.L}^{-1}.\text{h}^{-1}$  initial volume. The optimum feeding rate was maintained until the end of the fermentation. If at any time the DO dropped below 15% the methanol feed pump was stopped, allowing the culture to metabolize the methanol. Only when the DO was rising again was any adjustments made to the oxygen flow rate. Adding oxygen to immediately increase DO was detrimental to the culture as the excess methanol and oxygen combined to produce formaldehyde, which pickled the culture.

#### **3.1.6.1.2 Methanol feeding with DO control (Lee *et al.*, 2003a,b).**

In this feeding strategy, DO is used as an indication of carbon source availability. As the proposed method of having a set point bandwidth was not feasible with the current command system, the DO set point was set at 40%. Through BioCommand a basic control program was initiated. It consisted of defining a set point condition and then comparing true or false operators to this condition. The output then either indicated that methanol feeding be halted or recommenced. Methanol feeding rate was set at  $12 \text{ mL.L}^{-1}.\text{h}^{-1}$  initial volume. The methanol adaptation phase in which methanol feeding was ramped up from  $1.8$  to  $6 \text{ mL.L}^{-1}.\text{h}^{-1}$  initial volume was similar in both control strategies.

#### **3.1.6.1.3 Dissolved oxygen measurement**

Culture DO was measured using a DO probe (Metler Toledo InPro 6000, Inc., USA). Calibration was conducted according to New Brunswick Scientific BioFlo 110 protocol using a standard two point referencing technique, with 0% saturation corresponding to a nitrogen equilibration and 100% corresponding to air saturation at  $30^{\circ}\text{C}$ .

### **3.1.7 Analytical methods**

#### **3.1.7.1 Measurements of optical density and cell dry weight**

Optical densities were determined with a spectrophotometer at 600 nm. Dilutions were made to keep the OD reading below 1 unit. Wet cell weight ( $g_{WCW} \cdot L^{-1}$ ) was determined by washing a 2 mL volume of cell culture with breaking buffer, centrifuging for 2 min at 8000 g and then discarding the supernatant. The wet cell pellet was then weighed. Dry cell weight ( $g_{DCW} \cdot L^{-1}$ ) was determined by drying the wet cell pellet at 75°C for 96 h.

#### **3.1.7.2 Preparation of cell lysates**

Cell lysates were prepared according to the Invitrogen Corporation (Version G) *Pichia* expression kit with the following modifications: A washing step in breaking buffer containing per liter: 6 g Sodium phosphate (monobasic), 372 mg EDTA, 50 mL glycerol. Centrifugation at 3000 g and 4°C was followed by a resuspension step in breaking buffer followed by centrifuging at 8000 g for 1 min. 0.15 g of spun down cells were then made up to 1.5 mL with breaking buffer and 0.5 mL aliquotted to three impact-resistant glycerol tubes filled with an equal volume of acid washed glass beads (0.45 mm, Sigma). The tubes were then subjected to 9 cycles in a FastPrep® FP120A Instrument (120V) for 20 sec per cycle at  $5.5 \text{ m} \cdot \text{s}^{-1}$ . Each lysing cycle was separated by a 2 min rest cycle on ice. After centrifuging the samples at 12000 g and 4°C for 15 min the supernatant was extracted (first extraction). Another 200 µl of breaking buffer is then added to the samples and subjected to 1 cycle of breaking for 20 sec at  $5.5 \text{ m} \cdot \text{s}^{-1}$ . The sample was then spun down again at 12000 g and 4°C for 10 min. The supernatant was extracted and this second extract was then added to the first extract to obtain the soluble protein or cytoplasmic fraction. 200 µl of breaking buffer containing 1% Triton X-100 was then added and subjected to 1 cycle of breaking for 20 sec at 5.5 K. The sample was then spun down again at 12000 g and 4°C for 10 min. This supernatant was extracted and termed the insoluble or membrane fraction.

#### **3.1.7.3 Determination of HPV monomer and VLP levels**

HPV L1 protein was quantified from *P. pastoris* cell extracts by capture ELISA, which was modified from a polyvinyl alcohol (PVA)-blocking ELISA method (Studentsov *et al.*, 2002). A 96-well microtitre plate was coated with monoclonal

antibodies HPV16 J4 (recognises a linear HPV16 L1 epitope) or HPV16 V5 (recognises conformational epitope) overnight at 4°C, washed and blocked. Cell extracts were then added and incubated for 1h at 37°C, followed by a washing step and addition and incubation of rabbit anti-HPV16 VLP polyclonal serum (1:1000) for 1 h at 37°C. This serum was detected with swine anti-rabbit-HRP conjugate (1:5000; DAKO, Denmark) and 1,2-phenylenediamine dihydrochloride (OPD; DAKO) substrate. The colorimetric reaction was stopped with 2.8% sulphuric acid solution. Optical density of the samples was then detected at 450 nm in a Titertrek Multiskan MKII plate reader. Baculovirus-derived VLPs of known concentrations were used as standards.

#### **3.1.7.4 Total protein determination**

Total protein concentration of extract samples were determined using the Bio-Rad protein assay, based on the method of Bradford (Bradford, 1976), (Bio-Rad Laboratories, Hercules, CA), using bovine serum albumin (Pierce Chemical Co., Rockford, IL) as the reference standard.

#### **3.1.7.5 Analysis of m-RNA**

##### **3.1.7.5.1 Total RNA isolation**

Total RNA was isolated from frozen cell samples according to the method described in Invitrogen (Version G) *Pichia* expression kit manual. Total RNA concentrations were determined using the NanoDrop system (NanoDrop Spectrophotometer ND1000 3.2.1, Coleman Technologies, 1998).

##### **3.1.7.5.2 Slot Blot analysis**

Equal amounts of RNA were transferred to a nylon membrane (MSI, Westboro, MA) with slot blotting (Ausubel *et al.*, 1995). Hybridisation was done with PCR generated DIG-labeled fragments of the h-L1 gene and a *S. cerevisiae* actin gene (*ACT1*) with a high homology to *P. pastoris* actin gene. The latter was used as an internal control to comparatively quantify mRNA levels of different fermentations. Transcripts were visualised with the Chemiluminescent Detection Kit (Roche Biochemicals, Germany).

### 3.1.8 Plate $\beta$ -Galactosidase activity assays

$\beta$ -Galactosidase activity was determined on plates made of minimal methanol media containing 70  $\mu\text{g}/\text{mL}$  X-Gal. Methanol concentrations of 0.5%, 1% and 3% were investigated. An *E.coli* strain containing the plasmid pBluescript was used as a positive control for  $\beta$ -Galactosidase activity and a X-33 pPIC-X33[h-L1] strain was included as a negative control. Blue zones on the plates indicated  $\beta$ -galactosidase activity.

## 3.2 Experimental philosophy

From the literature review it was evident that cultivation methods for *P. pastoris* were very well developed and documented. It was also obvious that although a standard type of fermentation could be performed, various parameters could be altered to suit the specific heterologous protein being produced. Factors that influenced parameter changes have already been mentioned, but they can be roughly divided into two types of influences namely strain dependent parameters and cultivation dependent parameters. As HPV16 VLPs have never been produced in *P. pastoris* it was necessary to vary some of these basic parameters to determine optimal yield of the product.

### 3.2.1 Type of cultivation and mode of operation

Submerged fermentation is the fermentation type of choice for the production of heterologous gene products in *P. pastoris*. Additionally, the infrastructure and knowledge base for submerged fermentations were already well established in the Biotechnology laboratory (Biochemistry, University of Stellenbosch) where the experiments were carried out.

The preferred mode of operation for cultivation of *P. pastoris* is fed-batch. This was an obvious choice as the inducer (methanol) was only added to the fermentation once sufficient biomass had accumulated. Fed-batch feeding of glycerol after the initial batch phase further increased biomass while adapting the culture to limited feeding and switch-over to methanol. Due to the toxic effect of high methanol concentrations on the culture, it seemed prudent to add methanol in a fed-batch mode.

The size of bioreactor used was influenced mainly by two factors. Firstly, availability and therefore ability to conduct more than one fermentation at a time necessitated the use of the 1.3 L NBS bioreactors (FIG 3.3). Secondly the current infrastructure in the laboratory did not allow for the use of 10 L bioreactors as massive heat generation associated with high biomass formation could not be controlled with the current water or ethanol cooling systems available.



**FIGURE 3.3: Experimental setup of two 1.3 L NBS bioreactors.**

### **3.2.2 Factors influencing product yield and the effect on experimental development**

Strain dependent factors have been discussed in depth in the literature study, but it is necessary to elucidate why we were only able to vary a few of these factors. Three strains were available in the Mut<sup>+</sup> phenotype. This allowed us to evaluate the effect of different gene constructs on the production of HPV16 L1 proteins and VLPs. It was, however, not possible to evaluate the effect of methanol phenotype on the production levels. We were able to obtain, through antibiotic screening, up to four copies of the h-L1 gene in a Mut<sup>+</sup> strain and this could be used in comparison with a single copy strain to determine the effect of gene dosage on expression of heterologous proteins.

With the cultivation dependent parameters it was decided to first focus on the most clear-cut fermentation method as described by Higgins et al., (1998), namely constant

(CMF) or predetermined (PMF) methanol feed rates. This would be a base case scenario and hopefully all other cultivation methods could be measured against this method. It would also closely imitate the feeding profile one would use for a Mut<sup>s</sup> strain as this phenotype is preferred for intracellular production. The methanol feeding strategy and its effect on the dissolved oxygen concentration has been known to have a great effect on production capacity in *Pichia* fermentations (Romanos, 1995). This has been established through literature and was one of the main focuses in this study of the production of HPV16 L1 VLPs.

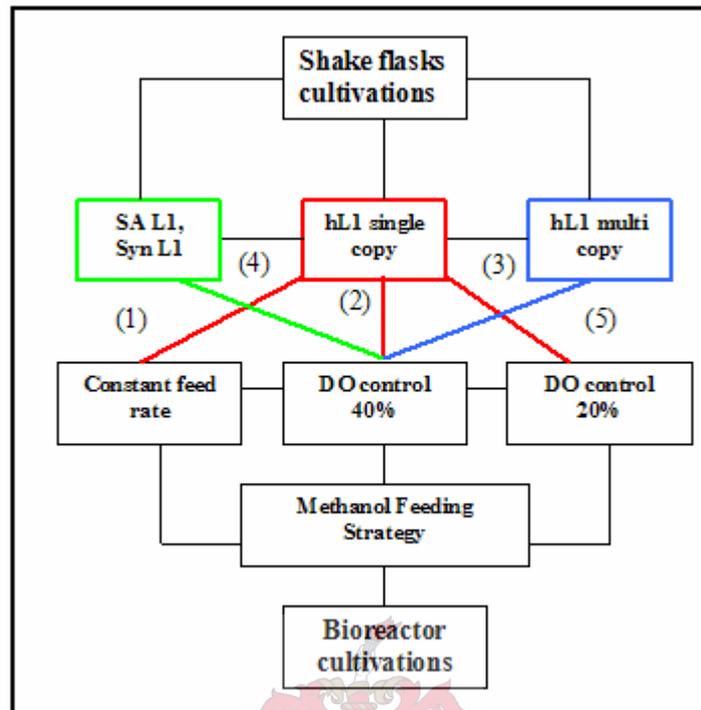
The second control strategy namely DO control would keep DO levels above a predetermined set point and it was decided that because this method gave very good results in other instances to compare both the single and multi copy gene integration to each other with this method. Based on the outcome of these experiments one could then decide on further parameter changes such as DO set point and strain to optimize production .

### 3.2.3 Experimental Design

Shake flask experiments were conducted to determine whether expression of the heterologous protein occurred. This was done for a range of methanol concentrations and for all strains. As the control, a strain containing the lacZ gene (to produce  $\beta$ -galactosidase), did not show any  $\beta$ -galactosidase activity it was determined that conditions were not adequate for induction.

The experimental design is shown in FIG 3.4. After unsuccessful shake flask cultivations the 40% DO control strategy was used to compare different strains. Subsequently feeding strategies for the best strain was compared. The single copy X33[h-L1] strain was cultivated with predetermined feed rates (1) and with DO control at a 40% DO set point (2). The multi copy h-L1 strain was cultivated under 40% DO control (3). At this point we were able to quantify expression of HPV L1 monomers as well as VLPs. It was obvious from calculated data that the single copy X33[h-L1] strain had to be optimised further and to ensure that sufficient induction occurred the dissolved oxygen set point was lowered to 20% DO (4). This increased the volume and interval at which methanol was fed. Both the X33[SA-L1] and the

X33[Syn-L1] strains were subjected to bioreactor cultivations (5) and data obtained compared to previous cultivations.



**FIGURE 3.4: Experimental design strategy including both shake flask and bioreactor cultivations (1-5 discussed in text).**

All cultivations were carried out in triplicate except for those with the X33[Syn-L1] and X33[SA-L1] strain where time constraints limited experiments to only one cultivation per strain. The single copy X33[h-L1] strain cultivation at 20% DO was carried out in duplicate.

### 3.2.4 Experimental analysis

As already described both shake flasks and bioreactors were used to determine HPV16 L1 production levels. The following parameters were investigated in this study:

- 1) Strains with different L1 gene constructs to determining success of developed constructs.
- 2) Multiple integration of the h-L1 gene to determine the effect of gene dosage on VLP production.
- 3) The effect of methanol feeding strategy on expression.

The effect of parameter changes (TABLE 3.1) were evaluated by completing the following analysis (a \* indicates which analyses were done for shake flask experiments):

- i) Intracellular HPV16, monomer and VLPs ( $\mu\text{g}\cdot\text{L}^{-1}$ ;  $\mu\text{g}\cdot\text{g}_{\text{DCW}}^{-1}$ )\*
- ii) Yield of product on carbon source ( $\mu\text{g}\cdot\text{L}^{-1}\cdot\text{L}^{-1}_{\text{methanol}}$ )
- iii) Total intracellular protein ( $\text{mg}\cdot\text{L}^{-1}$ )\*
- iv) Biomass growth curve (Optical Density at  $\text{OD}_{600}$ , Wet cell weight (WCW), Dry cell weight (DCW),  $\text{g}_{\text{DCW}}\cdot\text{L}^{-1}$ )\*

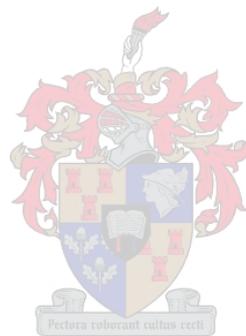
Of these parameters the main evaluation criteria for determining an increased performance was volumetric ( $\mu\text{g}\cdot\text{L}^{-1}$ ) and specific production levels ( $\mu\text{g}\cdot\text{g}_{\text{CDW}}^{-1}$ ). The other measured and calculated values were used to establish possible reasons for performance and to further differentiate between different cultivation methods.

TABLE 3.2 shows a summary of parameters changed in both the shake flasks and bioreactors.

**TABLE 3.2: Parameters varied in cultivations**

Cultivation type	Process parameter	Range
Shake flasks	Different gene constructs	X33[SA-L1], X33[Syn-L1], X33[h-L1], X33[multi-h-L1]
	Methanol concentration	0.5%, 1%, 2%, 3%
Bioreactor		Constant flow rate
	Methanol feeding control strategy	DO control (20%) DO control (40%)
	Different gene constructs	X33[SA-L1], X33[Syn-L1], X33[h-L1], X33[multi-h-L1]

HPV16 L1 monomer and VLP quantification was carried out for both the cytoplasmic (soluble) fraction and the membrane (insoluble) fraction. Because HPV16 L1 protein has not been produced in *P. pastoris* before, it was unknown where the heterologous product would be located. We did however know that the protein did not have a secretion signal and therefore would not accumulate extracellularly. The native virion has a tendency to accumulate in the cytoplasm.



|

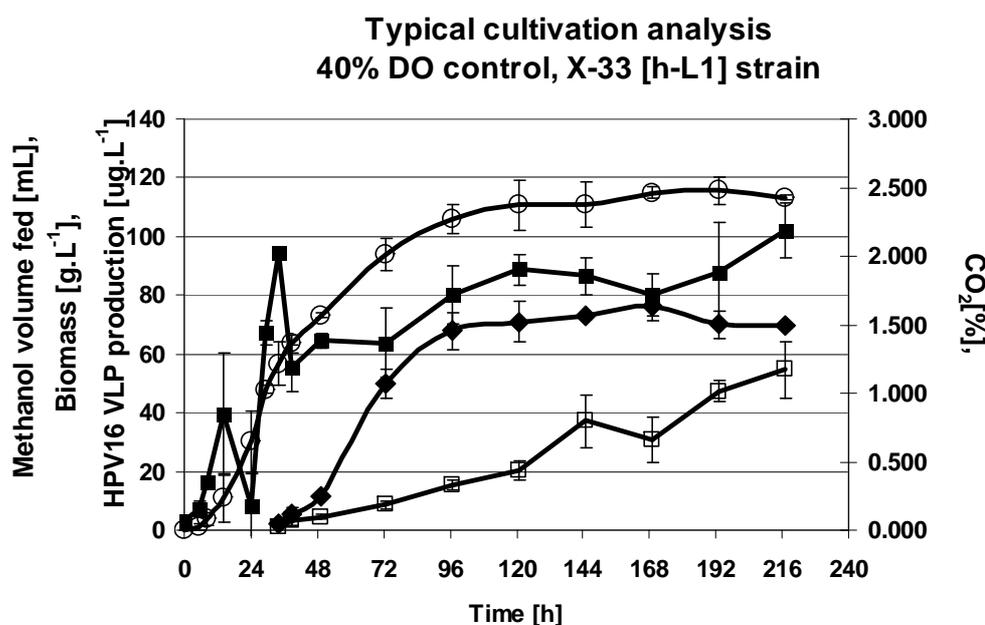
## 4. Results

### 4.1 Introduction

The effect of gene copy number, gene codon bias and cultivation strategy on HPV monomer and VLP production by *P. pastoris* was determined in fed-batch culture. Volumetric and specific monomer and VLP production levels, biomass accumulation, HPV VLPs per total protein production, yield of product on substrate and substrate consumption rates were compared. The ratio between VLPs and monomers was investigated and the location of the VLPs within the yeast cell was determined.

### 4.2 Typical cultivation analysis

The progression of a typical DO controlled cultivation is displayed in FIG 4.1. At the end of the glycerol batch phase (24 h) the biomass reached levels of approximately  $30.2 \pm 4.58 \text{ g}_{\text{DCW}}\cdot\text{L}^{-1}$ . The exponential growth phase (not indicated in FIG 4.1) followed a ten hour lag phase. After 5.5 h of the glycerol fed-batch phase, the biomass had accumulated to  $47.89 \pm 1.47 \text{ g}_{\text{DCW}}\cdot\text{L}^{-1}$ . This was sufficient biomass to commence the methanol fed-batch phase. At the end of the cultivation biomass reached a level of  $115.68 \pm 4.75 \text{ g}_{\text{DCW}}\cdot\text{L}^{-1}$ . Methanol feed rate increased from induction until approximately 120 h to levels of  $71.03 \pm 6.09 \text{ mL}\cdot\text{day}^{-1}$  from where it stayed constant. Carbon dioxide levels followed a trend similar to that of the methanol feeding and levelled off when the methanol feed rate did not increase. VLP production started when induction took place and VLP levels steadily increased until the end of the cultivation at 217 h. A typical CMF control strategy fermentation had the same glycerol batch and glycerol fed-batch phases and only differed in the methanol induction phase. The methanol feed rate stayed constant in the methanol induction period at the predetermined flow rate of  $3.6 \text{ mL}\cdot\text{h}^{-1}$  for the first 48 h of induction and was then decreased to  $1.8 \text{ mL}\cdot\text{h}^{-1}$  for the rest of the induction phase.



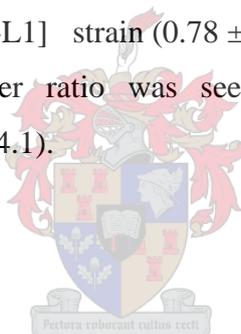
**FIGURE 4.1 : Typical fermentation analysis of the X33[h-L1] strain under 40% DO control. CO<sub>2</sub> (%) [■], Methanol addition (mL) [♦], HPV16 VLP Production [◊], Cell dry weight (g<sub>DCW</sub>.L<sup>-1</sup>) [○]. Error bars indicate one standard deviation.**

### 4.3 Comparison of the effect of different strain constructs and gene dosage

In TABLE 4.1 a comparison was made between the three gene constructs and the effect of gene dosage and codon bias under 40% DO control. The humanized gene construct has a significantly lower CAI (0.581) than the Syn-L1 (0.713) and the SA-L1 (0.719) genes. The humanized gene also has a much higher G+C content (62%) in comparison with the Syn-L1 (47.5%) and SA-L1 genes (49%).

No significant difference was noted (TABLE 4.1) for the specific VLP production levels ( $\mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}$ ) between the X33[h-L1] and the X33[multi-h-L1] strain under 40% dissolved oxygen (DO) control, while there was a significantly lower specific VLP production level for the X33[Syn-L1] strain ( $0.2 \pm 0.011 \mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}$ ). The X33[SA-L1] strain produced no measurable VLP levels. A significant difference in specific monomer levels between the X33[multi-h-L1] and the X33[h-L1] strain existed, while no significant difference existed in the specific monomer levels for the X33[h-L1] ( $0.48 \pm 0.012 \mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}$ ) and X33[Syn-L1] ( $0.50 \pm 0.02 \mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}$ ) strain (TABLE 4.1). The VLP yield on methanol ( $\mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}\cdot\text{L}^{-1}_{\text{methanol}}$ ) for the X33[Syn-L1] strain was much lower than for the other strains ( $0.34 \pm 0.019 \mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}\cdot\text{L}^{-1}_{\text{methanol}}$ ) (refer to TABLE 4.1). The methanol consumption rate ( $\text{ml}\cdot\text{h}^{-1}$ ) for the X33[Syn-L1]

and X33[SA-L1] strains under 40% DO control ( $3.16$  and  $3.18 \text{ ml.h}^{-1}$ ) respectively was not statistically lower than for the X33[h-L1] ( $3.42 \pm 0.24 \text{ ml.h}^{-1}$ ) and X33[multi-h-L1] strains ( $3.48 \pm 0.26 \text{ ml.h}^{-1}$ ) under the same control strategy (TABLE 4.1). A notable increase in biomass was observed for the X33[Syn-L1] and the X33[SA-L1] strains relative to the X33[h-L1] and X33[multi-h-L1] strains (TABLE 4.1). The biomass increase during induction was very similar between different strains as evident in comparable average linear rate of biomass increases ( $\text{g}_{\text{DCW}}\cdot\text{L}^{-1}\cdot\text{h}^{-1}$ ). The average biomass yield did not differ considerably between the 40% DO control with the X33[h-L1] or X33[multi-h-L1] strains ( $0.146 \pm 0.0033 \text{ g}_{\text{DCW}}\cdot\text{g}^{-1}_{\text{methanol}}$  vs.  $0.122 \pm 0.0061 \text{ g}_{\text{DCW}}\cdot\text{g}^{-1}_{\text{methanol}}$ ). The average biomass yield of both the X33[Syn-L1] and X33[SA-L1] strains do not differ statistically under 40% DO control and are  $0.135 \pm 0.005 \text{ g}_{\text{DCW}}\cdot\text{g}^{-1}_{\text{methanol}}$  and  $0.131 \pm 0.0066 \text{ g}_{\text{DCW}}\cdot\text{g}^{-1}_{\text{methanol}}$  respectively. Compared to the X33[Syn-L1] strain ( $0.4 \pm 0.006$ ) a higher VLP to monomer ratio was observed in the X33[h-L1] ( $0.76 \pm 0.04$ ) and X33[multi-h-L1] strain ( $0.78 \pm 0.02$ ), although no significant difference in VLP to monomer ratio was seen between the X33[h-L1] and X33[multi-h-L1] strain (TABLE 4.1).



**TABLE 4.1:** Comparison of the effect of different strain constructs and gene dosage under 40% DO control\*.

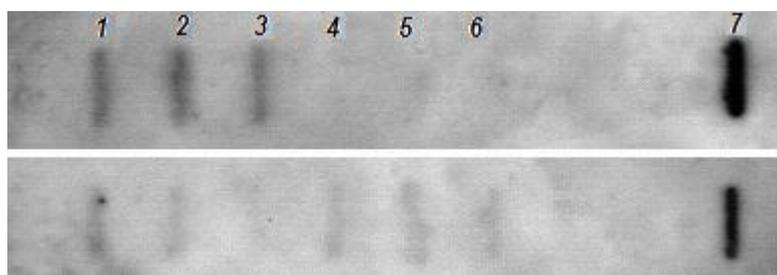
<b>STRAIN</b>	X-33[h-L1]	X-33[Syn-L1]	X-33[SA-L1]	X-33[multi-h-L1]
<b>Codon Adaptation Index (CAI)</b>	0.581	0.713	0.719	0.581
<b>G+C content (%)</b>	62	47.5	49	62
<b>Specific VLP production</b> [ $\mu\text{g}\cdot\text{g}^{-1}_{\text{dcw}}$ ]	$0.36 \pm 0.026$	$0.2 \pm 0.011$	0**	$0.32 \pm 0.018$
<b>Specific L1 monomer production</b> [ $\mu\text{g}\cdot\text{g}^{-1}_{\text{dcw}}$ ]	$0.48 \pm 0.012$	$0.50 \pm 0.021$	0**	$0.41 \pm 0.01$
<b>HPV VLP per total protein</b> [ $\text{ng}\cdot\text{mg}^{-1}$ ]	$7.43 \pm 0.11$	$8.94 \pm 0.13$	0**	$3.59 \pm 0.42$
<b>VLP Yield on methanol</b> [ $\mu\text{g}\cdot\text{g}^{-1}_{\text{dcw}}\cdot\text{L}^{-1}_{\text{methanol}}$ ]	$0.72 \pm 0.052$	$0.34 \pm 0.019$	0**	$0.65 \pm 0.004$
<b>Biomass</b> [ $\text{g}_{\text{dcw}}\cdot\text{L}^{-1}$ ]	$115.68 \pm 4.74$	$127.88 \pm 0.45$	$124.52 \pm 1.26$	$112.19 \pm 3.13$
<b>Average linear rate of biomass increase</b> [ $\text{g}_{\text{dcw}}\cdot\text{L}^{-1}\cdot\text{h}^{-1}$ ]	$0.348 \pm 0.006$	0.344	0.346	$0.35 \pm 0.011$
<b>Average Methanol consumption rate</b> [ $\text{ml}\cdot\text{h}^{-1}$ ]	$3.42 \pm 0.24$	3.16	3.18	$3.48 \pm 0.26$
<b>Average biomass yield</b> [ $\text{g}_{\text{DCW}}\cdot\text{g}^{-1}_{\text{methanol}}$ ]	$0.126 \pm 0.0062$	$0.135 \pm 0.005$	$0.131 \pm 0.0066$	$0.122 \pm 0.0061$
<b>VLP : monomer ratio</b>	$0.76 \pm 0.04$	$0.40 \pm 0.006$	0**	$0.78 \pm 0.02$

\*Statistical analysis: Single factor ANOVA done to evaluate significance of differences observed between variables.  $\pm$  value indicates one standard deviation. (See Appendix for ANOVA analysis)

\*\* Below detectable levels

When comparing mRNA levels of the X33[h-L1] and X33[multi-h-L1] strain under 40% DO control (FIG 4.2) higher band intensities were observed for bands 1 to 3 (X33[h-L1]) compared to bands 4 to 6. The *ACT1* control showed similar band

intensities for both the X33[h-L1] and the X33[multi-h-L1] strains, indicating that equivalent amounts of total mRNA were introduced to the slot blot.

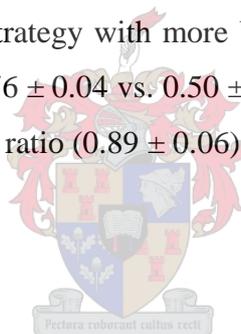


**FIGURE 4.2: mRNA slot blots comparing mRNA concentration between X33[h-L1] and X33[multi-h-L1]. Top: Bands 1-3 are a triplicate of the X33[h-L1] 40% DO control strategy mRNA followed by bands 4-6 of X33[multi-h-L1] mRNA under the 40% DO control strategy. Bottom: Comparison of *ACT1* concentrations for the corresponding X33[h-L1] and X33[multi-h-L1] strain mRNA. Band number 7 is the positive control for the h-L1 mRNA at the top and *ACT1* at the bottom.**

#### 4.4 Comparison of the effect of methanol addition strategy on the X33[h-L1] strain

The effect of the methanol addition strategy was investigated in the X33[h-L1] strain (TABLE 4.2) by comparing CMF, 40% DO and 20% DO control strategies. The specific VLP production level ( $\mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}$ ) was not significantly affected by the methanol addition strategy while a significant difference in specific monomer levels between three methanol addition control strategies existed. The 20% DO control strategy had the highest specific concentration of monomers ( $0.67 \pm 0.011 \mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}$ ) while the CMF (Constant methanol feed rate) control strategy had the lowest specific monomer concentration ( $0.38 \pm 0.016 \mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}$ ) (TABLE 4.2). The HPV VLPs per total protein ( $\text{ng}\cdot\text{mg}^{-1}$ ) for the 20% DO strategy was significantly higher ( $11.01 \pm 0.66 \text{ ng}\cdot\text{mg}^{-1}$ ) than for the CMF ( $4.8 \pm 0.4 \text{ ng}\cdot\text{mg}^{-1}$ ) and the 40% DO strategy ( $7.43 \pm 0.11 \text{ ng}\cdot\text{mg}^{-1}$ ). The total proteins per biomass levels were the same for all the control strategies; therefore the increase in VLPs per total protein was due to increased VLP production levels per unit biomass. Although no significant difference existed between the HPV VLP yields on methanol ( $\mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}\cdot\text{L}^{-1}_{\text{methanol}}$ ) for the different control strategies, the biomass levels obtained differed significantly. The 20% DO control strategy had the highest biomass level ( $128.66 \pm 5.5 \text{ g}_{\text{DCW}}\cdot\text{L}^{-1}$ ) followed by a much lower level for the 40% DO control ( $115.68 \pm 4.74 \text{ g}_{\text{DCW}}\cdot\text{L}^{-1}$ ). The CMF control strategy had the lowest biomass level at  $104.2 \pm 2.08 \text{ g}_{\text{DCW}}\cdot\text{L}^{-1}$ .

(TABLE 4.2). Significant differences could also be observed in the average linear rate of biomass increases ( $g_{DCW}\cdot L^{-1}h^{-1}$ ) (TABLE 4.2). The 20% DO control strategy had the highest average linear rate of biomass increase ( $0.388 \pm 0.0038 g_{DCW}\cdot L^{-1}h^{-1}$ ) followed by the 40% DO control strategy ( $0.348 \pm 0.006 g_{DCW}\cdot L^{-1}h^{-1}$ ) and lastly the CFM control strategy at  $0.255 \pm 0.0035 g_{DCW}\cdot L^{-1}h^{-1}$ . The methanol consumption rate ( $ml\cdot h^{-1}$ ) mirrored this trend; the CFM control strategy had the lowest methanol consumption rate ( $1.8 ml\cdot h^{-1}$ ). A significant difference in methanol consumption existed between the 20% DO ( $3.70 \pm 0.05 ml\cdot h^{-1}$ ) and 40% DO control strategy ( $3.42 \pm 0.14 ml\cdot h^{-1}$ ) (TABLE 4.2). The average biomass yield on methanol follows a similar trend with the 20% DO control strategy having a higher average biomass yield ( $0.146 \pm 0.0033 g_{DCW}\cdot g^{-1}_{methanol}\cdot L^{-1}$ ) than both the 40% DO ( $0.126 \pm 0.0062 g_{DCW}\cdot g^{-1}_{methanol}\cdot L^{-1}$ ) and the CMF control strategies ( $0.1 \pm 0.005 g_{DCW}\cdot g^{-1}_{methanol}\cdot L^{-1}$ ) (TABLE 4.2). In terms of the amount of VLPs formed from monomers (TABLE 4.2), there existed a significant difference between the 40% DO and the 20% DO strategy with more VLPs forming from monomers in the 40% DO control strategy ( $0.76 \pm 0.04$  vs.  $0.50 \pm 0.03$ ). The CMF control strategy had the highest VLP to monomer ratio ( $0.89 \pm 0.06$ ).



**TABLE 4.2:** Comparison of the effect of methanol addition strategy on the X33[h-L1] strain.

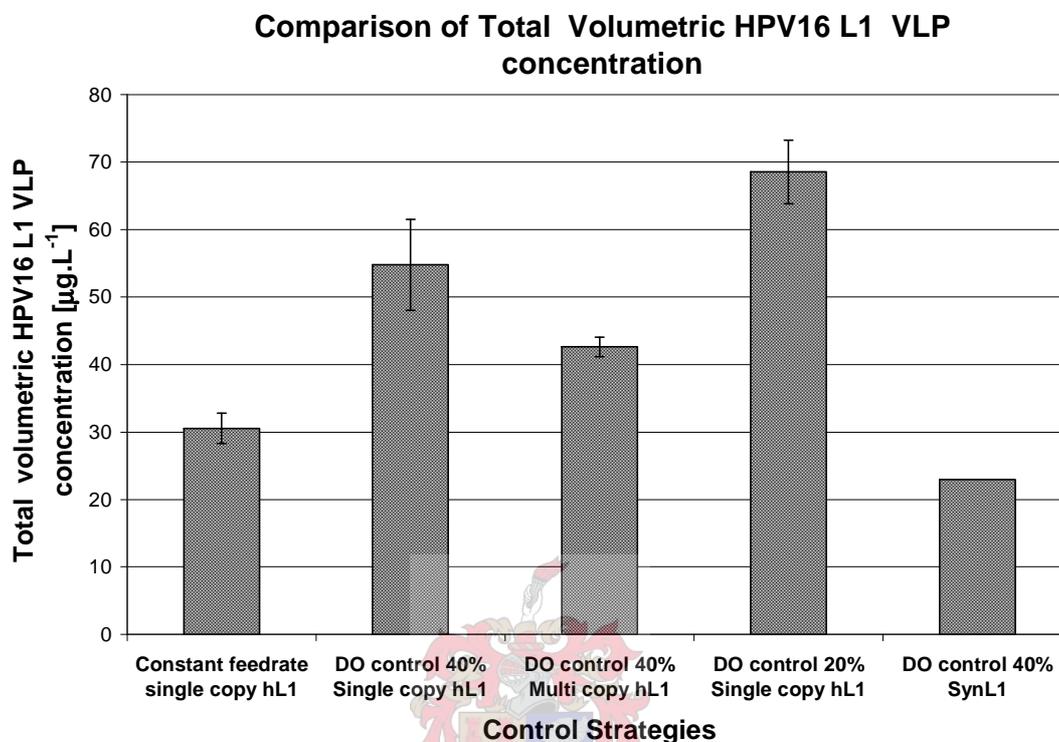
<b>CONTROL STRATEGY</b>	<b>CMF</b>	<b>40% DO</b>	<b>20% DO</b>
<b>Specific VLP production</b> [ $\mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}$ ]	0.30 $\pm$ 0.037	0.36 $\pm$ 0.026	0.33 $\pm$ 0.027
<b>Specific L1 monomer production</b> [ $\mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}$ ]	0.38 $\pm$ 0.016	0.48 $\pm$ 0.012	0.67 $\pm$ 0.011
<b>HPV VLP per total protein</b> [ $\text{ng}\cdot\text{mg}^{-1}$ ]	4.8 $\pm$ 0.4	7.43 $\pm$ 0.11	11.01 $\pm$ 0.66
<b>VLP Yield on methanol</b> [ $\mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}\cdot\text{L}^{-1}_{\text{methanol}}$ ]	0.6 $\pm$ 0.075	0.72 $\pm$ 0.0522	0.56 $\pm$ 0.051
<b>Average Methanol consumption rate</b> [ $\text{ml}\cdot\text{h}^{-1}$ ]	1.8	3.42 $\pm$ 0.14	3.70 $\pm$ 0.05
<b>Biomass</b> [ $\text{g}_{\text{DCW}}\cdot\text{L}^{-1}$ ]	104.20 $\pm$ 2.08	115.68 $\pm$ 4.74	128.66 $\pm$ 5.5
<b>Average linear rate of biomass increase</b> [ $\text{g}_{\text{DCW}}\cdot\text{L}^{-1}\cdot\text{h}^{-1}$ ]	0.255 $\pm$ 0.0035	0.348 $\pm$ 0.006	0.388 $\pm$ 0.0038
<b>Average biomass yield</b> [ $\text{g}_{\text{DCW}}\cdot\text{g}^{-1}_{\text{methanol}}$ ]	0.100 $\pm$ 0.005	0.126 $\pm$ 0.0062	0.146 $\pm$ 0.0033
<b>VLP : monomer ratio</b>	0.89 $\pm$ 0.06	0.76 $\pm$ 0.04	0.50 $\pm$ 0.03

\*Statistical analysis: Single factor ANOVA done to evaluate significance of differences observed between variables.  $\pm$  value indicates one standard deviation. (See Appendix for ANOVA analysis)

#### 4.5 Volumetric production levels

The volumetric production capacity of the VLP production system is of high importance for commercial manufacturing. The X33[h-L1] strain under 20% DO control performed the best of all the strains and control strategies with VLP levels peaking at  $68.53 \pm 4.74 \mu\text{g}\cdot\text{L}^{-1}$  after 187h of induction (FIG 4.3). There was a statistically significant difference in volumetric production of VLPs in the X33[h-L1] ( $54.76 \pm 6.73 \mu\text{g}\cdot\text{L}^{-1}$ ) and X33[multi-h-L1] ( $42.6 \pm 1.43 \mu\text{g}\cdot\text{L}^{-1}$ ) strains under 40% DO control. The X33[h-L1] strain under CMF produced VLP levels of  $30.53 \pm 2.25 \mu\text{g}\cdot\text{L}^{-1}$ . The X33[Syn-L1] strain under 40% DO control produced low

volumetric VLP levels ( $22.96 \pm 4.55 \mu\text{g}\cdot\text{L}^{-1}$ ), while no VLP production was observed with the X33[SA-L1] strain. In all cases volumetric monomer levels exceeded VLP levels.



**FIGURE 4.3: Comparison of total volumetric HPV16 L1 VLP concentrations of different strains and cultivation strategies. Error bars indicate one standard deviation.**

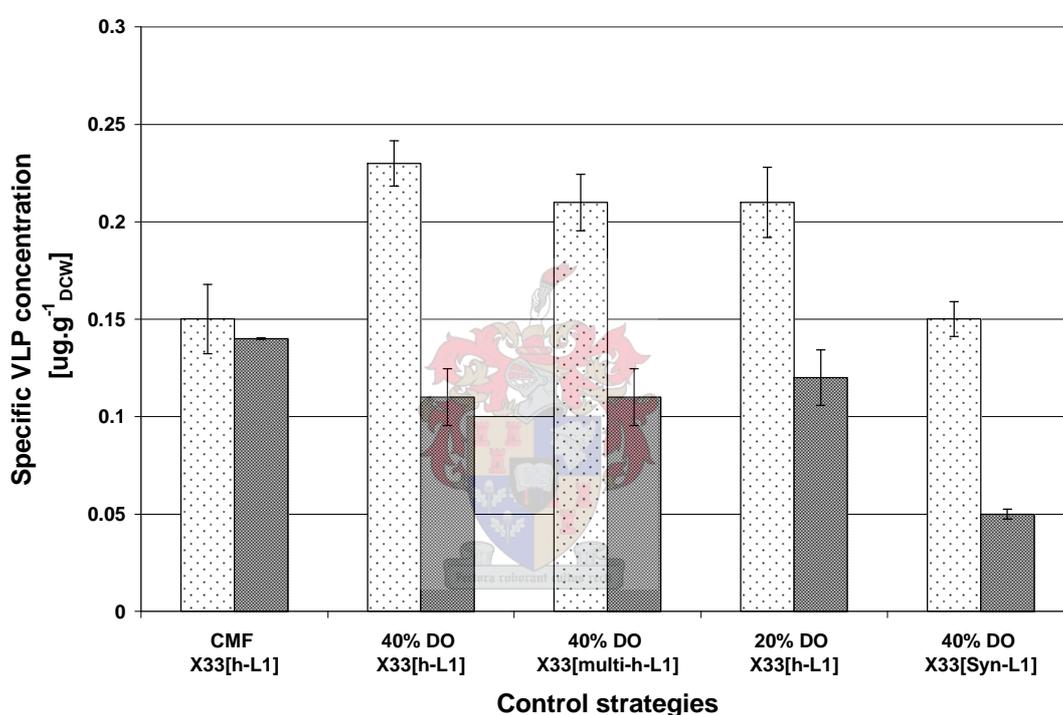
#### 4.6 Assembly efficiency and location of the VLPs

In all the cases studied the amount of VLPs found in the cytoplasm exceeded those found in the membrane fraction (FIG 4.4). A significant difference between the membrane and cytoplasmic specific VLP concentration existed under all conditions and strains investigated except with the CMF control strategy (FIG 4.4). Similar specific VLP concentrations were found in the membrane ( $0.14 \pm 0.0004 \text{ ug}\cdot\text{g}_{\text{DCW}}^{-1}$ ) and cytoplasmic fraction ( $0.15 \pm 0.018 \text{ ug}\cdot\text{g}_{\text{DCW}}^{-1}$ ).

No significant differences between the location of the VLPs in the 40% and 20% DO control strategies were found (40% DO: membrane  $0.11 \pm 0.0146 \text{ ug}\cdot\text{g}_{\text{DCW}}^{-1}$ , cytoplasm  $0.21 \pm 0.0268 \text{ ug}\cdot\text{g}_{\text{DCW}}^{-1}$ ; 20% DO: membrane  $0.12 \pm 0.0142 \text{ ug}\cdot\text{g}_{\text{DCW}}^{-1}$ , cytoplasm  $0.23 \pm 0.0117 \text{ ug}\cdot\text{g}_{\text{DCW}}^{-1}$ ), but a significantly higher amount of VLPs were

located in the membrane fraction ( $0.15 \pm 0.018 \text{ ug.g}_{\text{DCW}}^{-1}$ ) under the CMF control strategy (FIG 4.4).

In terms of different strains, the X33[Syn-L1] strain had significantly lower concentrations of VLPs in both the membrane ( $0.05 \pm 0.0025 \text{ ug.g}_{\text{DCW}}^{-1}$ ) and the cytoplasmic fraction ( $0.15 \pm 0.0089 \text{ ug.g}_{\text{DCW}}^{-1}$ ) under 40% DO control strategy than the X33[h-L1] strain. A comparison of location of the specific VLP concentration of the X33[h-L1] and the X33[multi-h-L1] strain showed no statistical difference.



**FIGURE 4.4: Specific HPV16 L1 VLP concentration in cytoplasmic and membrane fraction. (▨)Cytoplasmic fraction, (■)Membrane fraction. Error bars indicate one standard deviation.**

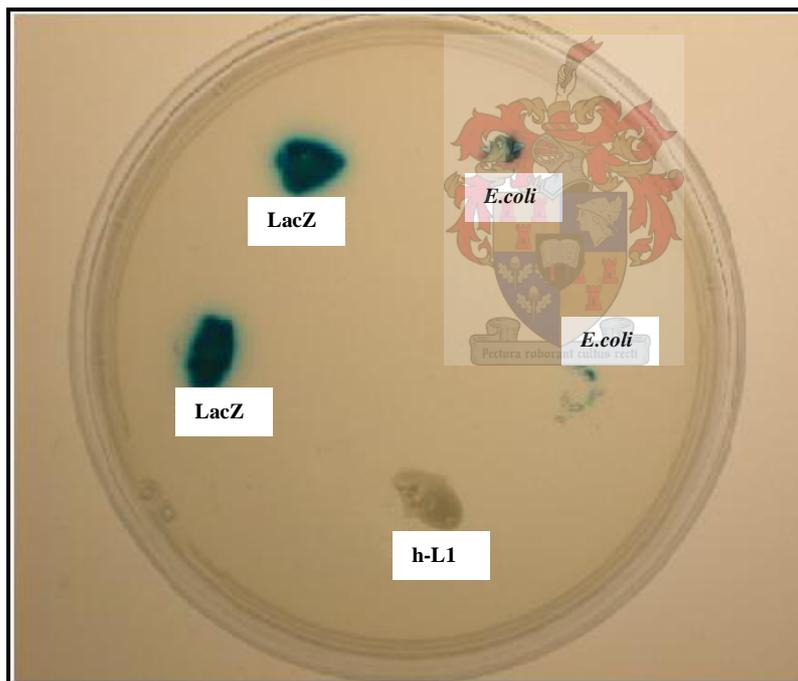
#### 4.7 Shake flask experiments

Shake flask experiments were carried out for all three single copy constructs and GS115[lacZ] control strain (producing intracellular  $\beta$ -galactosidase) at a 0.5, 1, 2, 3% methanol concentration prior to bioreactor experiments to investigate methanol induction efficiency. GS115[lacZ] was used as a positive control for induction.  $\beta$ -galactosidase assays showed no activity in the GS115[lacZ] strain compared to the pBluescript *E.coli* control. ELISAs of intracellular protein extracts from the strains containing the HPV16 gene indicated that there were no discernable levels of HPV16

L1 protein present either as VLPs or as monomers. A possible explanation was insufficient induction resulting in no expression of both  $\beta$ -galactosidase and the HPV16 L1 proteins. It was suggested that  $\beta$ -galactosidase plate assays should be carried out to ascertain whether induction occurred on plates.

#### 4.8 $\beta$ -galactosidase plate assays

Blue zones, indicating  $\beta$ -galactosidase activity, surrounded the *LacZ* expressing strain after one day on all the plates at 1, 2 and 3% methanol concentrations (FIG 4.5). The negative control (X33[h-L1] strain) showed no blue zone. It was not possible to quantify which methanol concentration led to the most efficient induction of the *LacZ* gene. It was however evident that induction conditions were sufficient on the plates but not in shake flasks. No further quantification of HPV VLP production levels in shake flasks was undertaken.



**FIGURE 4.5:  $\beta$ -galactosidase plate assays (blue zones indicate induction occurring and  $\beta$ -galactosidase being produced).**

## 5. Discussion

### 5.1 Introduction

HPV6 L1 VLPs were previously successfully produced in *P. pastoris* (Li *et al.*, 2003), while to our knowledge HPV16 L1 VLPs have not previously been produced in *P. pastoris*. The production of HPV16 L1 VLPs by recombinant expression in *P. pastoris* was therefore investigated as a production system and compared to the expression systems currently used for expression of HPV16 VLPs, namely baculovirus infected insect cells and *S. cerevisiae*. Three different gene constructs were investigated as well as a strain with multiple gene integrations. During the bioreactor cultivations three different induction strategies were applied. The effect of two different dissolved oxygen concentrations on heterologous protein production in the dissolved oxygen control strategy was investigated, while the second strategy involved a constant feed rate of the methanol inducer. Attention was given to where the VLPs were located intracellularly and the ratio of VLPs to monomer L1 proteins as an indication of VLP assimilation efficiency.

### 5.2 Comparison of strains with different gene constructs

#### 5.2.1 Codon adaptation, G+C content and the effect of Mut phenotype

The higher A+T content of the Syn-L1 and SA-L1 genes expressed in *P. pastoris* negatively affected HPV16 L1 VLP production. TABLE 4.1 indicates a high G+C content and subsequently a low A+T content for all genes transformed into the *P. pastoris* strains used in this study. The X33[h-L1] strain with a higher G+C content (62%) had much higher specific VLP concentrations in comparison with the X33[Syn-L1] and X33[SA-L1] strains. In the expression of HPV6 L1 proteins in *P. pastoris* (Li *et al.*, 2003), abolishing of the AT-rich regions (lowering the A+T content) on the gene proved crucial toward the successful production of the heterologous product. This supports the finding that the X33[h-L1] gene construct had a lower A+T content and therefore was a better HPV VLP producer than the other two gene constructs. In contrast, Hu and colleagues (2006) found that increasing the A+T content content in the design of their synthetic scFv gene from 44 to 57% had a positive effect on expression levels in *P. pastoris*. Sinclair and Choy (2002) state that codon bias in mammals tends towards a high G+C content and subsequently a lower A+T content. Transforming these genes into lower eukaryotes such as yeast may

affect mRNA secondary structure. As yeasts generally have more A+T biased sequences a high G+C content may actually decrease protein production, irrespective of the actual codon choice.

VLP production levels did not follow codon bias (CAI) under 40% DO control. Under 40% DO control the X33[h-L1] strain performed better than the X33[Syn-L1] or X33[SA-L1] strains (TABLE 4.1). A 2.39-fold difference in volumetric production level between the X33[h-L1] and X33[Syn-L1] strains support this observation (TABLE 4.1). The CAI as indicated in TABLE 4.1 is higher for the X33[Syn-L1] and SA-L1 genes than for the h-L1 gene ( $\pm 1.23$  fold higher). No detectable VLP production was observed in the X33[SA-L1] strain, while a small amount of monomers were detected. This may be explained by the fact that neither gene was optimized for expression in *P. pastoris* and may still contain codons that are rarely used in *P. pastoris* and would thus influence protein production negatively. Lu *et al.*, (2003) also found that use of preferred *Pichia* codons increased heterologous expression. Specific VLP levels in the X33[Syn-L1] strain were comparable, but still lower than the X33[h-L1] strains under any control strategy (TABLE 4.1). The same control strategies were applied to all three strains, therefore the inferior production results must be due to the different genes and their integration into the *P. pastoris* genome. Considering that the same cultivation strategy was used in all the strains, variations in specific VLP production could not be due to differences in biomass generation or methanol consumption rates. Average linear biomass increases were not significantly different, therefore G+C content of the different genes and their codon adaptation influenced production levels.

The Mut<sup>+</sup> methanol utilizing phenotype of the employed host strain may have contributed to low HPV16 VLP production levels. Cregg and co-workers (1987) found that the Mut<sup>+</sup> phenotype produced 10-fold less HBsAg than the Mut<sup>s</sup> phenotype. In the Mut<sup>+</sup> strain the maximum expression level obtained in a 0.1 L cultivation was 140  $\mu\text{g}\cdot\text{L}^{-1}$ . This correlates well with our study, where we obtained a maximum expression of 68.5  $\mu\text{g}\cdot\text{L}^{-1}$ . Vassileva and co-workers (2001) expressed HBsAg in *P. pastoris* using a Mut<sup>s</sup> phenotype. This indicates that for an intracellular

product that requires efficient assembly of monomer proteins to maintain immunogenicity the use of a Mut<sup>s</sup> or Mut<sup>r</sup> strain is favoured.

### 5.2.2 The effect of multiple gene integrations

The single copy h-L1 construct was superior to the multi copy h-L1 strain for the production of HPV16 L1 VLPs by *P. pastoris*. It is commonly believed that vector copy number is a determining factor affecting protein yield in intracellular foreign protein production (Vassileva *et al.*, 2001). When evaluating volumetric production levels (FIG 4.1) it was seen that the X33[multi-h-L1] strain had lower specific monomer and VLP concentrations compared to the X33[h-L1] strain, while average linear biomass increases, methanol consumption and VLP to monomer ratio were not significantly different. The higher levels of VLP production in the X33[h-L1] strain concurs with the higher levels of mRNA transcription observed in this strain compared to the X33[multi-h-L1] strain (FIG 4.2). Production differences between single and multi-copy gene integrations may be ascribed to the tempo of mRNA production and mRNA stability at steady state. The disparity in transcription levels (FIG 4.2) between the single and multi-copy gene integrations can be ascribed to gene integration factors. Multi-copy gene integration may have occurred in a genomic locus that is poorly transcribed. As the *ACT1* control gave similar band intensities for the X33[h-L1] and the X33[multi-h-L1], one would expect band intensities for the X33[multi-h-L1] strain to be of similar or increased values, although this was not the case in this study. Vassileva and co-workers (2001) studied the effect of increased copy number and found that increased copy number resulted in production levels that exceeded the limits of the particle assembly process. Unlike this study it was found that the particulate or VLP form increased with increasing copy number. This study seems to correlate better with that of Hu and co-workers (2006) who found that an increased copy number did not significantly increase expression levels.

Clare *et al.*, (1991) studied the effect of both gene dosage and Mut phenotype on expression levels and found that at one gene copy there was no significant difference between Mut phenotypes, but that at 9 gene copies the Mut<sup>s</sup> phenotype performed 20% better than the Mut<sup>r</sup> phenotype. A similar trend was found in the specific production levels (TABLE 4.1), although the difference was not statistically significant. The use of a Mut<sup>s</sup> phenotype for intracellular protein production in

*P. pastoris* is preferred and higher production levels are generally achieved with increased gene dosages in this phenotype in comparison with intact *AOX1* constructs (Cos *et al.*, 2005). TABLE 4.1 shows that the multi copy strain had lower biomass concentration than the single copy strain. This indicated that the decrease in volumetric production levels was in part due to the decreased biomass concentration observed for the X33[multi-h-L1] strain, but most likely to lower transcription level of the gene.

The amount of HPV monomers and VLPs per mg per biomass showed a marked decrease in the multi copy strain (TABLE 4.1) a consequence of the multiple gene integration. Vassileva and co-workers (2001) hypothesized that due to stress caused by induction with methanol, particulate deposition of HBsAg VLPs would increase in the membrane fraction. This was not the case in this study.

### **5.3 The influence of methanol addition control strategy**

Investigation of the volumetric production levels ( $\mu\text{g.L}^{-1}$ ) in FIG 4.1 showed that the 20% DO strategy produced the highest volumetric monomer ( $104.82 \mu\text{g.L}^{-1}$ ) and VLP concentrations ( $68.53 \mu\text{g.L}^{-1}$ ) due to the positive effect on biomass production, while HPV VLP yield per unit biomass was not affected (TABLE 4.2). The level of methanol induction was therefore similar in all control strategies and the main reason for volumetric VLP production differences was the increased biomass production (TABLE 4.2), resulting in higher biomass levels observed for the 20% DO control strategy in comparison to the other strategies. Both the CMF control strategy (2.24-fold lower) and the 40% DO strategy (1.25-fold lower) produced lower volumetric HPV16 L1 VLP concentrations compared to the 20% DO control strategy.

The higher biomass levels were obtained with the 20% DO strategy (TABLE 4.2). This was caused by the feeding of methanol for longer intervals and more continuously than for the 40% DO strategy, resulting in a higher total amount of methanol fed during the culture. The lowest biomass concentration ( $104.2 \text{ g}_{\text{DCW}}.\text{L}^{-1}$ ) was obtained in the CMF strategy where methanol was fed at very low constant rates compared to the DO strategies, resulting in the lowest total amount of methanol

addition. After approximately 120 h the biomass concentration leveled off in the CMF strategy and no further increase in biomass was observed. This was because the methanol feeding strategy supplied methanol at too low rates and the culture growth was limited by the methanol feed rate. This trend was also observed for the 40% DO strategy (maximum biomass level:  $115.7 \text{ g}_{\text{DCW}}\cdot\text{L}^{-1}$ ) although this leveling off was observed at approximately 150 h. The slackening off in biomass generation was better explained in FIG 4.1 which indicates that the methanol feed rate remains more or less constant after 120 h. The 20% DO control strategy has a much higher maximum biomass level (TABLE 4.2) ( $128.7 \text{ g}_{\text{DCW}}\cdot\text{L}^{-1}$ ) and no leveling off in biomass production was observed.

The methanol consumption rate was significantly higher for the 20% DO control strategy (TABLE 4.2). A lower dissolved oxygen set-point resulted in an increased methanol feed rate, which subsequently caused biomass generation at a faster rate. The average methanol consumption per hour for the 20% DO control strategy ( $3.7 \pm 0.05 \text{ ml}\cdot\text{h}^{-1}$ ) was approximately twice that of the CMF control strategy ( $1.8 \text{ ml}\cdot\text{h}^{-1}$ ), with the 40% DO control strategy ( $3.42 \pm 0.014 \text{ ml}\cdot\text{h}^{-1}$ ) methanol consumption was slightly slower than for the 20% DO control strategy (TABLE 4.2). The maximum yield of product on methanol ( $\mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}\cdot\text{L}^{-1}_{\text{methanol}}$ ) was very similar for the CMF control and DO control strategies. The yield of biomass on methanol (TABLE 4.2) indicates that the 20% DO control strategy had the greatest yield of biomass on methanol followed by the 40% DO control and the CMF control. This contributed greatly to the higher levels of VLPs observed in the 20% DO control strategy.

#### **5.4 Location and assembly efficiency of VLPs**

VLPs mostly occurred in the cytoplasm at higher levels than was present in the membrane fraction (FIG 4.4). The most significant difference between localisation of VLPs in the different strain constructs were between the X33[Syn-L1] and the X33[h-L1] strain. More VLPs (1.76-fold) occurred in the cytoplasm of the X33[h-L1] strain than in the X33[Syn-L1] strain. This was supported by the lower (1.85-fold) VLP to monomer ratio observed in the X33[Syn-L1] strain (TABLE 4.2). Therefore,

fewer VLPs were formed, which led to a lower VLP concentration in the cytoplasm. Similar concentrations of VLPs occurred in the membrane fraction. When comparing the X33[h-L1] and the X33[multi-h-L1] strains (FIG 4.3), no significant difference existed between the VLP concentrations in the cytoplasmic location and a notable decrease in VLPs located in the membrane of the X33[multi-h-L1] strain. No significant decrease in the VLP to monomer ratio was observed for the X33[multi-h-L1] strain relative to the X33[h-L1] strain.

Higher specific monomer concentrations for the DO control strategies relative to the CMF control strategy indicate better induction efficiency, while lower VLP to monomer ratios observed for the DO control strategies indicated a bottleneck in the assembly process. In terms of methanol addition control strategy the most notable difference in location of VLPs occurred between the 20% DO and the 40% DO control strategy (FIG 4.3). The concentration of VLPs located in the cytoplasm of the 20% DO control strategy was 1.48-fold lower than the VLP concentrations observed with the 40% DO control strategy. This correlated with a decrease in VLP to monomer ratio ( $0.48 \pm 0.032$  vs.  $0.4 \pm 0.018$ ) (TABLE 4.1) and supports the observation that at higher induction efficiencies a bottleneck in assembly occurs due to higher L1 monomer concentrations and VLPs are not assembled. The highest VLP to monomer ratio was observed with the CMF control strategy where methanol addition was much slower than for the DO control strategies. Therefore slower methanol addition produced lower L1 monomer concentrations and allowed time for the L1 monomers to assemble into VLPs.

The hydrophobicity of the L1 particles caused those that did not assemble into VLPs to associate with the cell membranes. FIG 2.4 indicated that there existed a large region of hydrophobicity on the L1 protein. It was also known from the study by Chen and co-workers (2000) that the pentamer-pentamer bonds were facilitated by the hydrophobicity of certain areas on the pentamer structure. Apparently, when pentameres did not find other pentameres to adhere to, they would attach to the cell membrane. The presence of higher L1 monomer concentrations in the 20% DO control strategy with the X33[h-L1] strain impaired the ability of the pentameres to find each other to assemble into VLPs and subsequently attached to the membrane. This was also illustrated in the X33[Syn-L1] strain where there was similar specific

L1 monomer concentrations but a lower VLP to monomer ratio ( $0.26 \pm 0.042$ ), indicating that there existed an ideal monomer concentration related to the folding capacity that facilitated VLP assembly so that the VLPs remained in the cytoplasm. In the native virion the L2 minor protein also has a hydrophobic region by which it binds to the HPV capsid (Finnen *et al.*, 2003). This attachment occurred in the hollow region the monomers make when a pentamer assembled. Therefore the HPV VLP that was assembled in the absence of the L2 protein had several sites on the surface of the VLP that should have distinct outward facing hydrophobic areas. This may have caused fully formed VLPs to attach to the membrane.

### 5.5 Shake flasks

Methanol induction was insufficient in shake flasks and the problem was compounded by low dissolved oxygen concentrations in the shake flask cultures. Oxygen transfer rate is one of the important factors in cultivation of *Pichia* and insufficient oxygen transfer in the shake flasks contributed to undetectable levels of heterologous gene expression in this system.

Shake flasks were baffled (3 indentations around the perimeter of the flasks) with loose cotton wool lugs covered by foil. The baffles increased mixing efficiency and dissolved oxygen levels in the flasks. In a 250 mL Erlenmeyer flask, a 50 mL starting volume was 20% of the flask volume. This was well within the suggested working volume of between 10 and 30%. Varying the methanol concentrations between the minimum necessary for induction (0.5%) and the maximum suggested concentration (3%) did not enhance HPV 16 L1 VLP production. Independent calibration of the shaking frame with a tachometer indicated that its maximum revolutions per minute were only 120 rpm (in contrast with the prescribed agitation of 300 rpm) and definitely contributed to no detectable HPV16 L1 VLP concentrations.

Another contributing factor to low expression levels was that the Mut<sup>+</sup> phenotype consumed methanol at a high rate, resulting in a quick loss of induction upon consumption of the inducing carbon source. This effect would result in the production of very low levels of heterologous product. A third contributing factor could be the low biomass concentrations associated with these shake flasks in comparison with

higher density in bioreactors (approx. 6 g<sub>DCW</sub>.L<sup>-1</sup> vs 120 g<sub>DCW</sub>.L<sup>-1</sup>). It was shown that even with high density cultivation very low concentrations of HPV16 L1 monomers or VLPs were present. In retrospect, HPV VLP levels in shake flasks, if any, would be below the detection level of the ELISA used in this study.

## 5.6 $\beta$ -galactosidase plate assays

The  $\beta$ -galactosidase plate assays verified that the *LacZ* gene was present in the GS115[LacZ] control strain and being induced. The blue zones (FIG 4.5) indicated that there was  $\beta$ -galactosidase activity. Therefore insufficient induction conditions in shake flasks inhibited heterologous protein production and necessitated the use of bioreactors where induction could be sufficiently maintained.

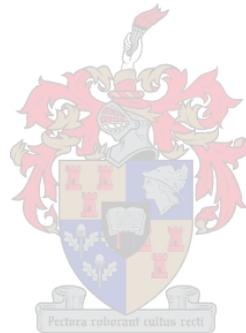
## 5.7 Benchmarking with alternative recombinant expression systems

HPV16 L1 VLP concentrations obtained in *P. pastoris* in this study are comparable to the study by Li et al., (2003), but much lower than expression levels obtained in baculovirus infected insect cells. However, based on the expression levels of HBsAg VLPs obtained in *P. pastoris*, this system has the ability to increase HPV VLP production capacity.

To our knowledge the only study on the production of HPV VLPs was done by Li et al., (2003) on HPV6. A maximum VLP concentration of 125  $\mu$ g.L<sup>-1</sup> was achieved. This level is comparable to our maximum VLP concentration of 68.5  $\mu$ g.L<sup>-1</sup>. We were unable to compare VLP concentration in *S. cerevisiae*, but through personal communication (E. Rybicki, Molecular and Biosciences Department, UCT, Cape Town) know that a concentration of approximately 1 g.L<sup>-1</sup> was obtained in the baculovirus infected insect cell system.

Although an HBV VLP vaccine is similar in concept to an HPV VLP vaccine, the HPV VLP is much more difficult to assemble and more complex than the HBsAg VLPs. A comparison of HPV VLP production levels with HBsAg production levels would, however, give an indication of what has been accomplished in *P. pastoris* with

other VLP vaccines. In a study by Cregg *et al.*, (1987) a maximum level of  $0.38 \text{ g.L}^{-1}$  HBsAg VLPs was obtained in a fed-batch system. In a later study by Chauhan and co-workers (1999) a maximum level of  $1 \text{ g.L}^{-1}$  was obtained with fed-batch cultivations. Compared to batch cultivation of *A. niger* (James *et al.*, 2007), in which HBsAg production levels reached maximum levels of  $3.6 \text{ mg.L}^{-1}$ . Maximum HBsAgVLP concentrations in *H. polymorpha* (also a methylotroph) were  $100 \text{ mg.L}^{-1}$  (R. Weyhenmeyer, Rhein Biotech, GmbH, Dusseldorf, Germany). It can be seen that *P. pastoris* has a significant capacity for VLP production and has proved itself as a valuable alternative expression system in the VLP vaccine arena.



## 6. Conclusions

This study has successfully proven that HPV16 VLPs and monomers can be expressed in *P. pastoris* at detectable levels. An investigation into the effect of codon bias, G+C content and gene dosage delivered the following results:

1. Despite a higher codon adaptation index for both the X33[Syn-L1] and X33[SA-L1] strains, the X33[h-L1] strain had significantly higher HPV16 VLP production levels. This discrepancy could be ascribed to the presence of rare codons on the gene that are rarely used in *P. pastoris*, negatively influencing heterologous protein production.
2. The lower A+T content of the h-L1 gene positively influenced HPV 16 VLP production in the X33[h-L1] strain.
3. Lower transcription levels in the X33[multi-h-L1] strain, as indicated by a mRNA slot-blot, caused decreased VLP production levels compared to a single gene integration. Lower transcription may be due to gene integration at a poorly transcribed genomic locus.
4. Multiple gene integration in Mut<sup>-</sup> or Mut<sup>s</sup> phenotype may still give increased HPV VLP production levels.

Subsequently, an examination of the effect on VLP production of three different methanol addition control strategies with the X33[h-L1] strain led us to conclude the following:

1. The 20% DO strategy produced the highest volumetric monomer ( $104.82 \mu\text{g.L}^{-1}$ ) and VLP concentrations ( $68.53 \mu\text{g.L}^{-1}$ ) and was therefore the preferred control strategy for HPV16 L1 VLP production.
2. Increased biomass concentrations in the DO control strategies due to greater biomass production and increased methanol consumption in the 20% DO control strategy resulted in higher volumetric VLP concentrations.
3. Increased HPV16 L1 monomer concentrations in the 40% and 20% DO control strategy indicated increased induction efficiency.

4. A decreased VLP to monomer ratio in the 20% and 40% DO control strategies indicated lower assembly efficiency and a bottleneck in the assembly process due to increased L1 monomer concentrations.
5. The CMF control strategy with a slower rate of methanol addition had a lower induction efficiency indicated by a lower HPV16 L1 monomer concentration.
6. During the 20% DO control strategy lower concentrations of VLPs were associated with the membrane fraction, indicating that higher L1 monomer concentrations facilitate VLP assembly in the cytoplasm and not in the membranes.
7. The HPV16 L1 monomers associated with the membranes due to hydrophobic areas on the L1 monomer when assembly to VLPs did not occur.

Sufficient induction of the Mut<sup>+</sup> strain was not possible in shake flasks although  $\beta$ -galactosidase plate assays confirmed that the  $\beta$ -galactosidase control strain (GS115[LacZ]) could be induced to produce  $\beta$ -galactosidase. Possible reasons for lack of induction were:

1. A contributing factor to low expression levels is that the Mut<sup>+</sup> phenotype consumes methanol at a high rate, resulting in a quick loss of induction upon consumption of the inducing carbon source. This effect would result in the production of very low to undetectable levels of heterologous product.
2. Insufficient oxygen transfer to the shake flask culture due to low shaking speeds.

HPV16 L1 VLP concentrations obtained in *P. pastoris* in this study are comparable to the study by Li et al., (2003), but much lower than expression levels obtained in baculovirus infected insect cells. Based on the expression levels of HBsAg VLPs obtained in *P. pastoris*, this system with the necessary recommended optimisation has the capacity for increased HPV VLP production ability.

## 7. Recommendations and Future prospects

*P. pastoris* has proven itself as heterologous host for a variety of both intracellular and extracellular proteins. *P. pastoris* can be cultivated in bioreactors with relative ease and obtain high density cultivations on a defined medium. HPV16 VLPs were produced in this system with a certain amount of success, subsequently the following strategy is proposed to possibly further enhance expression. This strategy is composed of two subsections namely molecular and process modifications.

### 7.1 Molecular modifications

One of the key features of the *Pichia* expression system is its ease of genetic manipulation and established molecular techniques similar to that of *S. cerevisiae*. Throughout the discussion reasons have been proposed for low expression levels in a host that is renowned for high intracellular expression levels. These modifications have been implemented successfully by other researchers as was highlighted in the literature study. In light of this wealth of information on molecular modifications I propose the following:

- Different **Mut phenotype**: Mut<sup>s</sup> or even Mut<sup>r</sup> have been proven to be successful at producing intracellular proteins. Specifically HBsAg VLP expression levels have been enhanced using this methodology.
- As gene dosage is an important factor in intracellular production, an increase of **gene dosage** within the new Mut phenotype is advised (Cos *et al.*, 2005).
- **Optimisation of codons** in the study by Li *et al.*, (2003) has proven an efficient strategy to improve HPV production in *P. pastoris*. Attention will need to be given to the following essential factors: Use of *P. pastoris* preferred codons and elimination of AT rich regions on the HPV16 L1 gene.

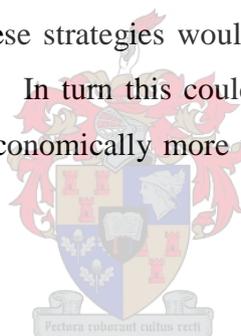
### 7.2 Process modifications

Methods for control in *Pichia* cultivations abound, but the process remains a simple three phase fed-batch cultivation. The generation of biomass is facilitated in the glycerol batch and fed-batch phase and it is only in the methanol induction phase that different induction control strategies can influence production of heterologous

proteins. Although dissolved oxygen control strategies are sufficient in most cases to facilitate production, it has some drawbacks as discussed in the literature survey. To overcome these, I recommend the implementation of a **methanol sensor** for online analysis of the methanol concentration of the induced culture. In support of this I cite the fact the *AOX* promoter is more strongly induced in the methanol limited culture. This methanol limitation can be better maintained with a direct methanol sensor.

Throughout the literature study it was rare to encounter bioreactor cultivations under 5L. **Scale up** of cultivations to 5 or 10L cultures would serve to enhance production and reduce the problems encountered in 1L bioreactors such as foaming and overflow. The available 10L bioreactors would however need to be adapted as currently the cooling loops in these bioreactors are not efficient at dissipating the heat generated by the high biomass density.

Hopefully implementation of these strategies would serve to enhance production of HPV16 L1 VLPs in *P. pastoris*. In turn this could give rise to an alternative HPV VLP vaccine candidate that is economically more viable than the current insect cell and *S. cerevisiae* systems.



## 8. References

- Anthony, C. (1982). The biochemistry of methylotrophs. Academic Press, New York, NY.
- Ausubel, F.M., Brent, R., Kingston, R., Moore, D.D., Seidman, J.G., Smith, J.A. and K. Struhl. (1995). Current Protocols in Molecular Biology. John Wiley and Sons, New York.
- Baay, M.F.D. Kjetland, E.F. Ndhlovu, P.D. Deschoolmeester, V. Mduluzi, T. Gomo, E. Friss, H. Midzi, N. Gwanzura, L. Mason, P.R. Vermorken, J.B. Gundersen, S.G. (2004). Human papillomavirus in rural community in Zimbabwe: The impact of HIV co-infection on HPV Genotype distribution. J. Med.Virol. 73: 481-485.
- Baines,W. (1998). Biotechnology from A to Z. (2<sup>nd</sup> Edition). New York. Oxford University Press.
- Baud D, Ponci F, Bobst M, De Grandi P, Nardelli-Haefliger D. (2004). Improved efficiency of a Salmonella-based vaccine against human papillomavirus type 16 virus-like particles achieved by using a codon-optimized version of L1. J. Virol. 78(23): 12901-12909
- Bosch, F.X. Manos, M.M. Munoz, N. (1995). Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. J Natl Cancer Inst. 87: 796-802.
- Bradford, M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248-254.
- Brady, C.P., Shimp, R.L., Miles, A.P., Whitmore, M., Stowers, A.W. (2001). High-level production and purification of P30P2MSP119 an important vaccine antigen for malaria, expressed in the methylotrophic yeast *Pichia pastoris*. Protein Expr. Purif. 23: 468-475.
- Cereghino, G.P., Cregg, J.M. (1999). Applications of yeast in biotechnology: protein production and genetic analysis. Curr. Opin. Biotechnol.10:422-427.
- Cereghino, J. Cregg, J.M. (2000). Heterologous protein expression in the methylotropic yeast *Pichia pastoris*. FEMS Microbiol.Rev. 24:45-66.
- Cereghino, G.P. Lin Cereghino, J. Ilgen, C. Cregg, J.M. (2002). Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. Curr. Opin. Biotechnol. 13: 329-332.
- Chauhan, A.K., Arora, D., Khanna, N. (1999). A novel feeding strategy for enhanced protein production by fed-batch fermentation in recombinant *Pichia pastoris*. Process Biochem. 34:139-145.

Chen, X.S., Garcea, R.L., Goldberg, I., Casini, G., Harrison, S.H. (2000). Structure of small virus-like particle assemble from the L1 protein of Human papillomavirus16. *Mol Cell*. 5:557-567.

Chiruvolu, V., Cregg, J.M., Meagher, M.M. (1997). Recombinant protein production in alcohol oxidase-defective strain of *Pichia pastoris* in fed-batch fermentations. *Enzyme Microb. Technol.* 21: 277-283.

Cid-Arregui, A. et al. (2003) A synthetic E7 gene of human papillomavirus type 16 that yields enhanced expression of the protein in mammalian cells and is useful for DNA immunization studies. *J. Virol.* 77:4928–4937.

Clare, J.J., Romanos, M.A., Rayment, F., Rowedder, J.E., Smith, M.A., Payne, M.M., Sreekrishna, K., Henwood, C.A. (1991a) Production of mouse epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strains containing multiple gene copies. *Gene*. 105:205-212.

Clare, J.J., Rayment, F.B., Ballantine, S.P., Sreekrishna, K., Romanos, M.A. (1991b). High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. *Bio/Technology*. 9:455-460.

Cos, O., Serrano, A., Montesinos, J.L., Ferrer, P., Cregg, J.M., Valero, F. (2005) Combined effect of the methanol utilization (Mut) phenotype and gene dosage on recombinant protein production in *Pichia pastoris* fed-batch cultures. *J.Biotechnol.*116: 321–335.

Cregg, J.M., Madden, K.N. (1987a). Developmetn of transformation systems and construction of methanol-utilisation-defective mutants of *Pichia pastoris* by gene disruption. In *Biological research on industrial yeast*. Edited by Stewart, G.G., Russel, I., Klein, R.D., Hiebsch, R.R. Boca Raton: CRC Press. 2:1<sup>18</sup>

Cregg, J.M., Tschopp, J.F., Stillman, C., Siegel, R., Akong, M., Craig, W.S., Buckholz, R.G., Madden, K.R., Kellaris, P.A., Davis, G.R., Smiley, B.L., Cruze, J., Torregrossa, R., Velicelebi, G., Thill, G.P. (1987b). High-level expression anf efficient assembly of Hepatits B surface antigen in the methylotrophic yeast, *Pichia pastoris*. *Bio/Technology*. 5:479-485.

Cregg, J.M., Vedvick, T.S., Raschke, W.C. (1993). Recent advances in the expression of foreign genes in *Pichia pastoris*. *Bio/Technology*. 11: 905-910.

DemL, L., Speth, C., Dierich, M.P., Wolf, H., Wagner, R. (2005). Recombinant HIV<sup>-1</sup> Pr55<sup>gag</sup> virus-like particles: potent stimulators of innate and acquired immune responses. *Mol. Immunol.* 42: 259-277.

De Villiers, E.M., Fauquet, C., Broker, T.R., Bernard, H.U., zur Hausen, H. (2004). Classification of papilloma virusses. *Virology*. 324:17-27.

Disbrow, G.L. et al. (2003) Codon optimization of the HPV<sup>-16</sup> E5 gene enhances protein expression. *Virology*. 311:105–114.

Finnen, R.L., Erickson, K.D., Chen, S.X., Garcea, R.L. (2003) Interactions between Papillomavirus L1 and L2 Capsid Proteins. *J.Virol.* 77(8):4818–4826.

Garcea, R.L. Gissmann, L. (2004). Virus-like particles as vaccines and vessels for the delivery of small molecules. *Curr Opin Biotechnol.* 15:513-517.

Gellissen, G., Hollenberg, C.P. (1997). Application of yeasts in gene expression studies: a comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* – a review. *Gene.* 190: 87-97.

Giroglou, T., Sapp, M., Lane, C., Fligge, C., Christensen, N.D., Streeck, R.E., Rose, R.C. (2001). Immunological analyses of human papillomavirus capsids. *Vaccine.* 19: 1783-1793.

Goldie, S.J. (2002). Health economics and cervical cancer prevention: A global perspective. *Virus Research.* 89:301-309.

Greenstone, H.L., Nieland, J.D., De Visser, K., De Bruijn, M.L.H., Kirnbauer, R., Roden, R.B.S., Lowy, D.R., Kast, W.M., Schiller, J.T. (1998). Chimeric papillomavirus virus-like particles elicit antitumor immunity against the E7 oncoprotein in an HPV16 tumor model. *Proc.Natl.Sci.USA.* 95:1800-1805.

Grgacic, E.V.L, Anderson, D.A. (2006).Virus-like particles: Passport to immune recognition. *Methods.* 40:60–65.

Gustafsson, C., Govindarajan, S., Minshull., J. (2004). Codon bias and heterologous protein expression. *Trends Biotechnol.* 22:346-353.

Hagensee, M.E. Yaegashi, N. Galloway, D.A. (1993). Self-assembly of human papillomavirus type 16 capsids by expression of the L1 protein alone or by coexpression of the L1 and L2 capsid proteins. *J.Virol.* 67: 315-322.

Hasslacher, M., Schall, M., Hayn, M., Bona, R., Rumbold, K., Luckl, J., Griengl, H., Kohlwein, S.D., Schwab, H. (1997). High-level intracellular expression of hydroxynitrile lyase from the tropical rubber tree *Hevea brasiliensis* in microbial hosts. *Protein Expr. Purif.* 11:61-71.

Heino, P., Dillner, J., Schwartz, S. (1995). Human Papillomavirus Type 16 Capsid Proteins Produced from Recombinant SemLiki Forest Virus Assemble into Virus-like Particles. *Virology.* 214:349–359.

Higgins, D.R., Cregg, J.M. (1998). Introduction to *Pichia pastoris*. In: Methods in molecular biology. Edited by: D.R. Higgins and J.M. Cregg. *Pichia Protocols* .Vol. 103.

Hohenblum, H., Gasser, B., Maurer, M., Borth, N., Mattanovich, D. (2004). Effects of gene dosage, promoters and substrates on unfolded protein stress of recombinant *Pichia pastoris*. *Biotechnol. Bioeng.* 85(4):367-375.

Hu, S., Li, L., Qiao, J., Guo, Y., Cheng, L., Liu, J. (2006). Codon optimisation, expression, and characterisation of an internalizing anti-ErbB2 single-chain antibody in *Pichia pastoris*. *Protein Expr. Purif.* 47: 249-257.

Huang, K., Lin, S. (2000) *Vaccine*. 18. Suppl1: S35-S38.

Houard, S., Heinderyckx, M., Bollen, A. (2002). Engineering of non-conventional yeasts for efficient synthesis of macromolecules: the methylotrophic genera. *Biochimie*. 84:1089–1093.

International Agency for Research on Cancer (IARC). Human papillomaviruses. Monographs on the evaluation of carcinogenic risks to humans. Lyon, IARC/WHO 64, (1995).

International Agency for Research on Cancer (IARC). The current status of development of prophylactic vaccines against human papillomavirus infection. Report of a technical meeting, Geneva. (1999).

International Aids Vaccine initiative (IAVI) Report. (2005). 9(4):23

Inan, M., Meagher, M.M. (2001). The effect of ethanol and acetate on protein expression in *Pichia pastoris*. *J. Biosci. Bioeng.* 92 (4): 337-341.

Inan, M., Meagher, M.M.. (2001). Non-repressing carbon sources for alcohol oxidase (*AOX1*) promoter of *Pichia pastoris*. *J. Biosci. Bioeng.* 92 (6): 585-589.

Invitrogen Corporation *Pichia* expression kit, "Protein Expression: A mAnnual of Methods for Expression of Recombinant Proteins using pPICZ and pPICZa in *Pichia pastoris*." Version G, Invitrogen, Carlsbad, CA USA.

Invitrogen Corporation. (2002) [www.invitrogen.com](http://www.invitrogen.com) *Pichia* Fermentation Process Guidelines. Version B.

James, E.R. (2005). Optimisation of a recombinant Hepatitis B vaccine through the cultivation and fermentation of *Aspergillus Niger*. Masters Thesis.

Kay, P., Soeters, R., Nevin, J., Denny, L., Dehaeck, C.M.C., Williamson, A. (2003). High Prevalence of HPV16 in South African Women with cancer of the cervix and Cervical Intraepithelial Neoplasia. *J. Med. Virol.* 71: 265-273.

Jahic, M., Gustavsson, M., Jansen, A., Martinelle, M., Enfors, S. (2003a). Analysis and control of proteolysis of a fusion protein in *Pichia pastoris* fed-batch processes. *J. Biotechnol.* 102:45-53.

Jahic, M., Wallberg, F., Bollok, M., Garcia, P., Enfors, S. (2003b). Temperature limited fed-batch technique for control of proteolysis in *Pichia pastoris* bioreactor cultures. *Microb. Cell Fact.* 2(1): 6.

Kaufmann, A.M., Schneider, A. (2006). New paradigm for prevention of cervical cancer. *Eur J Obstet Gynecol Reprod Biol.*

Kirnbauer, R., Booy, F., Cheng, N., Lowy, D.R., Schiller, J.T. (1992). Papillomavirus L1 major capsid protein self-assembles in to virus-like particles that are highly immunogenic. Proc. Natl. Acad. Sci. U S A. 89:12180-12184.

Kirnbauer, R., Taub, J., Greenstone, H. (1993). Efficient self-assembly of human papillomavirus type 16 L1 and L1-L2 into virus-like particles. J. Virol. 67: 6929-6939.

Kirnbauer, R., Hubbert, N.L., Wheeler, C.M., Becker, T.M., Lowy, D.R., Schiller, J.T. (1994). A virus-like particle enzyme-linked immunosorbent assay detects serum antibodies in a majority of women infected with human papillomavirus type 16. J. Natl Cancer Inst. 86: 494 - 499.

Kobayashi, K., Kuwae, S., Ohya, T., Ohda, T., Ohyama, M., Ohi, H., Tomomitsu, K., Ohmura, T. (2000). High-level expression of recombinant human serum albumin from the methylotrophic yeast *Pichia pastoris* with minimal protease production and activation. J. Biosci. Bioeng. 89 (1): 55-61.

Kolls, A., Sherris, J. (2000). HPV Vaccines: Promise and Challenges. PATH.

Leder, C. et al. (2001). Enhancement of capsid gene expression: preparing the human papillomavirus type 16 major structural gene L1 for DNA vaccination purposes. J. Virol. 75:9201-9209.

Lee, C.Y., Nakano, A., Shiomi, N., Lee, E.K., Katoh, S. (2003). Effects of substrate feed rates on heterologous protein expression by *Pichia pastoris* in DO-stat fed-batch fermentation. Enzyme Microb. Technol. 33:358-365.

Levi, J.E., Fernandes, S., Tateno, A.F., Motta, E., Lima, L.P., Eluf-Neto, J., Pannuti, C.S. (2004). Presence of multiple human papilloma types in cervical samples from HIV infected woman. Gynecol. Oncol. 92:225-231.

Li, P.C., Zhang, X.G., Zhou, L., Zeng, Y. (2003). Gene optimization is necessary to express HPV type 6 L1 protein in methylotrophic yeast *Pichia pastoris*. Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi. 17(4):310-314

Liu, H.L., Li, W.S., Lei, T., Zheng, J., Zhang, Z., Yan, X.F., Wang, Z.Z., Wang, Y.L., Si, L.S. (2005). Expression of Human papillomavirus type 16 L1 protein in transgenic tobacco plants. Acta Biochim Biophys Sin (Shanghai). 37(3):153-158.

Liu, S., Chou, W., Sheu, C., Changa, M.D. (2005). Improved secretory production of glucoamylase in *Pichia pastoris* by combination of genetic manipulations. Biochem. Biophys. Research Comm. 326 817-824.

Lowy, D.R., Kirnbauer, R., Schiller, J.T. (1994). Genital Human Papillomavirus Infection. PNAS. 91: 2436 - 2440.

Lowy, D.R., Schiller, J.T. (1998). Papilloma virus and cervical cancer: pathogenesis and vaccine development. J Natl Cancer Inst Monogr. 23: 27-30.

Maclean, J., Rybicki, E.P., Williamson, A.L. (2005). Vaccination strategies for the prevention of cervical cancer. *Expert Rev Anticancer Ther.* 5(1): 97-107.

Minning, S. Serrano, A., Ferrer, P., Solá, C., Schmid, R.D., Valero, F. (2001). Optimisation of high-level production of *Rhizopus oryzae* lipase in *Pichia pastoris*. *J. Biotechnol.* 86:59-70.

Munoz, N., Bosch, F.X., de Sanjose, S., Herrero, R., Castellsague, X., Shah, K.V., et al. (2003). Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N. Engl. J. Med.* 348(6): 518-27.

Neeper, M.P. Hofmann, K.J. Jansen, K.U. (1996). Expression of the major capsid protein of human papilloma virus type 11 in *Saccharomyces cerevisiae*. *Gene.* 180:1-6.

Outchkourov, N.S. et al. (2002). Optimization of the expression of equistatin in *Pichia pastoris*. *Protein Expr. Purif.* 24:18-24.

Parkin, D.M. (2002). Global cancer statistics in the year 2001. *Lancet Oncol.* 2(9), 533-543.

Pattenden, L.K., Middelberg, A.P.J., Niebert, M., Lipin, D.I. (2005). Towards the preparative and large-scale precision manufacture of virus-like particles. *Trends Biotechnol.* 23: 523-529.

Roden, R.B., Armstrong, A., Haderer, P., Christensen, N.D., Hubbert, N.L., Lowy, D.R., Schiller, J.T., Kirnbauer, R. (1997). Characterisation of a human papillomavirus type 16 variant-dependent neutralizing epitope. *J. Virol.* 71:6247-6252.

Romanos, M., Scorer, C.A., Clare, J.J. (1992). Foreign gene expression in yeast: a review. *Yeast.* 8:423-488.

Romanos, M. (1995). Advances in the use of *Pichia pastoris* for high-level gene expression. *Curr. Op. Biotechnol.* 6:527-533.

Sahm, H. (1977). Metabolism of methanol by yeasts. In: Ghose, T.K., Frechter, A., Blankenbrough, H. (Eds.) *Advances in biochemical engineering.* Vol 6, Springer Verlag, Berlin. p 77-103.

Santi, L., Huang, Z., Mason, H. (2006). Virus-like particles production in green plants. *Methods.* 40:66-76.

Sasagawa, T., Pushko, P., Steers, G., Gschmeissner, S.E., Hajibagheri, M.A.N., Finch, J., Crawford, L., Tommasino, M. (1995). Synthesis and Assembly of Virus-like Particles of Human Papillomaviruses Type 6 and Type 16 in Fission Yeast *Schizosaccharomyces pombe*. *Virology.* 206:126-135.

Sasagawa, T., Tani, M., Basha, T., Gigs-Hama, Y., Azar, K.K., Yasuda, H., Sakai, A., Inoue, M. (2005). A human papillomavirus type 16 vaccine by oral delivery of L1 protein. *Virus Research.* 110:81-90.

- Schiller, J.T., Lowy, D.R. (2001). *Expert. Opin. Biol. Ther.* 1: 571-581.
- Shi, X., Karkut, T., Chamankhah, M., Alting-Mees, M., Hemmingsen, S.M., Hegedus, D. (2003). Optimal conditions for the expression of a single-chain antibody (scFv) gene in *Pichia pastoris*. *Protein Expr. Purif.* 28:321-330.
- Sinclair, G. and Choy, F.Y. (2002). Synonymous codon usage bias and the expression of human glucocerebrosidase in the methylotrophic yeast, *Pichia pastoris*. *Protein Expr. Purif.* 26: 96–105.
- Sreekrishna, K., Brankamp, R.G., Kropp, K.E., Blankenship, D.T., Tsay, J., Simth, P.L., Wierschke, J.D., Subramaniam, A., Birkenberger, L.A. (1997). Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. *Gene.* 190:55-62.
- Stanley, M., Lowy, D.R., Frazer, I. (2006). Chapter 12: Prophylactic vaccines: Underlying mechanisms. *Vaccine.* 24S3: S3/106-S3/113.
- Steller, M.A. (2003). Human papillomavirus, it's genes...and cancer vaccines. *Cancer cell.* Jan: 7-8.
- Stratton, J., Chirovulo, V., Meagher, M. (1998). High cell density fermentation. *Methods. Mol. Biol.* 103:107-120.
- Studentsov, Y.Y., Schiffman, M., Strickler, H.D., Ho, G.Y.F., Pang, Y.S., Schiller, J., Herrero, R., Burk, R.D. (2000). Enhanced enzyme-linked immunosorbent assay for detection of antibodies to virus-like particles of Human Papillomavirus. *J. Clin. Microbiol.* 40(5):1755–1760.
- Valenzuela, P., Medina, A., Rutter, W.J., Ammerer, G., Hall, B.D. (1982). Synthesis and assembly of hepatitis B surface antigen particles in yeast. *Nature.* 298: 347-350.
- Varsani, A., Williamson, A.L., Rose, R.C., Jaffer, M., Rybicki, E.P. (2003). Expression of Human papillomavirus type 16 major capsid protein in transgenic *Nicotiana tabacum* cv Xanthi. *Arch. Virol.* 148(9): 1771-1786.
- Vassileva, A., Chugh D.A., Swaminathan, S., Khanna, N. (2001). Expression of hepatitis B surface antigen in the methylotrophic yeast *Pichia pastoris* using the GAP Promoter. *J. Biotechnol.* 88:21–35.
- Vassileva, A., Chugh D.A., Swaminathan, S., Khanna, N. (2001). Effect of copy number on the expression levels of Hepatitis B surface antigen in the methylotrophic yeast *Pichia pastoris*. *Protein Expr. Purif.* 21:71-80.
- Wallboomers, J.M., Jacobs, J.M., Manos, M.M., Bosch, F.X., Kummer, J.A., Shah, K.V., Snijders, P.J., Peto, J., Meijer, C.J., Munoz, N. (1999). Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J. Pathol.* 189:12-19.
- Wegner, G.H. (1983). Biochemical conversions by yeast fermentation at high cell densities. US Patent 4: 329, 414.

Merck. (2006). FDA Approves Merck's GARDASIL.  
[http://www.merck.com/newsroom/press\\_releases/product/2006\\_0608.html](http://www.merck.com/newsroom/press_releases/product/2006_0608.html)

Villa, L.L, Costa, L.R.R., Petta, C.A., Andrade, R.P., Ault, K.A., Giuliano, R.A., Wheeler, C.M., Koutsky, L.A., Malm, C., Lehtinen, M. *et al.*, (2005). Prophylactic quadrivalent human papillomavirus (types 6, 11, 16, and 18) L1 virus-like particle vaccine in young women: a randomised double-blind placebo-controlled multicentre phase II efficacy trial. *Lancet. Oncol.* 6(5): 271-278.

WHO,1. The current status of development of prophylactic vaccines against human papillomavirus infection. Report of a technical meeting, Geneva, 16<sup>th</sup> February 1999  
<http://www.who.int/gpv-documents/>

WHO,2. Proceedings of the fourth Global Vaccine Research Forum  
<http://www.who.int/vaccine-documents/>

World Health Organisation. (2005) Available online at: [http://www.who.int/vaccine\\_research/documents/new\\_vaccines/en/index8.html](http://www.who.int/vaccine_research/documents/new_vaccines/en/index8.html).

Woo, J.H. et al. (2002) Gene optimization is necessary to express a bivalent anti-human anti-T cell immunotoxin in *Pichia pastoris*. *Protein Expr. Purif.* 25:270–282.

Xie, J., Zhou, Q., Du, P., Gan, R., Ye, Q. (2005). Use of different carbon sources in cultivation of recombinant *Pichia pastoris* for angiostan production. *Enzyme Microb. Technol.* 36: 210-216.

Yadava, A. and Ockenhouse, C.F. (2003) Effect of codon optimization on expression levels of a functionally folded malaria vaccine candidate in prokaryotic and eukaryotic expression systems. *Infect. Immun.* 71:4961–4969.

Zhang, W., Bevins, M.A., Plantz, B.A., Smith, L.A., Meagher, M.M. (2000). Modeling *Pichia pastoris* growth on methanol and optimising the production of a recombinant protein, the heavy-chain fragment C of Botulinum neurotoxin, Serotype A. *Biotechnol. Bioeng.* 70(1):1-8.

Zhou, J. Sun, X.Y. Stanzel, D.J. Frazer, I.H. (1991). Expression of vaccinia recombinant HPV16 L1 and L2 ORF proteins in epithelial cells is sufficient for assembly of HPV virion-like particles. *Virology.* 185: 251-257.

## 9. Appendix

### 9.1 ANOVA

A single-factor ANOVA was carried out to determine the significance of the experimental data. The critical variables of different experiments were individually compared to each other to determine if a significant difference was apparent. When the p-value was larger than the  $\alpha$ -value (largest probability of rejecting  $H_0$ ) the null hypothesis ( $H_0$ ), which stated that there was a significant difference, was rejected. A 95% confidence interval ( $\alpha = 0.05$ ) was established for the results. The  $\pm$  indicated in Table 4.1 and Table 4.2 and the error bars on the charts signifies one standard deviation from the mean.

**TABLE 9.1.1:** ANOVA Statistical analysis of **TABLE 4.1: Comparison of strains and gene dosage effects.**

Anova analysis for different strains and multiple gene dosage

Measured variable	P-value	Significance
Specific VLP production [ $\mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}$ ] X33[Syn-L1] and X33[h-L1]	0.0006	Significant difference
Specific monomer production [ $\mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}$ ] X33[Syn-L1] and X33[h-L1]	0.169	No significant difference
HPV VLP per total protein [ $\text{ng}\cdot\text{mg}^{-1}$ ]	1.56E-10	Significant difference
VLP Yield on methanol [ $\mu\text{g}\cdot\text{g}^{-1}_{\text{DCW}}\cdot\text{L}^{-1}_{\text{methanol}}$ ] X33[Syn-L1] and X33[h-L1]	0.0007	Significant difference
Average Methanol consumption rate [ $\text{ml}\cdot\text{h}^{-1}$ ] X33[Syn-L1] and X33[h-L1]	0.098	No significant difference
Biomass [ $\text{g}\cdot\text{L}^{-1}$ ]	0.00013	Significant difference
Average linear biomass increase [ $\text{g}_{\text{CDW}}\cdot\text{L}^{-1}\cdot\text{h}^{-1}$ ]	0.54	No significant difference
VLP : monomer X33[Syn-L1] and X33[h-L1]	7.2E-5	Significant difference

**TABLE 9.1.2: ANOVA Statistical analysis of TABLE 4.2: Comparison of different control strategies**

Anova analysis for different control strategies

Measured variable	P-value	Significance
Specific VLP production [ $\mu\text{g}\cdot\text{g}^{-1}_{\text{dcw}}$ ]	0.509	No significant difference
Specific monomer production [ $\mu\text{g}\cdot\text{g}^{-1}_{\text{dcw}}$ ]		
CMF & 40% DO	9.47E-4	Significant difference
40% DO & 20% DO	2.91E-5	
HPV VLP per total protein [ $\text{ng}\cdot\text{mg}^{-1}$ ]	8.52E-06	Significant difference
VLP Yield on methanol [ $\mu\text{g}\cdot\text{g}^{-1}_{\text{dcw}}\cdot\text{L}^{-1}_{\text{methanol}}$ ]	0.103	No significant difference
Average Methanol consumption rate [ $\text{ml}\cdot\text{h}^{-1}$ ] (20%DO and 40% DO)	0.03	Significant difference
Biomass [ $\text{g}\cdot\text{L}^{-1}$ ]	8.92E-08	Significant difference
Average linear biomass increase [ $\text{g}_{\text{cdw}}\cdot\text{L}^{-1}\cdot\text{h}^{-1}$ ]	1.43E-03	Significant difference
VLP : monomer (CMF and 40% DO)	0.025	Significant difference

**TABLE 9.1.3:** ANOVA Statistical analysis of **FIGURE 4.3: Volumetric production levels**

Anova analysis for volumetric production levels

Measured variable	P-value	Significance
CMF and 40% DO X-33 h-L1	0.004	Significant difference
40% DO X-33 h-L1 and multi-h-L1	0.038	Significant difference
40% and 20% DO h-L1	0.044	Significant difference
40% DO X-33 h-L1 and Syn-L1	0.002	Significant difference

**TABLE 9.1.4:** ANOVA Statistical analysis of **FIGURE 4.4: Location of VLPs.**

Cytoplasmic fraction		
Measured variable	P-value	Significance
DO control (20% & 40%)	0.182	No significant difference
CMF and 40% DO control	0.003	Significant difference
DO control 40% (X33[h-L1] & X33[multi-hL1])	0.136	No significant difference
Strains X33[h-L1] & X33[Syn-L1]	0.001	Significant difference
Membrane fraction		
Measured variable	P-value	Significance
DO control (20% & 40%)	0.44328	No significant difference
CMF and 40% DO control	2.36E-02	Significant difference
DO control 40% (X33[h-L1] & X33[multi-hL1])	1	No significant difference
Strains X33[h-L1] & X33[Syn-L1]	2.17E-03	Significant difference