# Cloning and evaluation of expression of the open reading frames of a South African G9P[6] rotavirus strain encoding rotavirus structural proteins VP2 and VP6 in bacteria and yeast

**LA Naudé**

**20577095**

Dissertation submitted in fulfilment of the requirements for the degree *Magister Scientiae* in Biochemistry at the Potchefstroom Campus of the North-West University

Supervisor: Prof AA van Dijk

Co-supervisor: Dr HG O'Neill

November 2014

*“What lies behind you and what lies in front of you, pales in comparison to what lies inside of you”*

**Ralph Waldo Emerson**

# Acknowledgements:

I would like to show my appreciation and gratitude to the following people and institutions, for their contributions to the project as without them this project would not have been possible:

- Prof Albie van Dijk (North-West University) thank you so much for all the encouragement, support and guidance.

- Dr Trudi O'Neill (University of the Free State) thank you so much for all your patience, guidance, time and support throughout this study.

- The National Research Foundation and North-West University for financial support

- Dr Jacobus Albertyn (University of the Free State) for providing his knowledge and guidance of yeast expression

- Dr Christiaan Potgieter (Deltamune) for providing the pColdTF_VP2 construct

- Julian Ingram for all your continuous love, support, and encouragement throughout this project

- The support and encouragement of my family and friends especially my mother and father

- Lastly my Heavenly Father for giving me this great opportunity and strength to complete this project

# Summary:

Rotavirus infection causes severe gastroenteritis, affecting all children under the age of five regardless of hygiene or water quality. The currently licensed vaccines succeeded in reducing diarrhoea worldwide, but they still have shortcomings, especially the efficacy of the vaccines in developing countries. One of the main reasons for this can be due to the difference in strains, since the strains used to develop the currently licensed vaccines (RotaTeq and Rotarix) were selected from strains circulating in the developed world (G1, G2, G3 and G4), while the main strains present in Africa (G8, G9 and G12) were not included. A second shortcoming of the currently licensed vaccines is the cost of these vaccines. The vaccines are very expensive and most developing countries cannot afford the vaccines as well as the fact that the manufacturing companies cannot produce enough vaccines for all the countries. An attractive alternative to the currently licensed rotavirus vaccines is the non-live vaccine candidate, virus-like particles, which can provide a possible cheaper, safer and efficacious alternative or complement the currently licensed vaccines.

Therefore, in this study a South African G9P[6] rotavirus strain, RVA/Human-wt/ZAF/GR10924/1999/G9P[6], was used to determine whether or not co-expression of the structural proteins VP2 (genome segment 2) and VP6 (genome segment 6) was possible in bacteria and yeast. The South African GR10924 G9P[6] neonatal strain was previously obtained from a stool sample and the nucleotide consensus sequence was determined for both genome segment 2 (VP2) and genome segment 6 (VP6). Bacterial codon optimised coding regions or open reading frames were used in this study. The open reading frames (ORFs) of the genome segments encoding, VP2 and VP6, were cloned into the expression vector pETDuet-1, which allows for the simultaneous expression of two genes in bacteria. The ORF of genome segment 6 was purchased from GeneScript and the ORF of genome segment 2 was obtained from Dr AC Potgieter (Deltamune (Pty) Ltd R&D, South Africa). Compatible restriction enzyme sites were used to sub-clone the ORF of the bacterial codon optimised genome segments into the expression vector. Only the expression of the VP6 protein in bacteria was observed with Coomassie stained SDS-PAGE.

The ORFs encoding VP2 (genome segment 2) and VP6 (genome segment 6) of the wild type GR10924 G9P[6] strain were cloned into the wide range yeast expression system vector, pKM173, which allows for the simultaneous expression of more than one gene. Several yeast strains were used in this study namely *Kluyveromyces marxianus, Kluyveromyces lactis, Candida deformans, Saccharomyces cerevisiae, Yarrowia lipolytica, Arxula adeninivorans, Hansenula polymorpha* and *Debaryomyces hansenii*. Expression of both proteins was not detected in the several yeast strains, as seen with western blot

analysis. DNA extractions were done on two colonies of each yeast strain that were used for western blot analysis to evaluate successful integration into the yeast genomes. Only a few of the colonies contained either both of the genome segments or only one of the two genome segments of interest.

To summarise, the simultaneous expression of VP2 and VP6 from rotavirus GR10924 G9P[6] was not successful in bacteria or yeast, but it was possible to soluble express the bacterial codon optimised GR10924 G9P[6] VP6 in bacteria using the pETDuet-1 as expression vector.

**Keywords:** Rotavirus, gastroenteritis, non-live vaccine, virus-like particles, South Africa, GR10924 G9P[6], bacterial codon optimised, bacterial expression, yeast expression.

# Opsomming:

Rotavirus infeksie veroorsaak ernstige gastro-enteritis wat alle kinders jonger as vyf jaar affekteer, ongeag van higiëne en water gehalte. Die huidige gelisensieerde entstowwe is effektief in kliniese proewe wêreldwyd maar het nog steeds tekortkominge, soos die effektiwiteit van die entstowwe in ontwikkelende lande. Een van die hoofredes vir die lae effektiwiteit van die entstowwe kan wees as gevolg van die rotavirus stamme wat gebruik word om die entstowwe te vervaardig, aangesien die entstowwe ontwikkel is vanaf rotavirus stamme wat in ontwikkelde lande sirkuleer (G1, G2, G3 en G4) terwyl opkommende stamme in Afrika (G8, G9 en G12) nie ingesluit word nie. Die ander moontlike tekortkominge van die huidige gelisensieerde entstowwe is die koste van die entstowwe, aangesien dit baie duur is en die meeste ontwikkelende lande nie die entstowwe kan bekostig nie. Die vervaardigingsmaatskappye kan ook nie genoeg entstowwe produseer vir al die lande wat die entstowwe gebruik nie. Virusagtige partikels (VAPs) is n aantreklike nie-lewendige entstof kandidaat wat moontlik ´n goedkoper, veiliger en meer effektiewe alternatief of kompliment tot die huidige gelisensieerde entstowwe kan wees.

In die hudige studie is die Suid-Afrikaanse GR10924 G9P[6] rotavirus stam, RVA/Human-wt/ZAF/GR10924/1999/G9P[6] gebruik om te bepaal of die gelyktydige uitdrukking van die strukturele proteïen VP2 (genoom segment 2) en VP6 (genoom segment 6) in bakterieë en gis moontlik is. Die Suid-Afrikaanse GR10924 G9P[6] neonatale stam is voor die aanvang van die studie uit n stoel monster verkry en die nukleotiedvolgorde bepaal vir beide genoom segment 2 (VP2) en genoom segment 6 (VP6). Bakteriële kodon-geoptimaliseerde DNS volgordes is in die studie gebruik. Die oopleesraam van die genoom segmente wat vir VP2 en VP6 kodeer is in die uitdrukkingsvektor pETDuet-1 gekloneer. Die uitdrukkingsvektor laat die gelyktydige uitdrukking van twee gene toe. Die oopleesraam van genoom segment 6 is vanaf GeneScript verkry en die oopleesraam van genoom segment 2 vanaf Dr AC Potgieter (Deltamune (Pty) Ltd R&D, Suid-Afrika). Geskikte beperkingsensiem verterings is gebruik om die oopleesraam van die bakteriële kodon-geoptimiseerde genoom segmente in die uitdrukkingsvektor te sub-kloneer. Slegs die uitdrukking van die VP6 proteïen was suksesvol in bakterieë, soos op n Coomassie gekleurde SDS-PAGE jel waargeneem.

Die oopleesraam wat kodeer vir VP2 (genoom segment 2) en VP6 (genoom segment 6), van die wilde tipe GR10924 G9P[6] rotavirus stam, is in die gis uitdrukkingsvektor pKM173 gekloneer. Die pKM173 vektor laat die gelyktydige uitdrukking van meer as een geen toe. Verskeie gis stamme is in die studie gebruik naamlik *Kluyveromyces marxianus, Kluyveromyces lactis, Candida deformans, Saccharomyces cerevisiae, Yarrowia lipolytica,*

*Arxula adeninivorans, Hansenula polymorpha* en *Debaryomyces hansenii.* Uitdrukking van beide proteïne soos waargeneem met westelike klad analise was nie in enige van hierdie gis stamme suksesvol nie. DNS is ge-ekstraeer op twee kolonies van elke gis stam wat gebruik was vir die westelike klad analise eksperiment. Ons het bevind dat slegs ‘n paar kolonies beide rotavirus gene bevat terwyl die ander kolonies slegs een van die twee gene van belang bevat het.

Om op te som; die gelyktydige uitdrukking van VP2 (genoom segment 2) en VP6 (genoom segment 6) van die GR10924 G9P[6] rotavirus stam was nie suksesvol in beide gis of bakterieë nie. Die bakteriële kodon-geoptimiseerde VP6 proteïen kon in n oplosbare vorm in bakterieë uitgedruk word.

**Sleutelwoorde:** Rotavirus, gastro-enteritis, nie-lewendige entstowwe, virusagtige partikels, Suid-Afrika, GR10924 G9P[6], bakteriële kodon geoptimaliseerde, bakteriële uitdrukking, gis uitdrukking.

# Table of Contents

# Chapter 1:

# Literature review

## 1.1 An overview of rotavirus

In 1973, Dr Ruth Bishop and her colleagues discovered a virus that causes diarrhoea using electron microscopy. They identified virus particles in the cytoplasm of mature epithelial cells lining duodenal villi and in faeces, from children admitted to the Royal Children's Hospital in Melbourne, Australia (Bishop *et al.,* 2009). The virus had a diameter of 70 nm and had a wheel like appearance, from there the name rotavirus (rota= Latin word for wheel) (Bishop *et al.,* 2009).

Rotavirus mainly infects the young of humans and animals. In animals, rotavirus infection is more prevalent in sheep, pigs and cattle and can lead to significant economic loss in livestock (Martella *et al.,* 2010 and Midgley *et al.,* 2012). Rotavirus infects almost all children worldwide in industrialized and developing countries, before the age of 5 years (Glass *et al.,* 2008). Symptoms of rotavirus usually occur in the first 48 hours after infection. Symptoms include vomiting, watery diarrhoea and abdominal pains. These symptoms cause severe diarrhoea in children and infants (Surendran *et al.,* 2008). Rotavirus infections have also been documented in adults and are mainly spread by faecal-oral transmission. Symptoms in adults include nausea, malaise, headache, abdominal cramping, fever and diarrhoea (Anderson *et al.,* 2009); the infection in adults can also be asymptomatic.

## 1.2 Rotavirus structure

The rotavirus virion consists of structural and non- structural proteins.

### 1.2.1 Structural proteins

The rotavirus virion consists of six structural proteins, VP1-4, VP6 and VP7 that form the concentric three-layered particle that surround the 11 genome segments (Figure 1.1). The internal layer surrounds the viral genome and contains VP2, the RNA-dependent RNA polymerase VP1 and VP3. VP6 forms the middle layer and the outer layer consist of VP7 and VP4 (Estes and Kapikian, 2007)

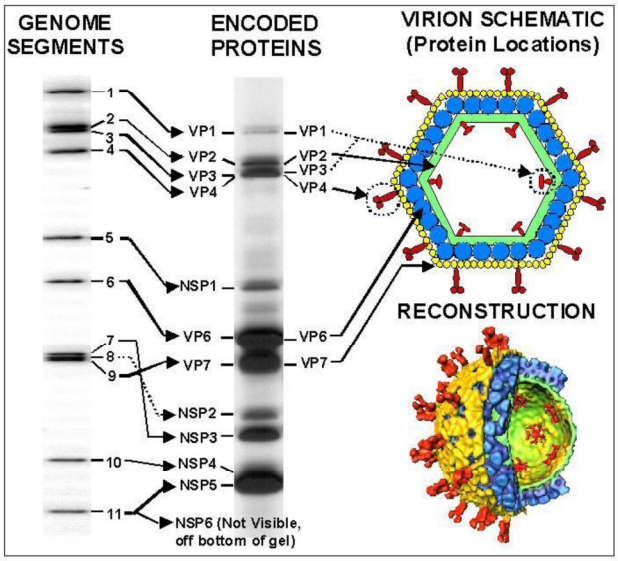


**Figure 1.1: An overview of the coding assignment and virion structure of rotavirus.** Rotavirus has 11 RNA genome segments. The genome segments encode six structural proteins (VP1-4 and VP6-7) and six non-structural proteins (NSP1-6). The six structural proteins form the concentric three-layered particle of rotavirus. Illustration taken from Greenberg and Estes, 2009, with permission.

#### 1.2.1.1 Inner layer

VP1 is encoded by genome segment 1 and is one of the three proteins that make up the inner layer of the concentric three-layered particle (Estes and Cohen, 1989). VP1, an RNA-dependent RNA polymerase is necessary for the recognition of the 3' end of the mRNA, but in the absence of other viral proteins (Table 1.1), does not have any replicase activity. It can, however, still bind to any viral mRNA (Estes and Kapikian, 2007). VP2 is encoded by genome segment 2 and is necessary for VP1 replicase activity (Table 1.1). VP2 interacts with plus strand RNA and VP2 interactions with the VP1 polymerase are required for replicase activity (Estes and Kapikian, 2007). VP3 is encoded by genome segment 3 and is a minor structural protein (Estes and Cohen, 1989). VP3 is a guanyl-transferase and methyl-transferase enzyme (Table 1.1) that is present in small quantities. Together with VP1 it provides enzymatic functions required for producing the capped mRNA transcripts (Jayarem *et al.,* 2004).

#### 1.2.1.2 Intermediate layer

VP6 is encoded by genome segment 6 and is the major structural protein in virus particles, located on the outer surface of single layered particles (Estes and Cohen, 1989). VP6 together with VP2 forms the double layered particle (DLPs). VP6 is the most abundant and immunogenic protein of the virus and contains group and subgroup antigenic determinates (Table 1.1).

Both biochemical and immunological approaches have been used to determine whether or not VP6 performs specific biological functions during virus replication. Earlier biochemical studies have indicated that none of the components of double-layered particles are capable of transcribing dsRNA and that VP6, despite the lack of any enzymatic function, is essential for endogenous transcription of the genome (Jayarem *et al.,* 2004).

#### 1.2.1.3 Outer layer

The rotavirus protein VP7 together with VP4 forms the outer capsid protein shell, with VP4 spikes that emanate through the outer capsid shell. These proteins (VP4 and VP7) induce neutralizing antibody responses and define the serotypes of the virus (Table 1.1).

Protease sensitive VP4 is encoded by genome segment 4 (Table 1.1) (Greenberg and Estes, 2009). It is also used to determine serotypes, namely P types (Dennehy *et al.,* 2008) and play a role in cell attachment and cell penetration of the protein. When VP4 is cleaved (prior to cell attachment) by a trypsin-like enzyme, it forms VP5* and VP8* that result in the enhancement of viral infectivity. The cleavage also enhances the penetration of the virus into the cells (Estes and Kapikian, 2007).

Glycoprotein VP7 is encoded by genome segment 9 (Table 1.1) (Greenberg and Estes, 2009). Serotypes determined by this protein are, therefore, termed G serotypes (Dennehy *et al.,* 2008). This capsid protein (VP7) induces neutralizing antibodies and is highly immunogenic. Appropriate calcium levels help maintain the structural integrity of the VP7 layer. However, low concentrations of calcium, similar to those in the cytoplasm, trigger the dissociation of VP7 trimers. This leads to the uncoating of the VP7 layer. The uncoating of the outer layer, which results in the formation of DLPs is a necessary event in the replication cycle of rotavirus (Jayarem *et al.,* 2004).

**Table 1.1: Properties of rotavirus genome segments, proteins and their functions.** The data in the table is based on the rotavirus SA11 strain.

| Genome segments | Size (bp) | Protein | Size (kDa) | Location | Protein function |
|---|---|---|---|---|---|
| 1 | 3302 | VP1 | 125 | Core | RNA-dependent RNA polymerase, RNA binding, interacts with VP2 and VP3 |
| 2 | 2690 | VP2 | 102 | Core | RNA binding, interacts with VP1 |
| 3 | 2591 | VP3 | 98 | Core | Guanylyl and methyl transferase, ssRNA binding, interacts with VP1 |
| 4 | 2362 | VP4 | 88 | Outer capsid | Hemagglutinin, neutralization antigen, virulence, protease-enhanced infectivity, cell attachment, fusion region |
| 5 | 1611 | NSP1 | 59 | Non-structural | RNA binding, antagonist of interferon response |
| 6 | 1356 | VP6 | 48 | Intermediate capsid | Hydrophobic trimer, group and subgroup antigen |
| 7 | 1105 | NSP2 | 35 | Non-structural | Important for genome replication/packaging, main constituent of viroplasm, NTPase, RNA binding, interacts with NSP5 |
| 8 | 1059 | NSP3 | 37 | Non-structural | Important for viral mRNA translation, PABP homologue, RNA binding, interacts with eIF4G |
| 9 | 1062 | VP7 | 37 | Outer capsid | RER transmembrane glycoprotein, neutralization antigen, $Ca^{++}$ binding |
| 10 | 751 | NSP4 | 20 | Non-structural | RER transmembrane glycoprotein, role in morphogenesis, viral enterotoxin |
| 11 | 667 | NSP5 | 22 | Non-structural | Constituent of viroplasm, interacts with NSP2, RNA binding, Protein kinase |
| 11 | | NSP6 | 11 | Non-structural | Constituent of the viroplasm, interacts with NSP5 |

*Table was compiled from Pesavento *et al.,* 2006 and Attoui *et al.,* 2011

### 1.2.2 Non-structural proteins

The rotavirus virion also contains six non-structural proteins (NSP), namely NSP1-6. Non-structural proteins are not incorporated into the mature virus particle, but are expressed in the infected cells from the viral genome (Estes and Kapikian, 2007). These proteins play an essential role in virus pathogenesis, morphogenesis and replication and most of them also show multifunctional properties. Non-structural proteins represent potential targets for the development of antiviral agents because of their essential roles in virus biology (Suguna *et al.,* 2010). The characteristics and known functions of the non-structural proteins are summarised in Table1.1.

## 1.3 Classification of Rotavirus

### *1.3.1 Reoviridae*

Rotavirus is one of fifteen genera of the *Reoviridae* family. *Reoviridae* has two subfamilies, each respectively consisting of six and nine genera (Table 1.2). *Reoviridae* is a family of linear, segmented double-stranded RNA genome segments (Patton *et al.,* 2008). The virion contains three capsids namely an outer, middle and inner capsid that has an icosahedral symmetry. They have between 10 - 12 segments and the genome size varies from 18 000-30 000 base pairs. *Reoviridae* viruses are named so because the first reovirus was isolated from the respiratory and enteric tracks of animals and humans as "orphans", *i.e.* they are not associated with any disease (Carter *et al.,* 2005). Viruses of the *Reoviridae* family infect vertebrates, invertebrates, higher plants, bacteria and fungi.

**Table 1.2: Classification of the *Reoviridae* family**

| ***Reoviridae* subfamilies** | **Genus** | **Number of genome segments** | **Genome size** |
|---|---|---|---|
| *Sedoreoviridae* | *Cardoreovirus* | 11 | N/A |
| | *Mimoreovirus* | 11 | 25 400bp |
| | *Orbivirus* | 10 | 19 200bp |
| | *Phytoreovirus* | 12 | 26 000bp |
| | *Rotavirus* | 11 | 18 500bp |
| | *Seadornavirus* | 11 | 21 000bp |
| *Spinareoviridae* | *Aquareovirus* | 11 | 30 500bp |
| | *Coltvirus* | 12 | 29 000bp |
| | *Cypovirus* | 10 | 25 000bp |
| | *Dinovernavirus* | 9 | N/A |
| | *Fijivirus* | 10 | 27 000 – 30 000bp |
| | *Idnoreovirus* | 10-11 | 27 000 – 30 000bp |
| | *Mycoreovirus* | 11-12 | 23 000bp |
| | *Orthoreovirus* | 10 | 23 000bp |
| | *Oryzavirus* | 10 | 20 000bp |

Table compiled from the online databases Universal Database for the International Committee on Taxanomy of viruses (ICTV) (http://ictvonline.org/) and viral zone database (http://viralzone.expasy.org/).

### 1.3.2 Rotavirus

Rotavirus is a segmented double-stranded RNA virus in the family *Reoviridae* and specifically the subfamily of *Sedoreoviridae.* The viral genome consists of 11 double-stranded RNA (dsRNA) segments that are contained within the virus capsid (Estes and Cohen, 1989) and has a total size of approximately 18,522 base pairs ranging from 667bp (genome segment 11) to 3,302bp (genome segment 1). Each of the genome segments encodes for one viral protein except for genome segment 11, which encodes for two proteins. The genome segment encodes for six structural proteins (VP1-4, VP6-7) and six non-structural proteins (NSP1-6). Rotavirus particles contain the dsRNA-dependent RNA polymerase that is important in producing mRNA in infected cells, which is required for gene expression and genome replication (Estes and Kapikian, 2007; Patton *et al.,* 1995).

### 1.3.3 Rotavirus groups and subgroups

Rotavirus can be classified into eight serological groups, namely A-H. Rotavirus groups A-C are found in humans and animals while group D-G have only been found in animals to date (Knipe and Estes, 2007; Matthijnssens *et al.,* 2011). The remainder of this literature review and study mainly focus on Group A rotaviruses. Group A rotaviruses are further classified into subgroups according to the presence of specific epitopes that are found on the VP6 protein (Greig *et al.,* 2006). Subgroup A can further be divided into serogroups that are determined by the protease sensitive protein, VP4 (P-serotypes) and the glycoprotein VP7 (G-serotypes) that induce neutralizing antibodies (Matthijnssens *et al.,* 2008). Due to the cross-reactivity of the monoclonal antibodies used in the serotyping, dual typing is now performed as genotyping.

The overall genetic relatedness among homologous genome segments has been assessed by RNA-RNA hybridization. RNA-RNA hybridization has provided molecular evidence to show close interspecies relationship between human and animal strains. Three human genogroups, represented by reference rotavirus strains Wa, DS-1 and Au-1 have been established. Sequence comparison of rotavirus genomes is critical to the assignment of genotypes (Matthijnssens *et al.,* 2011). In 2008, a classification system was proposed for rotaviruses in which all eleven genomic RNA segments are used. The full genome classification system is based on nucleotide cut-off percentages (Matthijnssens *et al.,* 2008). The full genome classification system is depicted on the notations Dx-P[x]-Ix-Rx-Cx-Mx-Ax-Nx-Tx-Ex-Hx, which represents, respectively, the genotype for the genome segments encoding V7-VP4-VP6-VP1-VP2-VP3-NSP1-NSP2-NSP3-NSP4-NSP5. The letter in each genotype were derived from the function associated with the structural and non-structural

protein for example **G**lycosylated, **P**rotease sensitive, **I**nner capsid protein, **R**NA-dependent RNA polymerase, **C**ore protein, **M**ethyltransferase, Interferon **A**ntagonist, **N**TPase, **T**ranslation enhancer, **E**nterotoxin, p**H**osphoprotein (Table 1.3) (Matthijnssens *et al.,* 2008). There are currently 27G, 35P, 16I, 9R, 9C, 8M, 17A, 9N, 12T, 15E and 11H genotypes identified as of September 2012 (10th International Rotavirus symposium. Bangkok 2012, Matthijnssens).

**Table 1.3: Nucleotide percentage identity cut-off values defining genotypes of 11 rotavirus genome segments (Matthijnssens *et al.,* 2008).**

| Genome product | Percentage identity cut-off values | Number of genotypes | Function of genotypes |
|---|---|---|---|
| VP7 | 80% | 27G | **G**lycosylated |
| VP4 | 80% | 35P | **P**rotease sensitive |
| VP6 | 85% | 19I | **I**nner capsid |
| VP1 | 83% | 9R | **R**NA dependant-RNA polymerase |
| VP2 | 84% | 9C | **C**ore protein |
| VP3 | 81% | 8M | **M**ethyltransferase |
| NSP1 | 79% | 17A | interferon **A**ntogonist |
| NSP2 | 85% | 9N | **N**TPase |
| NSP3 | 85% | 12T | **T**ranslation enhancer |
| NSP4 | 85% | 15E | **E**nterotoxin |
| NSP5 | 91% | 11H | p**H**osphoprotein |

### 1.4 Burden of Rotavirus disease

Gastroenteritis causes about 1.3 million deaths each year in children under the age of five, as estimated by the World Health Organisation (WHO) (Tate *et al.,* 2012). At least one third of all gastroenteritis cases are caused by rotavirus, in both the developed and developing countries. Other causes of gastroenteritis are due to bacterial infections, parasites, other virus infections and also by unknown causes. Rotavirus infects every child in its first few years of life with the first infection of rotavirus leading to acute diarrhoea. The proportion of rotavirus detected in children admitted to hospital with diarrhoea is highest in developed countries, but the majority of rotavirus deaths occur in developing countries (Tate *et al.,* 2012).

Children in developing countries account for 82% of rotavirus deaths (Khoury *et al.,* 2011). Figure 1.2 indicates that the majority of deaths occur in Sub-Saharan Africa and Asia, with more than 100-1000 deaths per 100 000 children younger than 5 years old (Tate *et al.,* 2012). A study conducted by Mwenda and co-workers in selected African countries showed that 40% of children included in this study were positive for rotavirus (Mwenda *et al.,* 2010). They also found that the prevalence of rotavirus infection amongst hospitalized children was

the highest in these regions than anywhere else (Mwenda *et al.,* 2010). Tate and co-workers conducted a study that also showed that 95% of rotavirus related deaths occurred in the 72 countries that are eligible to receive vaccines with financial support by the Global Alliance for Vaccine and Immunisation (GAVI) (Tate *et al.,* 2012). The national estimate of rotavirus deaths in children under the age of five ranges from less than five deaths in 74 countries to as many as 99 000 deaths in India alone (Figure 1.3). Tate and co-workers found that five countries accounted for half of all rotavirus deaths namely the Democratic Republic of the Congo, Ethiopia, India, Nigeria and Pakistan (Figure 1.3) with India having the greatest proportion of deaths (Tate *et al.,* 2012).

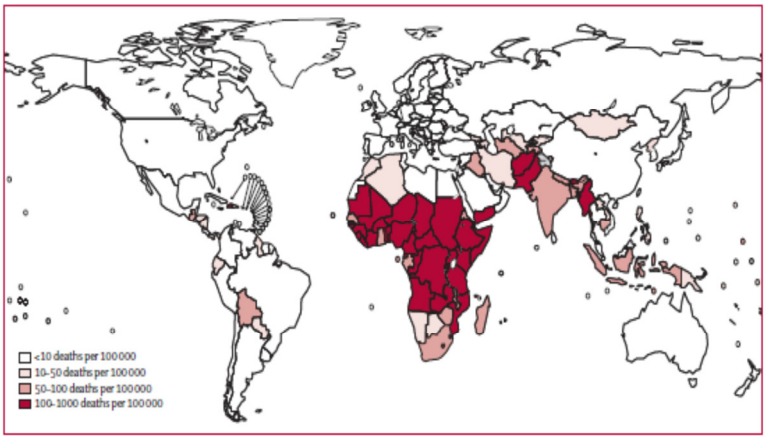


**Figure 1.2: Graphic representation of the number of deaths caused by rotavirus annually in children under the age of five. Each colour represents the number of deaths per 100 000 children younger than 5 years.** The dark pink regions represent 100 - 1000 deaths per 100 000 people, the medium pink 10 - 100 deaths per 100 000 people, the light pink 10-50 deaths per 100 000 people and the white regions <10 deaths per 100 000 people. Clearly the majority of deaths occur in the developing countries of South Asia and sub-Saharan Africa. Illustration taken from Tate *et al.,* 2012 with permission.

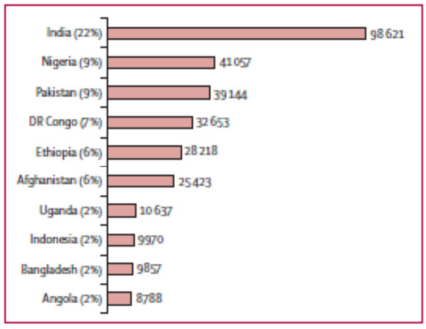


**Figure 1.3: Countries with the greatest number of rotavirus-related deaths in 2008.** Number of deaths due to rotavirus diarrhoea, with India accounting for 22% (99 000) of rotavirus related deaths, in children under the age of five (http://www.who.int/immunization/monitotringsurveillance/burden/estimates/rotavirus/en/index.html).

### 1.5 Molecular epidemiology

Before 1995, human rotavirus genotypes G1, G2, G3 and G4 were worldwide the most prevalent. Since 1995, the G9 genotype has emerged worldwide and is considered to be the fifth major human rotavirus genotype. The G12 genotype has since been found to circulate in most parts of the world and might become the sixth major human rotavirus genotype (Matthijnssens *et al.,* 2010a). In the developed countries; North America, Europe and Australia, the G/P types G1P[8], G2P[8], G3P[8], G4P[8] and G9P[8] represent more than 90% of the circulating genotypes (Seheri *et al.,* 2014). A study conducted in sixteen African countries (East, West, South and North Africa) showed that strains different to those prevalent in the developed countries were prevalent and circulating in these regions. The G1 genotype was most prevalent followed by G9, G2, G12 and G3. Similarly, the P[8] genotype was most prevalent followed by P[6] and P[4]. The most prevalent G/P combinations detected in Africa were G1P[8], G2P[4], G2P[6], G3P[6], G8P[6] and G12P[8] (Seheri *et al.,* 2014). These findings show that a high genetic diversity of rotavirus strains is circulating in the African regions.

### 1.6 Replication cycle of rotavirus

The rotavirus replication cycle has five main steps namely cell membrane attachment, membrane penetration, replication, assembly of new particles and the release of the newly formed virus.

Rotavirus attachment is achieved by the outer layer VP4 spike protein (Jayarem *et al.,* 2004), since newly assembled rotavirus virions are not infectious. The VP4 spike protein first has to be proteolitically cleaved by a trypsin-like protease of the host gastrointestinal tract into two proteins namely VP5$^{*}$ and VP8$^{*}$ (Figure 1.4) (Trask *et al.,* 2012). Endocytosis delivers the virion to the early endosome wherein the reduced $Ca^{2+}$ concentration is thought to trigger uncoating (the loss of VP7) of the triple-layered particle and membrane penetration by VP5*. The loss of the outer capsid proteins and the release of the double layered particle (DLP) into the cytosol activate the internal polymerase complex (VP1 and VP3), to transcribe capped positive sense RNA from each of the eleven double-stranded RNA (dsRNA) genome segments. Positive (+) RNAs serve either as mRNAs for the synthesis of viral proteins by cellular ribosomes or as templates for synthesis of negative-sense RNA [(-) RNA] during genome replication (Trask *et al.,* 2012). The non-structural proteins, NSP2 and NSP5 interact to form large inclusion bodies (viroplasms) that produce components necessary for genome replication and assembly of sub-viral particles.

Genome packaging is initiated when VP1 (and possible VP3) bind to the 3' end of the viral (+) RNAs (Trask *et al.,* 2012). It is thought that the interactions among the eleven (+) RNAs drive the formation of the "assortment complex". Following assortment, an assembling VP2 core shell engages the polymerase component of polymerase (+) RNA complexes, thereby activating the enzymes to initiate (-) RNA synthesis to produce the dsRNA genome (Figure 1.4). The intermediate capsid protein VP6 assembles onto the core to form the DLP, assembly of the outer capsid is not well understood. To assemble fully, the DLP must exit the viroplasm, associate with the VP4 spike protein and breach the endoplasmic reticulum (ER) membrane to gain access to the glycoprotein VP7 protein (Trask *et al.,* 2013). The current model proposes that the interaction of DLPs with the rotavirus transmembrane protein, NSP4, recruits DLPs and the outer capsid protein VP4 to the cytosolic face of the ER membrane. Through an undefined mechanism the DLP-VP4-NSP4 complex buds into the ER. Removal of the ER membrane and NSP4 permit assembly of the ER-resident outer capsid VP7 protein and the formation of the triple-layered particle (TLP). The release from the infected cell exposes the virion to trypsin-like proteases of the gastrointestinal tract, resulting in the specific cleavage of the VP4 into VP5* and VP8* to produce the fully infected virion.

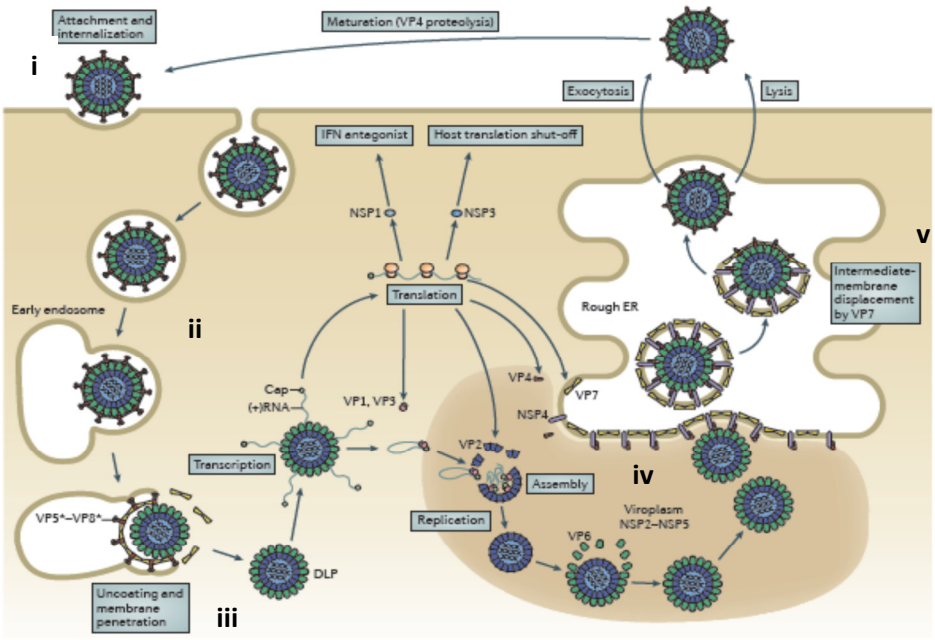


**Figure 1.4: Rotavirus replication cycle.** i) Attachment and neutralization of the non-infectious rotavirus viron to the target cell, ii) virion is delivered to early endosome, iii) uncoating of triple-layered particles and membrane penetration by VP5, iv) the assembly of the outer capsid, v) intermediate displacement by VP7 occurs. Illustration taken from Trask *et al.,* 2012 with permission.

## 1.7 Pathology of Rotavirus

Children are infected with rotavirus within their first few years of life, regardless of the level of hygiene or quality of food, water and sanitation. Rotavirus is mainly transmitted through the faecal oral route, from person to person, but can also be transmitted through airborne droplets or fomites on toys and countertops (Parez *et al.,* 2008). The virus is highly contagious and low concentrations are needed for infection. It can survive for days on environmental surfaces and for months in stool samples and at room temperature (Parez *et al.,* 2008).

Rotavirus infection can be symptomatic or asymptomatic. Symptoms can occur within 48 hours. Symptoms usually last for 4-8 days, depending on the severity of infection. Clinical symptoms of children infected are mild fever, nausea and vomiting, loss of appetite, abdominal pain and watery diarrhoea (Parez *et al.,* 2008). Once rotavirus is ingested and not neutralized by stomach acid it will attach to the enterocytes of the small intestine, were rotavirus infection mainly occur. During the first 18-36 hours after infection with rotavirus (the incubation period) extensive cellular necrosis of the epithelium will form, resulting in the

loss of digestive enzymes, lower absorption of fluids, higher osmotic pressure in the gut lumen and villous atrophy. Collectively, these effects will result in diarrhoea (Desselberger *et al.,* 2009). A secreted form of the endoplasmic reticulum (ER) transmembrane protein NSP4 is involved in diarrhoea induction though interactions with cellular plasma membrane integrin domains and possibly other receptors. It signals a phospholipase C-dependent increase in intracellular $Ca^{2+}$ and subsequently the release of chloride. Finally, the NSP4 C-terminus acts as an intracellular receptor for double-layered particles to facilitate infectious particle assembly in the ER (Hu *et al.,* 2012). During the post incubation period the epithelial surface is destroyed leading to blunted villi, extensive damage, water and nutrient loss, dehydration and massive quantities of villus shedding ($>10^2$ particles per gram) in stool samples of infected individuals (Glass *et al.,* 2006).

### 1.8 Immunogenicity of Rotavirus

Rotavirus infection has been associated with antigenemia (presence of antigen in the blood) and RNAemia (presence of viral RNA in the blood) (Moon *et al.,* 2012). However, rotavirus infection mainly induces both innate and acquired (humoral and cellular) immune responses upon natural infection (Angel *et al.,* 2012). Mechanisms responsible for immunity to rotavirus infection are not completely understood in humans, therefore, animal models are mainly used.

#### 1.8.1 Innate immune response

The innate immune response is the first barrier to infection (Liu *et al.,* 2009). The secretion and cytokines belonging to the interferon (IFN) family (type I, II and III) play an important role in the innate immune response by producing IFN stimulated gene (ISG) products (Arnold *et al.,* 2013). RNA viruses are recognized by the invaded cell through several pattern recognition receptors (PRR). Following PRR activation, signal transduction can be expected to activate the transcription factors, IFN regulatory factors (IRF3) and nuclear factor (NF)-ϰB, promoting optimal IFN-β expression (Arnold *et al.,* 2013). The IFN-β expression in rotavirus infected cells is inhibited by the NSP1 viral non-structural protein that has an affinity for IRF3. The suppression of IFN-β expression is not only mediated by the effect of NSP1 on IRF3, since the protein (NSP1) can also induce the degradation of other members of the IRF family including IRF5 and IRF7 (Angel *et al.,* 2012). NSP1 is a broad spectrum antagonist of type I IFN expression in infected cells since NSP1 has the capacity to target multiple members of the IRF family. NSP1 proteins of human rotavirus rely only on the degradation

of IRF5 and IRF7 to undermine the IFN signalling. In animal rotaviruses the NSP1 proteins tend to target IRF3, IRF5 and IRF7 (Arnold *et al.,* 2013). When IFN signalling is blocked, systematic virulence of particular strains are increased and lethal biliary and pancreatic diseases are induced.

### 1.8.2 Humoral immune response

Primary rotavirus infection causes a serotype-specific humoral immune response (Desselberger *et al.,* 2009). It has been suggested that humoral rotavirus immunity is correlated with protection. Homotypic and heterotypic neutralizing antibody response have been found in children after primary rotavirus infection. This suggests the presence of cross-reactive neutralizing antibodies (Desselberger *et al.,* 2011). Protection against subsequent rotavirus infections by different serotypes also develops and increases with the number of rotavirus re-infections. The exact correlates of protective immunity are not known but since rotavirus replication takes place in the intestinal enterocytes it is assumed that the effector mechanism must be active at the intestinal mucosa (Ward *et al.,* 2008). The most obvious immunological effector is immunoglobulin A (IgA). The first rotavirus infection is usually the most severe with severity decreasing as the number of rotavirus infection increases. Symptomatic as well as asymptomatic infection has similar degrees of protection against rotavirus infection (Franco *et al.,* 2006). Faecal specimens, which have been investigated for the presence of rotavirus-specific IgA antibodies, showed that at high levels IgA correlate well with protection. Rotavirus-specific serum IgA antibodies have neutralizing activity that also reacts with epitopes known to elicit heterotypic protection. Individuals with selective IgA-D deficiency may be protected from severe rotavirus disease by developing compensatory rotavirus-specific IgG responses that are higher than those in IgA competent persons (Desselberger *et al.,* 2011).

### 1.8.3 Cellular immune response

T-cells are cells that protect the host against the invading virus or antigen. Two mayor types of T-cells are found namely $CD8^+$ (cytotoxic T-cells) and $CD4^+$ (also known as T-helper cells). T-cell responses have been documented in mouse models. The $CD8^+$ cells kills target cells using one of two pathways. It gives complete protection (up to two weeks) after primary infection and partial protection (three months) after re-infection (Malik *et al.,* 2005). The $CD4^+$ T-cells are not only involved in supplying help to the $CD8^+$ T-cells, but also generate rotavirus specific immunoglobulin A (IgA) which mediates long term protection against rotavirus infection (Angel *et al.,* 2007).

## 1.9 Vaccines

An ideal vaccine would have the following qualities; (i) safe with no or few side-effects, (ii) easy and cheap to manufacture, (iii) stable for storage or transport, (iv) easy to administer, (v) could be given to infants alongside other childhood vaccinations and (vi) would stimulate life-long protection against all forms of disease (http://www.jenner.ac.uk). An effective rotavirus vaccine should, therefore, decrease mortality rates which occur as a result of severe diarrhoea as well as decrease the number of hospitalizations and doctor's visits due to mild infections. Such vaccines are needed in resource-poor countries because of the high mortality rates in these regions.

### 1.9.1 Vaccine history

In 1998, the first rotavirus vaccine, Rotashield, was released by Wyeth-Ayerst Laboratories now known as American Home products. It was a tetravalent, live attenuated rhesus monkey vaccine (RRV-TV) and was administered in three doses (Shadman *et al.,* 2000). In May 1999, nine cases of intussusception (developing of bowel obstruction) associated with children receiving Rotashield, were reported to the Vaccine Adverse Events Reporting System (VAERS) (Murphy *et al.,* 2001). The manufacturing company voluntarily withdrew Rotashield from the market in October 1999 and the Centre for Disease Control and Prevention (CDC) stopped the recommendation of Rotashield for routine immunisation of children (Glass *et al.,* 2004). Recent scientific studies conducted support the safety of Rotashield when it is administered at the appropriate age to infants and shows that it is not associated with intussusception. The International Medica Foundation is currently conducting a phase II clinical trial of the Rotashield vaccine in Ghana, Africa in association with the Noguchi Memorial Institute for Medical research (http://www.intl-medica.org/rotashield.asp).

### 1.9.2 Currently licensed vaccines

There are currently four licensed rotavirus vaccines worldwide namely RotaTeq, Rotarix, Lanzhou Lamb (only in China) and Rotavac (only in India). In 2009, the WHO Strategic Advisory Group of Experts (SAGE) recommended rotavirus vaccines to be introduced in all national immunisation programmes. They strongly recommended the introduction in countries where the rotavirus mortality rate of children under the age of five is more than 10% (Marlow *et al.,* 2012). It is, however, still the countries decision whether or not they

want to introduce the vaccine. These decisions are also influenced by the healthcare systems, economics and the burden of the disease.

As of April 2015, 77 countries (Figure 1.5) mostly middle, low and high income countries have introduced the rotavirus vaccine in their National Immunization Programme, including 30 GAVI eligible countries. Other countries such as Canada and Thailand have introduced rotavirus vaccine in pilot or regional introductions.

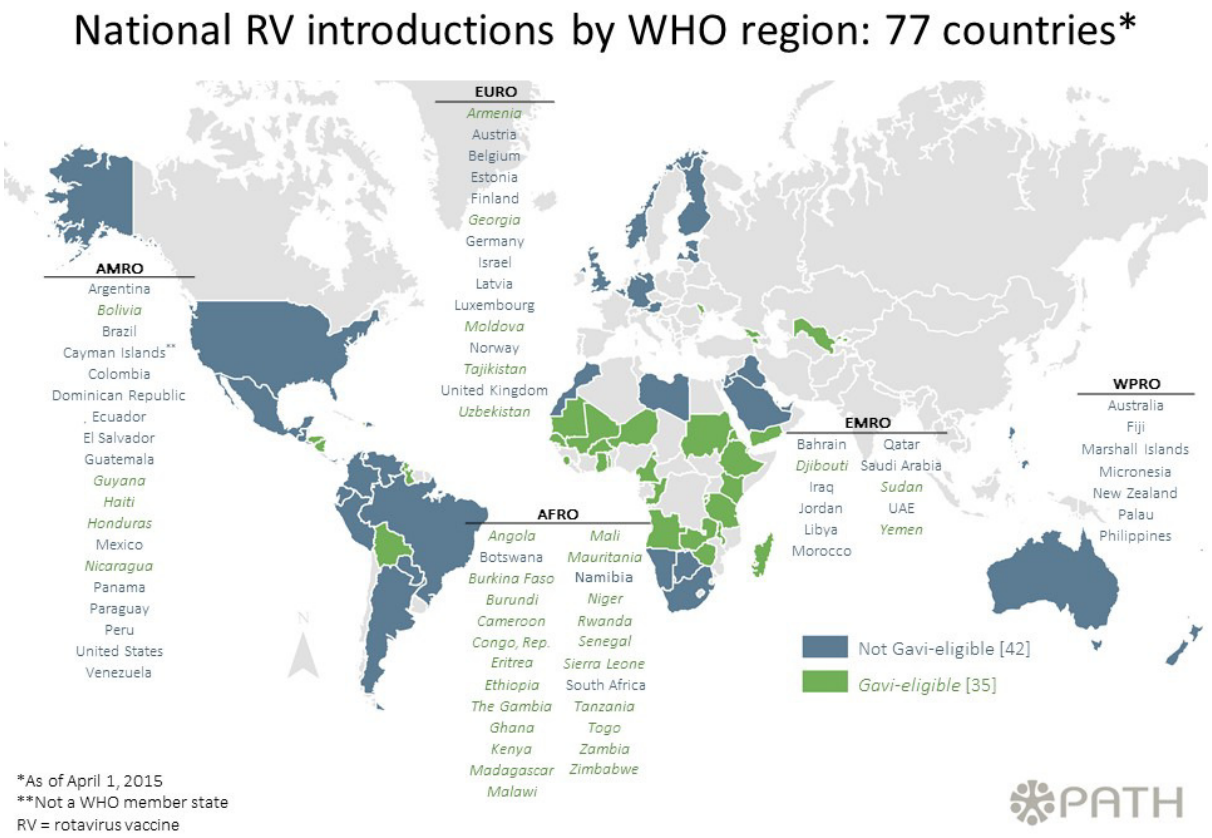


**Figure 1.5: Map demonstrating the national rotavirus vaccine introduction.** Countries indicated in green are GAVI eligible countries that have introduced the rotavirus vaccine. The countries indicated in blue are not GAVI eligible countries that have introduced the rotavirus vaccine. (http://sites.paath.org/rotavirusvaccine/rotavirus-advocacy-and-communications-toolkit/country-introduction-maps-and-list/).

The Global Alliance on Vaccines and Immunisation (GAVI) gives financial support to the developing countries that want to introduce the vaccines. GAVI began offering support for rotavirus vaccines in 2006. Since then the rotavirus vaccines have been introduced in 35

GAVI eligible countries. GAVI has approved four additional countries for rotavirus vaccine support (Figure 1.6) (http://sites.paath.org/rotavirusvaccine/rotavirus-advocacy-and-communications-toolkit/country-introduction-maps-and-list/).

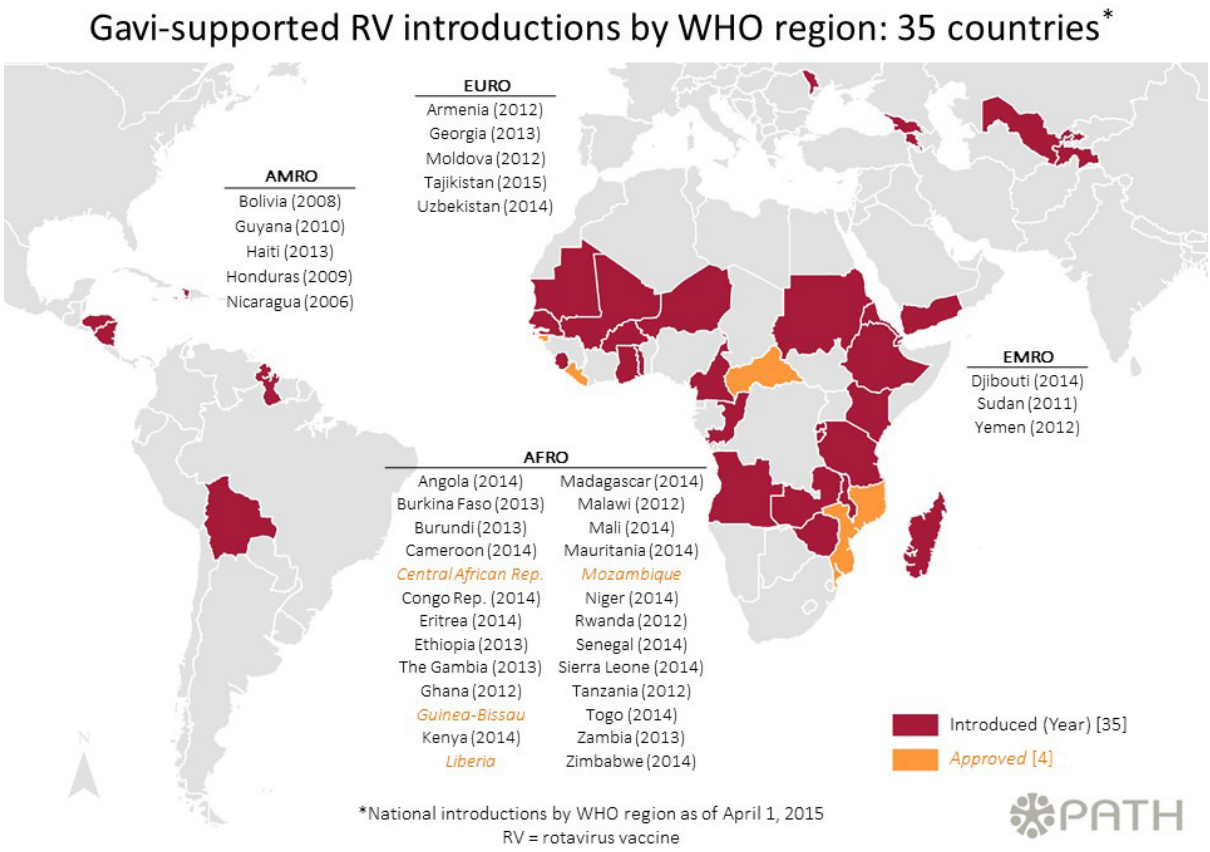


**Figure 1.6: Map demonstrating GAVI-supported countries and approved countries for rotavirus vaccine introduction.** Countries indicated in maroon have already introduced the vaccine while countries indicated in orange are approved by GAVI for rotavirus vaccine introduction (http://sites.paath.org/rotavirusvaccine/rotavirus-advocacy-and-communications-toolkit/country-introduction-maps-and-list/).

### 1.9.2.1 RotaTeq

RotaTeq is a pentavalent human bovine (WC3) live-attenuated vaccine produced by Merck Research Corporation (Matson *et al.,* 2006). The bovine rotavirus strain (WC3) was isolated from a calf in Pennsylvania, in 1981 and used as the starting point to develop a multivalent vaccine (Ciarlet *et al.,* 2009; Mathijnssens *et al.,* 2010*b*). WC3 is one of the parent strains and has a genotype of G6P[7]. WC3 as a monovalent was found to be immunogenic but

gave inconsistent results in efficacy studies done on humans (Tom-Revzon *et al.,* 2009). Improvement on the bovine rotavirus (WC3) by developing reassortants with human VP4 or VP7 led to the current rotavirus vaccine, RotaTeq (Tom-Revzon *et al.,* 2009). The currently licensed RotaTeq vaccine contains five human bovine reassortant strains (G1-G4 and P1[8]) each expressing a different VP7 or VP4 surface protein on the backbone of the naturally attenuated tissue culture adapted parental bovine rotavirus strain (G6P[7]) (Matthijnssens *et al.,* 2010b).

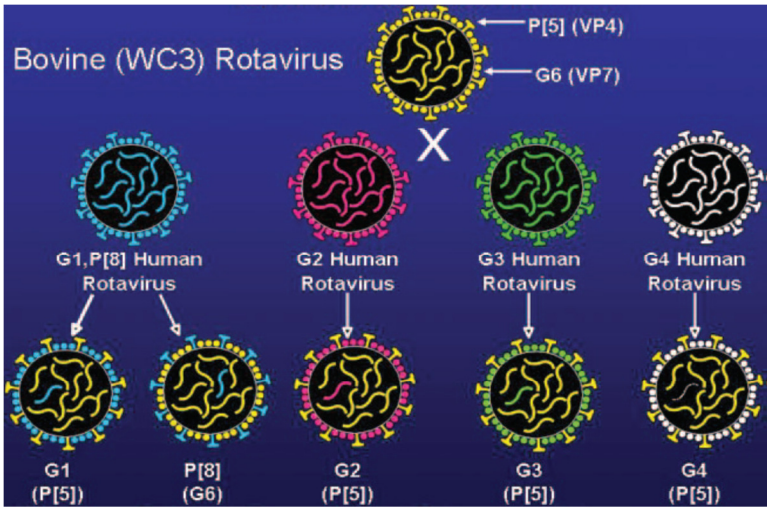


**Figure 1.7: Attenuated human-bovine rotavirus reassortant vaccine (RotaTeq).** The vaccine contains five reassortant rotaviruses. Four reassortant rotaviruses express the VP7 protein (G1, G2, G3 or G4) from the human rotavirus parent strain and the VP4 protein (P7[5]) from the bovine rotavirus parent strain. The fifth reassortant virus expresses the VP4 proteins (P1A[8]) from the human rotavirus parent strain and the outer capsid protein G6 from the bovine parent strain. Illustration taken from Dennehy *et al.,* 2008 with permission.

The US Food and Drug administration approved RotaTeq on February 3, 2006 and on February 21, 2006 the Advisory Committee on Immunization Practice (ACIP) recommended RotaTeq for the routine immunization programme in the US. The vaccine is administered with other licensed vaccines in the routine immunization schedule at 2, 4 and 6 months of age (Ciarlet *et al.,* 2009).

### 1.9.2.2 Rotarix

Rotarix is a live-attenuated vaccine, produced by GlaxoSmithKline Biological. The vaccine contains the RIX4414 strain of G1P[8] rotavirus (Figure 1.8) (Bernstein *et al.,* 2006). Rotarix was developed from an isolate that was isolated from an infant in Cincinnati in 1989. The isolate was designated 89-12 and showed protection against rotavirus. This evidence provided the recognition that a human rotavirus rather than an animal rotavirus strain can be used to develop a vaccine. An initial randomized placebo-controlled double blind efficacy trial was conducted with the vaccine in 213 infants. Vaccine efficacy in the first year of life was 89% against rotavirus disease and 100% against very severe disease (Bernstein *et al.,* 2006). The vaccine was initially named RIX4414 and further development was done by limiting dilution cloning of 89-12 in Vero cells leading to the currently licensed vaccine.

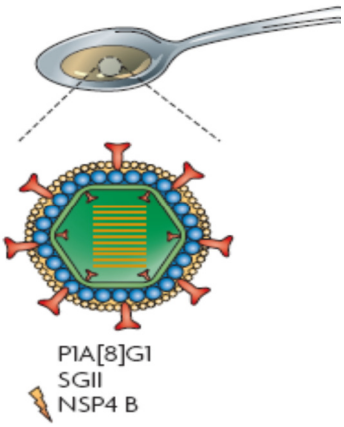


**Figure 1.8: Rotarix attenuated human rotavirus vaccine.** Vaccine is produced in a tissue culture adapted human P1A[8]G, VP6 subgroup II and NSP4 geno-group B strain. Illustration taken from Angel *et al.,* 2007 with permission.

In 2007, the World Health Organisation prequalified Rotarix in the United States, based on the efficacy and safety data obtained from Latin America and Western Europe trials. The Food and Drug Administration (FDA) licensed Rotarix, in the United States, in April 2008. Rotarix was, however, registered first in Mexico, in 2004. In August 2009, South Africa became the first African country to include Rotarix into the Expanded Programme on Immunisation (Madhi *et al.,* 2014). The vaccine is administered in two doses with other childhood licensed vaccines on the routine immunization schedule at 2 and 4 months of age.

#### 1.9.2.3 Lanzhou lamb

Lanzhou lamb is a live monovalent serotype (P[2]G10] of group A rotavirus. The vaccine was isolated from a local lamb (in China) with diarrhoea and grown in kidney cells for 37 generations (Fu *et al.,* 2007). It is manufactured by the Lanzhou lamb Institute of Biological Products (Lanzhou, China) and was licensed in 2000 for gastroenteritis among children (Fu *et al.,* 2007).

Lanzhou lamb is administered in one dose annually for children 2 months – 3 years and 3-5 years old. Since it was launched, nearly 500 000 children younger than 5 years have been immunised (Fu *et al.,* 2007). However, the vaccine is only licensed in China (Parasher *et al.,* 2006) due to the little data that is available of the vaccine's safety, immunogenicity and efficacy. The efficacy is unknown due to the fact that the vaccine was not tested against placebo in a controlled phase III trial (Fu *et al.,* 2007).

#### 1.9.2.4 Rotavac (*Bovine/neonatal 116E strain)*

The 116E strain was isolated from asymptomatic newborns at a hospital in New Delhi, India in 1985 (Glass *et al.,* 2005). Sequence analyses showed that the 116E is a G9P[1] strain, with VP4 being similar to that of several bovine rotavirus strains (Rippinger *et al.,* 2010). The 116E monovalent vaccine was adapted to grow on Vero cells by the Manufacturer Bharat Biotech International Ltd (BBIL) and was formulated as a vaccine candidate (Parasher *et al.,* 2006).

Phase I/II immunogenicity and safety trials were conducted in infants aged 8 to 20 weeks at both low ($1 \times 10^4$ ffu) and high ($1 \times 10^5$ ffu) doses. The vaccine was administered three times separately from the routine childhood vaccines (Bhandari *et al.,* 2009). Immunogenicity was seen after the first dose and reached immunogenic rates of 89.7% after the third administration of the high dose (Bhandari *et al.,* 2009). The results were encouraging enough to conduct a large clinical trial to evaluate protective efficacy of the 116E vaccine in a field setting and with the other immunisation vaccines (Bhandari *et al.,* 2009). Since 2001, PATH (Programme for Appropriate Technology in Health) has been part of a collaborative effort to develop and evaluate 116E. PATH is supporting India's National Institute of Immunology, in close collaboration with Indian Department of Biotechnology and BBIL, to prepare the vaccine for phase III efficacy trials (www.path.org; developing new vaccine against rotavirus).

A multicentre double blind placebo controlled phase III trial was conducted in India (March 2011-September 2013). The study was conducted in infants age 6-7 weeks which received

three doses of the oral human bovine natural reassortant vaccine (116E) at ages 6-7, 10 and 14 weeks along with other childhood vaccines (Bhandari *et al.,* 2014). Vaccine efficacy against severe rotavirus gastroenteritis in children up to two years of age was 55.1% (first year of life was 56.4% and second year of life was 48.9%). The 116E vaccine, now known as Rotavac, was licensed in India in January 2014 (http://www.who.int/vaccine_safety/committee/topics/rotavirus/rotavac/Jun_2014/en/).

### 1.9.3 Shortcomings of currently licensed vaccines

The currently licensed vaccines have shown to be effective in clinical studies conducted worldwide. However, both Rotarix and RotaTeq have some shortcomings as will be discussed below.

#### 1.9.3.1 Efficacy trials

Efficacy trials performed with Rotarix in South Africa and Malawi showed that it was less effective in Malawi (49.9%) and South Africa (76.9%) (Madhi *et al.,* 2010) than in Europe and North America (with 95-98%) against severe rotavirus gastroenteritis (Vesikari *et al.,* 2007; Linhares *et al.,* 2008; Ruis-palacois *et al.,* 2006). The same was observed for efficacy trials performed with RotaTeq in Africa and Asia, which showed that it was less effective in Africa (64.2%) and Asia (51%) (Armah *et al.,* 2010; Zaman *et al.,* 2010) than in the United States ($\pm$90%) (Vesikari *et al.,* 2006; Ciarlet *et al.,* 2009). These efficacy trials, therefore, suggest that Rotarix and RotaTeq are more effective in developed countries than in some developing countries.

The difference of efficacy between developed and developing countries have yet to be clearly identified, but can be due to a number of factors. The one reason can be due to the difference in strains, since the strains used to develop RotaTeq and Rotarix were selected from strains circulating in the developed world (USA and Europe), namely G1, G2, G3 and G4, while the emerging strains in Africa G8, G9, G12 and G2 (Steele *et al.,* 2012; Seheri *et al.,* 2014) were not included. The variability in strains could offer an explanation for the lower efficacy of these vaccines observed in the developing countries.

It should be considered that unlike the developed countries, children with HIV and malnourishment were included in some of the studies conducted in Africa, which can also impair their immune response to a rotavirus vaccine (Patel *et al.,* 2009). The higher background of rotavirus, other enteric co-infections and chronic diseases (malaria and

tuberculosis) that are prevalent in these populations can play a role in vaccine efficacy in developing countries. The co-administration of oral polio vaccine (OPV) and breastfeeding at the time of vaccine administration may also play a role, as well as the interference of passively acquired maternal antibodies (Chan *et al.,* 2011).

#### 1.9.3.2 Severe combined immunodeficiency (SCID)

Severe combined immunodeficiency (SCID) is defined as a group of genetic disorders that results in the lack of T-cell and B-cell immunity. The diseases are usually characterized by life threatening infection during the first year of life and can be fatal unless it is treated, usually with stem-cell transplantation (Patel *et al.,* 2010).

The safety and effectiveness of Rotarix and RotaTeq in infants with primary or secondary immuno-deficiencies have not been evaluated. These include infants on immuno-suppressive therapy and infants with malignant neoplasms affecting bone marrow or lymphatic system. However, current guidelines support the administration of rotavirus vaccines to children infected with HIV, the largest immunosuppressed population study to date. The first reported case of SCID was of a 9 month old baby that presented with a history of faltering growth and chronic diarrhoea. She was fully immunized and also received the oral RotaTeq vaccine at 2, 4 and 6 months. She had mild diarrhoea after the first dose and remained well until four months, after which she developed persistent vomiting and diarrhoea with poor weight gain that worsened at six months. Assessment of her chronic diarrhoea revealed rotavirus in her stool, lymphoenia, and undetectable IgG, IgA and IgM. Lymphocyte subsets confirmed absent T-cells with absent lymphocyte function and normal levels of B- and natural killer cells (Werther *et al.,* 2009). A diagnosis of severe SCID was made (Werther *et al.,* 2009). An additional three cases of SCID babies presenting with diarrhoea were reported following vaccination with RotaTeq. Three patients with SCID were given two doses (at 2 and 4 months) live, oral pentavalent rotavirus vaccine (Rotateq). All three patients presented symptoms of diarrhoea and failure to thrive. Symptoms occurred after receiving the second dose in two of the patients (patient 1 and 2) while symptoms already appeared after the first dose in the third patient. This can be an indication of early protection in infancy by transplacentally acquired maternal antibodies (Patel *et al.,* 2010). Rotavirus was detected in stool specimens of all three patients with one patient having co-infection with giardia and adenovirus as well (Patel *et al.,* 2010). Rotavirus was still present in stool specimens up to the age of 9 months (patient 1), 8.5 months (patient 2) and 14 months (patient 3) in the infants. After these reports of rotavirus vaccines and SCID the Food and Drug administration updated the package insert for both vaccines and listed SCID as a conflict for the use of live oral rotavirus vaccines.

These cases of rotavirus infection in patients with SCID raise concern regarding the safety of live-attenuated rotavirus vaccines in children with severe combined immunodeficiency disease. This is especially of interest for developing countries that account for the most rotavirus deaths and have a high prevalence of severe immunocompromised (HIV) patients.

#### 1.9.3.3 Reassortment

Reassortment is the mixing of genetic material of two different virus strains infecting the same cell.

There are different ways that reassortment can occur in the case of rotavirus, namely (Gentsch *et al.,* 2005)

- *reassortment between the common rotavirus strains*
- *Animal human rotavirus reassortment:* where animal rotavirus genes can be introduced to human rotavirus through reassortment.

Rotavirus vaccines are administered through the oral route and have the risk of reassortment with field strains, leading to a virulent virus. Payne and co-workers (2010) documented a case where reassortment occurred between the RotaTeq vaccine strains of genotypes G1P7[5] and G6P1A[8], during intestinal replication. Transmission occurred from the young vaccinated to the older unvaccinated sibling causing symptomatic rotavirus gastroenteritis that required medical care (Payne *et al.,* 2010). Another study conducted by Donate and co-workers (Donate *et al.,* 2012) documented cases where viral specimens were associated with a G1P[8] strain, resulting through genetic reassortment between two component RotaTeq strains. This study showed that during the replication and excretion of RotaTeq vaccine, reassortment of parental strains can occur (Donate *et al.,* 2012). However, the benefits of vaccination outweigh any small risk of vaccine associated gastroenteritis. These studies showed that reassortment is possible when virus shedding and transmission occur from a vaccinated to an unvaccinated infant as well as between two vaccine strains.

#### 1.9.3.4 Breastfeeding

Maternal antibodies are transmitted from mother to foetus via the placenta and breast milk. Such antibodies can provide immunity for the foetus and new born for up to 6 months. There are three factors that can alter the effective titer of the vaccine, namely amount of neutralizing activity in the breast milk, the effect of breastfeeding and the practises around the time of breastfeeding. Both IgA antibodies (neutralizes rotavirus) and receptor

analogues (absorb to virus and inhibit attachment) are found in breast milk. These factors will decrease as breastfeeding progresses but are the highest during the first days after birth (Patel *et al.,* 2009).

Therefore, if the vaccine was administered during a time in which the child was not fed, the vaccine strain can pass through to the gut unaltered and start replicating. However, if the child received the vaccine after a feeding, the mouth and gastrointestinal tract can become *in vivo* sites for virus neutralization and the vaccine will not be effective. A study conducted by Moon and co-workers (Moon *et al.,* 2013) showed that higher levels of lactoferin, loactohedrin, IgA and neutralizing activity were present in breast milk specimens from women in India and Africa than from women in America. They also demonstrated positive associations between levels of lactoferin or IgA and neutralizing activity in Indian and African women. Therefore, the lower immunogenicity and efficacy of rotavirus vaccines in developing countries could be explained in part by the co-active inhibitory effect of high levels of antibody and non-antibody components in breast milk that are consumed by infants at the time of immunisation (Moon *et al.,* 2013).

Recently Groome and co-workers (Groome *et al.,* 2014) conducted a study that showed that breastfeeding does not have an effect on infant immune response to the rotavirus vaccine. The study was conducted in South Africa (Soweto) to determine whether or not abstence from breastfeeding an hour before or after each vaccination had an impact on the immune response of infants receiving two doses of the rotavirus vaccine, Rotarix. The study was conducted on only HIV uninfected infants at the 6 week infant immunisation (Groome *et al.,* 2014). Groome and co-workers found that the abstention from breastfeeding for at least one hour before or after each vaccination dose had no significant effect on the frequency of sero-conversion among the vaccinated infants or the titres of anti-rotavirus IgA in the sera of the same infant (Groome *et al.,* 2014). This study showed that breastfeeding is probably not the cause of the lower efficacy of rotavirus vaccines in lower income countries. Therefore, the reasons that oral rotavirus vaccines appear to have relatively low efficacy in low income countries require further investigation. These studies include co-administration of oral polio vaccine, micronutrient deficiency, enteric co-infection and other diseases such as HIV (Groome *et al.,* 2014).

#### 1.9.3.5 Porcine circovirus (PCV)

Porcine circovirus (PCV) is a single-stranded DNA virus that is non-enveloped with an unsegmented circular genome (Ma *et al.,* 2011). A next generation sequencing approach led to the discovery of PCV nucleic acid sequences in Rotarix (Victoria *et al.,* 2010) and it

showed that the sequences represented infectious PCV particles. The Food and Drug Administration (FDA) temporarily suspended the use of Rotarix, on March 22, 2010 (Dore *et al.,* 2012). Traces of PCV1 and PCV2 DNA fragments were also identified in RotaTeq. On May 7, 2010 the suspension of Rotarix was lifted due to the fact that the benefits of rotavirus vaccination outweigh any associated risk with the use of Rotarix or RotaTeq (http://www.who.int/vaccine-safety/topics/rotavirus/rotarix-statment-march-2010/en/index.html). The FDA and both vaccine companies have updated the labelling of both Rotarix and Rotateq stating the presence of PCV1 (in Rotarix) and PCV1 and PCV2 (in RotaTeq) in the vaccines (American Academy of Paediatrics, 2010).

#### 1.9.3.6 Intussusception

Intussusception is a form of bowel obstruction which occurs when one segment of the bowl becomes enfolded within another segment, which if not treated, can be fatal (Tate *et al.,* 2012). The first case of intussusception associated with rotavirus vaccines was with the first rotavirus vaccine, Rotashield in 1999 (Patel *et al.,* 2009). Due to the concerns regarding a potential age dependant risk of intussusception with the previous rotavirus vaccine, strict age of administration guidelines were, therefore, implemented for the new rotavirus vaccines. The currently licensed vaccines have been carefully monitored, initially by large safety and efficacy studies and by post marketing surveillance as well. Both currently licensed rotavirus vaccines, Rotarix and RotaTeq, were associated with intussusception during clinical trials (Tate *et al.,* 2012). However, post-marketing surveillance of the currently licensed vaccines has indicated a small risk of intussusception (1-2 cases per 100 000 infants vaccinated) detected in some populations (Europe, Mexico, Brazil, United States, Australia) following immunisation with the first dose of both currently licensed rotavirus vaccines (Tate *et al.,* 2012). None the less the immunisation committees continue to recommend the use of rotavirus vaccines given that the benefits of the vaccine exceed the risk, but further research is needed to understand fully the association between rotavirus vaccination and intussusception (Tate *et al.,* 2012).

#### 1.9.3.7 Other shortcomings

***Oral polio vaccine (OPV):*** The simultaneous administration of the OPV has the potential to interfere with the oral rotavirus immune response and is one of the reasons for the lower rotavirus vaccine efficacy in developing countries. OPV does not interfere with rotavirus in developed countries, since inactivated polio vaccines (IPV) are used in these countries. Over 140 countries use OPV as part of their routine immunisation programme (Patel *et al.,* 2012).

The potential of mutual interference between OPV and rotavirus exists since both are live, attenuated vaccines and have the potential that their virus strains can replicate in the gut. Several studies for both Rotarix (South Africa, Bangladesh, and Latin America) and RotaTeq (four Latin America countries) have been done to assess the effect of OPV and rotavirus co-administration. It is clear from the data obtained that OPV does interfere with the immune response of the first dose of rotavirus vaccines. The interference can, however, be overcome after the completion of the vaccine series. Further research is still necessary to understand the full impact of OPV interference with the rotavirus vaccine (Patel *et al.,* 2012).

***Interfering gut flora:*** The overgrowth of small bowel bacteria can impair the immune response to live oral vaccines. Therefore, co-inhibiting bacteria and viruses could decrease the immune response to live vaccine viruses (Patel *et al.,* 2009). This is especially of interest to developing countries where children may have an abundance of enteric pathogens in their gut by the age of three months (this also includes enteroviruses, that are not part of the normal gut flora).

***Other medical conditions:*** Infants in developing countries such as Africa and Asia may also be exposed to other acute and chronic conditions such as HIV infection, tuberculosis, malaria, diarrhoea and fever, which can lower the vaccine efficacy of the rotavirus vaccines (Patel *et al.,* 2009).

### 1.9.4 Alternative vaccines

#### 1.9.4.1 Live vaccines

The objective of a successful live oral vaccine is that it must provide a high level of protection against the disease. Data has implied that an attenuated rotavirus vaccine, which mimics natural infection, can provide protection against the disease (Tate *et al.,* 2010). There are currently four licensed live attenuated rotavirus vaccines (as described in section 1.9.2), however these vaccines are usually not highly effective in some countries and the vaccine manufacturing companies also do not have enough production capacity. Therefore, research in rotavirus vaccines has to continue to make the existing vaccines more affordable as well as to develop new, safer and more affordable vaccines. There are currently two alternative live vaccines that are being investigated.

***Bovine reassortant vaccine***

The Bovine UK Compton strain (G6P[7]) was obtained from a stool sample of a colostrum derived calf, with diarrhoea and was isolated from calf kidney cells in the United Kingdom (Clements-Mann *et al.,* 1999). It is a tetravalent vaccine candidate containing 10 genome segments from the G6P[7]) bovine UK Compton strain and the VP7 genome segment of human rotavirus serotypes (G1-4) (Clements-Mann *et al.,* 2001).

The vaccine was tested in both the United States and Finland and was developed in parallel with Rotashield (RRV-TV). In Finland it showed non-reactogenicity in phase I trials as well as satisfactory immunogenicity in phase II trials (Vesikari *et al.,* 2006). This vaccine is seen as a "designer vaccine" meaning that the VP7 can encode for specific serotypes to accommodate the emergence of new and different serotypes (Kapikian *et al.,* 2005). The vaccine candidate was developed at the Laboratory of Infectious Diseases of the National Institute of Allergy and Infectious diseases (NIAID, Bethesda, MD) (Clements-Mann *et al.,* 1999) and the National Institute of Health (NIH). The NIH has licensed the vaccine to seven companies in three countries namely Brazil, India and China. PATH (Programme for Appropriate Technology in Health) is supporting the development of the BRV, by providing financial support and scientific assistance to two selected manufacturers. These manufacturers are Shantha Biotechnics Ltd in India since 2006 and the National Biotec Group's Wuhan Institute of Biological Science in China since 2007 (PATH).

***Human neonatal rotavirus RV3 strain***

The RV3 (serotype G3 strain) was found in healthy newborns at the Melbourne obstetric hospital in Australia, 1970 (Barnes *et al.,* 2002). Sequence analyses showed that RV3 is a naturally attenuated G3P[6] strain (Riepinger *et al.,* 2010).

A phase I/II trial was conducted in children receiving the vaccine at age 3, 5 and 7 months. The trial showed the vaccine to be safe but it showed low immunogenicity (Barnes *et al.,* 2002). A second generation, increased titre RV3-BB vaccine was developed with the expectation of improved immunogenicity. A single centre, double-blind randomised placebo controlled phase I study was conducted to evaluate the safety, tolerability and immunogenicity of a single oral dose of the second generation RV3-BB vaccine. Sixty participants across three age groups, 20 adults (age 18-50 years), 20 children (age 3-5 years) and 20 infants (age 16-18 weeks) were included in the study. The RV3-BB rotavirus vaccine was well tolerated in all three age groups in the phase I clinical trial. There was evidence of immunogenicity with vaccine intake in the majority of infants (89%) with a single dose of RV3-BB. Phase II trials to assess the immunogenicity and efficacy of the RV3-BB

rotavirus vaccine, with the first dose delivered at birth or at 6-8 weeks of age, commenced in 2011 (data not yet available) (Danchin *et al.,* 2013).

#### 1.9.4.2 Non-live vaccines

Although live oral vaccines are the primary approach for vaccine development, other approaches and routes of administration are being studied (Kang *et al.,* 2006). Non-live rotavirus vaccines are being studied as the next generation of vaccines. Three of these non-live approaches used for non-live rotavirus vaccines are discussed below.

***Inactivated rotavirus particles (IRV)***

An inactivated vaccine (or killed vaccine) is defined as a vaccine that consists of virus particles which are grown in culture and then killed using a method such as thermal (heat) or chemical inactivation (formaldehyde, aziridines or β-propiolactone). Chemical inactivation has been used to prepare inactivated vaccines against viruses such as polio, influenza and others. However, chemical treatments can cause physical and biochemical damages to the virus. Thermal inactivated rotavirus particles maintain their structural, biochemical and antigenic integrity (Jiang *et al.,* 2008). Inactivated rotavirus vaccines (IRV) have advantages over live oral vaccines, such as that IRVs do not have to be administered orally, thereby causing ease of administration.

Inactivated rotavirus vaccines have been tested in rabbits, mice and gnotobiotic piglets. Jiang and co-workers used a heat inactivated vaccine in the study they conducted (Jiang *et al.,* 2008). It was shown in mice that the vaccine was highly immunogenic. They also examined whether or not the addition of aluminium hydroxide ($Al(OH)_3$) to the vaccine will enhance the immune response (Jiang *et al.,* 2008). Results obtained indicated that a very low dose of thermally inactivated vaccine (formulated with $Al(OH)_3$) was highly immunogenic (Jiang *et al.,* 2008). Another study conducted by Wang and co-workers (Wang *et al*., 2010) assessed the immunogenicity and protection of a possible inactivated rotavirus vaccine candidate with the human strain CDC-9 (G1P[8]) formulated with aluminium phosphate against rotavirus infected piglets (Wang *et al.,* 2010). Results obtained showed partial protective immunity in piglets (Wang *et al.,* 2010). The study's findings were also in agreement with a study conducted by Dr Nakagomi, in Japan (Wang *et al.,* 2010). Dr Nakagomi's study showed that gnotobiotic piglets were completely protected from a partially homologous challenge with a rotavirus G1P[8] strain, after they were immunised with an inactivated G1P[4] rotavirus strain vaccine candidate in the absence of an adjuvant. Although these studies conducted are promising, these vaccines will still have to be tested in clinical trials.

***Triple- and double-layered virus-like particle vaccines***

Triple-layered and double-layered particles (VLPs and DLPs) are formed from expressed recombinant proteins and usually contain VP2 and VP6, with or without VP4 and VP7. There are several expression systems that can be used to produce the recombinant proteins (that become incorporated into VLPs). However, the most common expression system is the baculovirus system (Ward *et al.,* 2008). VLPs have been tested as a rotavirus vaccine in mice, rabbits and gnotobiotic piglets. Partial or total protection was obtained against virus shedding in mice and rabbits (Conner *et al.,* 1996). VLPs are more protective in gnotobiotic piglets when given with an adjuvant and are also more effective when given intranasally than orally (Ward *et al.,* 2008). Azvedo and co-workers showed that VLPs can be used as a booster after priming mice with a live attenuated virus (Azvedo *et al.,* 2010). It was shown that prime/booster vaccines are a more effective approach than multiple doses of live or VLP vaccines alone (Azvedo *et al.,* 2010). The absence of the viral genome leaves space for components and, therefore, VLPs can be used as a delivery agent to target organs or cells, as shown by Cortes and co-workers (Cortes *et al.,* 2010). They found that VLPs can be used as a drug delivery system for the efficient delivery of bio-molecules into the colon (Cortes *et al.,* 2010).

Since VLPs cannot replicate in the intestine they are anticipated to be safe when administered to infants, especially those with malnourishment and diseases that render them immuno-compromised. VLPs may be especially useful for children in developing countries where rotavirus infections are high (Jiang *et al.,* 1999). Therefore, VLPs might provide a safer and efficacious alternative to, or complement live rotavirus vaccines, for both human and veterinary use (Conner *et al.,* 1996).

**VP6 based vaccine candidate**

VP6 is one of the six structural proteins of rotavirus and together with VP2 is necessary for the formation of the double-layered particle. VP6 is essential in the viral replication cycle and determines the antigenic subgroups (A-H). It also has a possible role in immunity making it a good candidate for a non-live vaccine (Ward *et al.,* 2010).

It has been shown in the literature that recombinant rotavirus VP6 can induce protection in mice. In one study they used recombinant VP6 expressed in *Escherichia coli* as a fusion chimeric protein. The recombinant chimeric fusion protein suppressed intestinal rotavirus antigen production >93% when it was administered to mice after a murine rotavirus challenge (Choi *et al.,* 1999). Another study showed that immunisation with DNA encoding VP6 induced protection in mice (Hermann *et al.,* 1996). Ward and co-workers expressed

VP6 in *Escherichia coli* as a fusion protein with the maltose binding protein (MBP:VP6). The VP6 vaccine candidate was tested in mice. Mice received two doses of the MBP:VP6 purified protein intranasally with the adjuvant LT (R1926G). Mice were challenged with the EDIM (Epizootic diarrhoea of infant mice) virus. Results showed that protection against EDIM virus shedding was greater in mice receiving the MBP:VP6 with adjuvant than the mice receiving only the MBP:VP6 without adjuvant. The results showed that one dose of MBP:VP6 with adjuvant can promote long term protection (Ward *et al.,* 2010). The study only showed the protection obtained in mice.

Another study conducted by Lappalainen and co-workers produced a human recombinant VP6 protein in Sf9 cells by baculovirus expression system (Lappalainen *et al.,* 2013). BALB/C mice received two doses of the recombinant VP6 intramuscularly. Recombinant VP6 induced a balanced T-helper response (Th-1 and Th-2 type) and high levels of serum IgG antibodies with cross reactivity against six different rotavirus strains (Wa, SC2, BrB, 69M, L26, WC3 and RRV). Mucosal VP6 specific IgG and IgA antibodies were also detected in faeces and vaginal washes of the immunized animals (Lappalainen *et al.,* 2013). These studies gave evidence that VP6 can be considered as a non-live vaccine candidate against rotavirus. If a VP6 vaccine is found to be safe and effective in humans, as observed in animals, it possibly has the potential to be one of the next generation vaccines.

## 1.10 Expression systems used for generating recombinant proteins and VLPs as vaccine candidates

Expression systems can be classified into two main groups: prokaryotic expression systems (bacteria) and eukaryotic expression systems (yeast, fungi, insect/baculovirus, plant cells and mammalian cells). A wide range of expression systems are available, however, only three expression systems will be discussed here namely baculovirus expression (gold standard for rotavirus DLPs and VLPs), yeast and bacteria (the two expression systems that will be used in this project).

### *Insect/Baculovirus expression system*

The baculovirus expression system is the main expression system used in the expression of rotavirus virus-like particles as well as in studies to show that the VLPs induce protection. Commercialised baculovirus expression systems (BVES) have many advantages, namely (i) they are safe to use (non-pathogenic) and(ii) easy to scale-up. Baculovirus have been reproducibly scaled up for large-scale production of biologically active recombinant products,

(iii) high levels of recombinant gene expression are possible and (iv) simultaneous expression is possible. The most commonly used insect cell lines in the baculovirus expression system is Sf9, Sf21 and High Five cells. The Bac-to-Bac expression system is mainly used (Invitrogen manual, 2004). A few examples of Baculovirus studies include:

**(i) Rotavirus-like particles administered mucosally induce protective immunity (O'Neal *et al.,* 1997)**

O'Neal and co-workers produced the VP2/6 and VP2/6/7 VLPs by self-assembly of individual rotavirus structural proteins, co-expressed by baculovirus recombinants in insect cells (O'Neal *et al.,* 1997). Mice were administered with the VLP as well as with the VLP and the cholera toxin adjuvant. They found that protection levels were higher when the VLPs were administered with the cholera toxin adjuvant (O'Neal *et al.,* 1997).

**(ii) Rotavirus 2/6 virus-like particles administered intranasally with Cholera Toxin, *Escherichia coli* Heat labile Toxin (LT), and LT-R192G induce protection from rotavirus challenge (O'Neal *et al.,* 1998)**

O'Neal and co-workers made rotavirus VLPs by co-infecting cells with baculovirus recombinants that expressed rotavirus structural proteins (O'Neal *et al.,* 1998). The rotavirus VP2/6 VLPs was administered to mice along with three different adjuvants namely the cholera toxin (CT), *E.coli* heat labile toxin (LT) and LT-R192G (mutant of *E.coli* heat labile toxin). All three adjuvants induced protection, in mice, when administered with the 2/6 VLPs (O'Neal *et al.,* 1998).

**(iii) Characterization of virus-like particles produced by the expression of rotavirus capsid proteins in insect cells (Crawford *et al.,* 1994).**

Crawford and co-workers cloned each of the rotavirus genes, coding for a major capsid protein (VP1, VP2, VP3, VP4, VP6 and VP7), into the baculovirus expression system and expressed each protein in insect cells. They also found double-layered particles containing VP2/6 with or without VP4 and triple layered particles VP2/6/7 with or without VP4 (Crawford *et al.,* 1994). Crawford and co-workers also found that VLPs maintained the structural and functional characteristics of rotavirus particles (determined by electron microscope examination of particles), the presence of non-neutralizing and neutralizing epitopes on VP4 and VP7 and hemagglutination activity of VP2/4/6/7 VLPs. The production of VP2/6/4 particles indicated that the VP4 protein interacts with VP6. Cell binding assays performed by Crawford and co-

workers also indicated that the VP4 protein is the viral attachment protein (Crawford *et al.,* 1994).

These are only a few of many studies that have been conducted so far. Insect/baculovirus expression systems do have disadvantages mainly that insect cells do not grow rapidly, the complexity of the growth medium but the biggest, however, is that they are very expensive.

***Bacterial expression system***

Bacterial expression systems are attractive since bacteria have the ability to grow rapidly, reach high cell densities, often have well characterized genetics, are low in costs and have the availability of increasingly large number of cloning and expression vectors (Yin *et al.,* 2007). *Escherichia coli* cells are mainly used due to the fact that the organism is well characterized. However, other bacterial strains (such *as B.subtillus, Lactococcus, and Lactobacillus*) are also available for protein expression. Bacterial expression does, however, have restrictions since bacteria cannot perform post-translational modifications such as glycosylation (Reyes-Ruiz *et al.,* 2006). These modifications can affect the solubility, structure, stability and immunogenicity of the functional protein (Yin *et al.,* 2007).

Rotavirus proteins have been expressed in bacteria as mentioned previously in section 1.9.4.2. Examples of some of these studies include:

**(i) Bacterial expression of major antigenic regions of porcine rotavirus VP7 induces a neutralizing immune response in mice (Wang *et al.,* 1999)**

Wang and co-workers expressed a chimeric protein comprising *E.coli* outer membrane protein A (OmpA) and part of a porcine rotavirus VP7 (containing all three antigenic regions) in *Salmonella* and *E.coli* as an outer membrane associated protein (Wang *et al.,* 1999). When mice were immunized with *E.coli* or *Salmonella* cells expressing the chimeric protein, antibodies were produced against native VP7 (Wang *et al.,* 1999).

**(ii) Development of porcine rotavirus VP6 protein based ELISA for differentiation of this virus and other viruses (Zhu *et al.,* 2013)**

The purpose of this study was to express the viral protein VP6 protein of porcine rotavirus (PRV) in bacteria and to generate rabbit polyclonal antiserum to the VP6 protein (Zhu *et al.,* 2013). They also established a discrimination ELISA to distinguish PRV from a panel of other porcine viruses. Zhu and co-workers found that VP6 could be expressed in *E.coli* and that the anti-VP6 antibody was capable of

distinguishing PRV from porcine transmissible gastroenteritis virus, porcine epidemic diarrhoea virus, porcine circovirus type II, porcine reproductive and respiratory syndrome virus, porcine pseudorabies virus and porcine parvovirus (Zhu *et al.,* 2013). In conclusion, they found that PRV VP6 could be expressed in *E.coli* to generate antibodies in rabbits and that the anti-VP6 serum antibody could be used as a good diagnostic reagent for the detection of PRV.

**(iii) Antibody independent protection against rotavirus infection of mice stimulated by intranasal immunization with chimeric VP4 or VP6 protein (Choi *et al.,* 1999).**

Choi and co-workers administered mice with purified *E.coli* VP4, VP6 or a truncated VP7 (TrVP7) protein fused to the maltose binding protein (MBP). Mice were challenged with EDIM (Epizootic diarrhoea of mice) a month after the last immunisation (Choi *et al.,* 1999). Choi and co-workers found that when mice were administered with the proteins and the adjuvant Saponin adjuvant (QS-21), rotavirus serum IgG was induced by each protein, however, shedding was only significantly reduced in MP:VP6 immunized mice. When mice were administered with the proteins and the adjuvant *E.coli* heat labile toxin LT (R192G), rotavirus serum IgG was also induced but MBP:VP4 immunization stimulated no detectable rotavirus antibody (Choi *et al.,* 1999). No protection against shedding was found in TrVP7 but was found in MBP:VP6 and MBP:VP4.

**(iv) Intranasal administration of an *Escherichia coli* expressed codon-optimised rotavirus VP6 protein induces protection in mice (Choi *et al.,* 2004)**

Choi and co-workers developed a prototype vaccine candidate consisting of MBP:VP6:His6, a chimeric protein of maltose binding protein, VP6 and hexahistidine expressed in *E.coli.* Mice were administered with the MBP:VP6:His6 protein and *E.coli* heat labile toxin LT (R192G) adjuvant (Choi *et al.,* 2004). They found that protection was produced in mice and that the expression of the complete rotavirus VP6 protein was enhanced by codon-optimization and that the protection elicted was not affected by the denaturation of the recombinant VP6 (Choi *et al.,* 2004).

These are only a few studies of rotavirus proteins expressed in bacteria.

***Yeast expression system***

Yeast expression systems are attractive, since yeast cells can grow quickly and in defined media, are easy and less expensive to work with than mammalian or insect cells, they can

be adapted to fermentation and they are also safe since they do not contain any endotoxins or oncogenes (Yin *et al.,* 2007). *Saccharomyces cerevisiae* is the most often used yeast and has been engineered to express different heterologous genes. *Pichia pastoris* and *Kluyveromyces marxianus* are the other two most common yeast systems for heterologous expression (Yin *et al.,* 2007). There are also four other "non-conventional" yeast that have been established for expression namely *Arxula adeninivorans, Hansenula polymorpha, Kluyveromyces lactis* and *Yarrowia lipolytica* (Yin *et al.,* 2007). However, yeast does have drawbacks. The most important drawback is the cell wall if the protein is not to be secreted. Breaking of the cell wall and recovering the interior of the cell intact can be very difficult (Reyes-Ruiz *et al.,* 2006). The typical way to break the cell wall is by using glass beads, but this has limitations on the size of the sample.

There are not many examples of rotavirus proteins expressed in yeast. Rodriguez-Limas and co-workers have shown in two studies that yeast can be used as a platform to express multi-layered rotavirus-like particles and as a possible alternative for veterinary vaccine against rotavirus. In the first study, Rodriguez-Limas and co-workers produced triple-layered particles (VP2, VP6 and VP7) in *Saccharomyces cerevisiae* (Rodriguez-Limas *et al.,* 2011). In the second study, Rodriguez-Limas and co-workers showed that *Saccharomyces cerevisiae* yeast extract containing rotavirus proteins can efficiently induce protection against infection in adult mice models depending of the route of immunisation serum and antibodies used (Rodriguez-Limas *et al.,* 2014).

## 1.11 Motivation and aims of project

### 1.11.1 Motivation of project

Rotavirus infections are extremely problematic despite the success of the commercially available rotavirus vaccines. Rotavirus infection causes deaths of young children worldwide, as discussed above. Currently licensed vaccines have drawbacks such as their efficacy, specificity of strains circulating in specific regions and cost in developing countries.

This project will mainly focus on the expression of the rotavirus structural proteins, VP2 and VP6, in bacteria and yeast. The long term objective for this project is the production of VLPs, specifically double-layered particles (DLPs) and to contribute to the production of cheaper and safer rotavirus-like particle vaccines (VLP vaccine) that can possibly help with the rotavirus infections in developing countries, in particular South Africa. Therefore, the South African GR10924 G9P[6] genotype is used in this study.

It has been shown in the literature that VLPs can been expressed in various expression systems. At the North-West University the rotavirus vaccine development approach is to

produce virus-like particles (VLPs) of local rotavirus strains in insect cells using the baculovirus expression system. However, the baculovirus system is very expensive, whereas bacterial and yeast expression systems are low in cost, have a simple physiology, short generation time and up-scaling can be easily achieved.

### 1.11.2 Aims of project

There are two different types of expression systems that will be used in this study, namely bacterial expression and yeast expression. The aims of this project for each of these systems are given below.

*1) Bacterial expression*

The aim is to express the open reading frames of the bacterial codon optimised rotavirus genome segment 6 (VP6) and 2 (VP2) of a local GR10924 G9P[6] rotavirus strain. The expression plasmid pETDuet-1 will be used.

- To construct the expression vector pETDuet-1_VP2/6 for the simultaneous expression of both structural proteins (VP2 and VP6)
- To verify the co-expression of VP2 and VP6 in bacteria using SDS-PAGE

*2) Yeast expression:*

The aim is to express the open reading frames of the wild type rotavirus genome segment 6 (VP6) and 2 (VP2) of a local GR10924 G9P[6] rotavirus strain. The yeast expression plasmid pKM173 and pKM177 will be used.

- To construct the expression vector pKM173_VP2/6 for the simultaneous expression of both rotavirus structural proteins VP2 and VP6
- To verify the co-expression of VP2 and VP6 in several yeast strains using SDS-PAGE and western blot analysis

# Chapter 2:

# Cloning and bacterial expression of rotavirus structural proteins VP2 and VP6 open reading frames of a South African G9P[6] strain, optimised for expression in bacteria

## 2.1 Introduction

As mentioned previously, bacterial expression systems are attractive because of their ability to grow rapidly and reach high cell densities. In the literature there are reports that showed that rotavirus structural proteins VP2, VP6 and VP7 have been expressed in bacteria, as discussed in Chapter 1 (section 1.10).

In our laboratory the human rotavirus strain RVA/Human-wt/ZAF/GR10924/1999/G9P[6] (G9-P[6]-I2-R2-C2-M2-A2-N2-T2-E2-H2 genotype) was used in the study (Jere *et al.,* 2011). In the current study, the bacterial expression vector pETDuet-1 was used since it allows for the simultaneous expression of the rotavirus structural proteins VP2 (genome segment 2) and VP6 (genome segment 6). To our knowledge this expression plasmid has not yet been used for the expression of rotavirus structural proteins. This study had two main objectives. Firstly, to clone genome segment 2 (VP2) and genome segment 6 (VP6) into the expression plasmid, pETDuet-1. Secondly, to determine whether or not co-expression of the rotavirus structural proteins VP2 and VP6 is possible in bacteria. The bacterial codon optimised genome segment 2 has been expressed before by Dr AC Potgieter (Deltamune (Pty) Ltd R&D, South Africa) but using a different expression vector namely the cold shock vector, pColdTF.

Our long term goal for this study is to determine whether or not it is possible to produce double-layered particles in bacteria. If double-layered particles can form in bacteria it can be used as a booster, together with a live attenuated rotavirus vaccine in developing countries.

## 2.2 Materials and methods

To simplify the flow of the text, the names, suppliers, and catalogue numbers of the reagents used in this study, are included in Appendix A. Therefore, the supplier and catalogue number information is not mentioned in the text but can be found in Appendix A which is organised alphabetically.

### 2.2.1 Rotavirus genome segments, plasmids and bacterial cell lines

The bacterial codon optimised genome segment 2 (VP2) and genome segment 6 (VP6) from the human rotavirus strain GR10924 G9P[6], was used in this study. A stool sample containing the GR10924 G9P[6] strain was obtained from the stool sample collection at the Diarrhoeal Pathogens Research Unit (MRC/DPRU), University of Limpopo, Medunsa, South Africa (Potgieter *et al.,* 2009), dsRNA was extracted and the consensus sequence determined using next generation sequencing (Potgieter *et al.,* 2009). Genome segment 6 (VP6) was optimised for expression in bacteria and purchased from Geneart, which provided genome segment 6 (VP6) in the pGA15 plasmid. The bacterial codon optimised genome segment 2 (VP2) was obtained from Dr AC Potgieter (Deltamune (Pty) Ltd R&D, South Africa) in the pColdTF vector, previously shown to express VP2 in bacteria.

JM109 is an *Escherichia coli (E. coli) strain* that was used for plasmid construction and propagation. JM109 strains are all purpose cloning strains that are ideal for generation of high quality plasmid DNA. JM109 is a K strain bacterium that carries the *recA1 and endA1* mutation that aids in plasmid stability and provides high quality DNA plasmid preparations, respectively.

Origami cells and Origami cells containing pGro7 (provided by Dr Rencia van der Sluis) were used for expression experiments. Origami cells are K-12 derivatives that have mutations in both the thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes, which greatly enhance disulphide bond formation in *Escherichia coli*. Disulphide bond formation is essential for capsid assembly of the virus-like particles. The pGro7 chaperone plasmid is one of five chaperone plasmids commercialized by Takara and contains the groES-groEL chaperone. These chaperone plasmids are developed to enable efficient expression of multiple molecular chaperones known to work in co-operation in the folding process of the expressed proteins. It has been reported that co-expression of a target protein with the chaperone increases recovery of the expressed proteins in the soluble fraction.

### 2.2.2 Cloning vector

The expression vector pETDuet-1 (Novagen) was used in this study (Figure 2.1). The pETDuet™-1 vector is designed for the simultaneous expression of two genes of interest. The vector contains two multiple cloning sites (MCS), each of which is preceded by a T7 promoter/*lac* operator, for high protein expression and a ribosome binding site (rbs). The vector also carries the ColE1 replicon, *lacI* gene and ampicillin resistance gene.

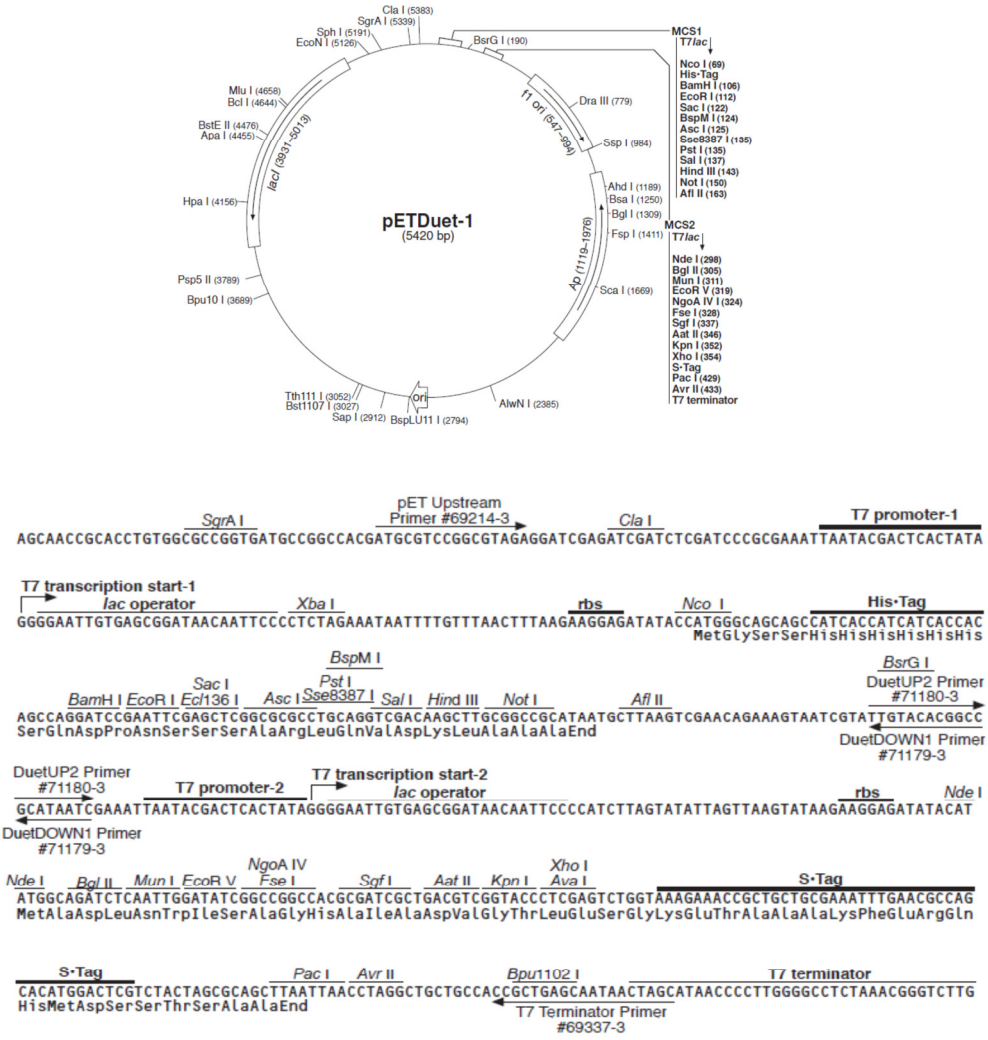


**Figure 2.1: The plasmid map and cloning/expression regions of pETDuet-1**. The plasmid map and cloning/expression region illustrate the properties of pETDuet-1. The vector contains two multiple cloning sites (MCS), each of which is preceded by a T7 promoter/*lac* operator and a ribosome binding site sequence (rbs). The multiple cloning sites each contain several restriction enzymes, with the second multiple cloning site containing a T7 terminator. The vector also carries the *lacI* gene sequence and ampicillin resistance gene. The commercially available primers, used for sequencing for MCS1 and MSC2, are also indicated.

The co-expression of multiple target genes in *E. coli* is advantageous for studying protein complexes. Optimal yield, solubility, and activity are often achieved with co-expression and may protect individual subunits from degradation. The pETDuet vectors carry compatible replicons and antibiotic resistance markers and may be used together in appropriate host strains for effective propagation and maintenance of four plasmids in a single cell that allow for the co-expression of up to eight proteins. The capability of pETDuet vectors to be co-transformed, propagated, and induced for robust target protein co-expression make them ideal for the analysis of protein complexes. The multiple cloning regions have restriction sites that facilitate the cloning of two genes and the transfer from other Novagen pET constructs. The pETDuet vectors provide the option of producing native unfused proteins, or fusions to His•Tag® and S•Tag™ sequences for detection and purification of protein complexes.

### 2.2.3 DNA recombinant techniques

Recombinant DNA techniques used in this study, namely polymerase chain reaction (PCR), restriction endonuclease digestion, ligation and transformation were performed according to Sambrook and Russell (2001). Commercial kits were used for plasmid isolation, gel extraction and PCR clean-up. The recombinant techniques used will be discussed below.

#### 2.2.3.1 Plasmid isolation

Two types of plasmid isolation methods were used namely the Promega Pureyield midi-prep kit (Promega) and the Birnboim and Doly alkaline lysis mini-prep method (Birnboim and Doly, 1979).

##### 2.2.3.1.1 Promega Pureyield™ plasmid midi-prep system

Propagation of bacteria containing the plasmid of interest was inoculated into 50 ml LB broth (10 mg/ml tryptone, 5 mg/ml yeast extract and 10 mg/ml NaCl), containing the appropriate antibiotics (50 mg/ml ampicillin) and grown for 16 hours at 37°C, while shaking at 180 rpm. The next day these cultures, containing the appropriate plasmid, were centrifuged at 6 000 x g for 15 minutes and the supernatants were discarded.

The pellets of the cultures were used to isolate plasmids using the Promega Pureyield™ plasmid midi-prep kit. The isolation was done according to the instructions of the manufacturer. This kit is designed to isolate high-quality plasmid DNA and uses a silica membrane column to isolate plasmid DNA. This system also incorporates a unique

endotoxin removable wash that reduces contaminants such as endotoxins, RNA and endonuclease. Removal of these contaminants aids in the efficiency of transformation.

This procedure consists of four steps namely i) preparation and lysis of bacterial cell culture, ii) DNA purification iii) wash and iv) elution. The alkaline conditions denature the chromosomal DNA and proteins. The lysis time ensures the optimum release of plasmid DNA without the release of chromosomal DNA and without exposing the plasmid DNA to denaturing conditions for too long. The lysate is then neutralized which causes denatured components to precipitate while only small plasmid DNA renaturates and stay in the solution. The plasmid DNA binds to the membrane of the binding column but contaminants such as endotoxin, protein, RNA and endonucleases do not bind and pass through the column. The membrane is then washed with the washing buffer, followed by the plasmid DNA being eluted in 400 ųl nuclease free water.

#### 2.2.3.1.2 Mini plasmid preparation

Mini-plasmid DNA preparation was prepared using the rapid alkaline extraction procedure of Birnboim and Doly (1979). The principle of this method is a selective alkaline denaturation method of high molecular weight chromosomal DNA while closed circular DNA remains double-stranded. The chromosomal DNA re-natures upon neutralization to form an insoluble clot, leaving the plasmid DNA in the supernatant (Birnboim and Doly, 1979). Ten millilitres of LB cultures containing the appropriate antibiotics were inoculated with the desired clone and incubated at 37°C overnight with shaking. The cells were harvested by centrifugation at 13 000 x g for 30 seconds.

The pellet was re-suspended by using 100 µl solution I, the lysozyme solution (containing 2 mg/ml lysozyme, 50 mM glucose, 10 mM EDTA and 25 mM Tris-HCl, pH 8.0). Plasmids containing cells are treated with lysozyme to weaken the cell wall. EDTA chelates divalent cations in the solution preventing DNases from damaging the plasmid and also helps by destabilizing the cell wall. Glucose maintains the osmotic pressure so the cells do not burst. The next step was to add 150 µl solution II; the alkaline SDS solution (0.2 N sodium hydroxide (NaOH) and 1% sodium dodecyl (lauryl) sulphate (SDS)) to lyse the cell wall completely. The chromosomal DNA is selectively denatured and when the lysate is neutralized by acidic sodium acetate the mass of chromosomal DNA renatures and aggregates to form an insoluble network. Finally 150 µl solution III, high salt solution, was

added (3 M potassium acetate) to cause precipitation of the protein-SDS complexes and high molecular weight RNA.

The microfuge tube, containing the DNA of interest, was left at 0°C for 60 minutes to allow most of the protein, high molecular weight RNA and chromosomal DNA to precipitate. The sample was centrifuged (10 000 x g for 5 minutes) and 1 ml ethanol was added and the microfuge tube left at -20°C for 30 min. The precipitant was collected by centrifugation (10 000 x g for 2 minutes) and the supernatant removed by aspiration. The pellet was dissolved in 100 µl of 0.1 M sodium acetate/0.05 M Tris-HCl (pH 8.0) and re-precipitated with 2 volumes of cold ethanol. The tube was left at -20°C for 10 minutes and the precipitated collected again by centrifugation. The pellet was dissolved in 40 µl nuclease free water. The plasmid preparations were analysed as described in section 2.2.3.4.

#### 2.2.3.2 PCR amplification of the coding sequences

PCR is a technique to amplify the amount of a specific DNA segment. A preparation of denatured DNA that contains the segment of interest serves as template for DNA polymerase and two specific oligonucleotides serves as primers for DNA synthesis. They prime the DNA polymerase-catalyzed synthesis of the two complementary strands of the desired segment, effectively doubling its concentration in the solution (Sambrook and Russell, 2001). The DNA is then heated to dissociate the DNA duplexes and then cooled so that the primers bind to both the old and newly formed strands. These steps form one cycle which is repeated several times, usually between 20 - 30 times. The protocol has been automated by the invention of the thermal cyclers that alternately heat the reaction mixture to 95ºC to dissociate the DNA, followed by cooling, annealing of primers and another round of DNA synthesis (Sambrook and Russell, 2001).

The isolation of the thermostable DNA polymerase from the thermophillic bacterium *Thermus aquaticus* has made it unnecessary to add fresh enzyme for each round of synthesis (Sambrook and Russell, 2001). However, TaKaRa Ex Taq™ provides more efficient amplification and higher fidelity than conventional Taq DNA polymerase under conventional PCR conditions, since TaKaRa Ex Taq has a lower mutation rate than standard Taq.

Amplification was performed with the use of a BioRad™ thermocycler. The coding regions were amplified in reaction mixtures containing 0.5 µM of both the forward and reverse primer (Table 2.1) and 0.5 µl template (57 ng), 5 µl 10 x Takara Ex-Taq buffer (1X), Takara Ex-Taq

(1.25 units/50µl), 2.5 mM dNTP and nuclease free water added to a final volume of 50 µl. The negative controls contained no template.

**Table 2.1: Oligonucleotide primers used in this study**

| Primer name | Oligonucleotide sequence (5' → 3') | $T_m$ (°C) | Length (bp) |
|---|---|---|---|
| VP6bacpETD_F | CGA G**CC ATG G**AC GTG CTG TAT AG | 59.5°C | 23 |
| VP6bactpETD_R | CTC C**GA ATT C**TT ACT ATT TCA CCA GC | 55.6°C | 26 |
| VP2bacpETD_F | CCA G**AG ATC T**AT GGC GTA TCG | 57.7°C | 21 |
| VP2bactpETD_R | CTC T**CC TAG G**TT ACT ACA GTT CGT TCA TG | 54.4°C | 29 |

* All primers were synthesized by Integrated DNA technologies (IDT)
* Restriction enzyme sites, in primers, used for amplification is indicated in bold

The samples were denaturated at 98°C for 1.5 minutes followed by 30 cycles of amplification, unless otherwise stated. Each cycle consisted of 15 seconds denaturation at 95°C, 30 seconds of annealing (temperature depends on annealing temperature of primers, indicated in Table 2.2) and 2 minutes of extension at 72°C. Amplification was followed by one elongation cycle at 72°C for 10 minutes. PCR products were analysed by means of gel electrophoresis using a 1% agarose gel.

**Table 2.2: Annealing temperatures of different primer sets**

| Genome segment amplified | Primers used | Annealing $T_m$ (°C) |
|---|---|---|
| Genome segment 2 Bacterial optimised | VP2bacpETD_F and VP2bactpETD_R | 57°C |
| Genome segment 6 Bacterial optimised | VP6bacpETD_F and VP6bactpETD_R | 50°C |

#### 2.2.3.3 Purification of PCR amplicons

After PCR amplification the amplicons were purified by means of PCR clean-up, using the commercial kit NucleoSpin Extract II (Clonetech Laboratories, Inc). Purification was done according to the instructions of the manufacturer with a few modifications. With the NucleoSpin Extract II method, DNA binds in the presence of chaoptropic salts to a silica membrane. The binding mixture that contains the NT buffer, (5 mM Tris-Cl, pH 8.5) and DNA is loaded directly onto the NucleoSpin Extract II columns. Contaminations such as salts and soluble macromolecular components are removed by a washing step using ethanolic buffer NT3 and the pure DNA eluted in nuclease free water.

The NucleoSpin Extract II buffer ensures complete removal of primers from PCR reactions while small DNA fragments are still bound and purified with high yields. The procedure consists of five steps i) adjustment of the DNA binding conditions by means of the chaotropic salts in the NT buffer, ii) binding of the DNA, iii) washing of the silica membrane to remove salts and macromolecular components, iv) drying of the silica membrane and v) eluting DNA. The following modifications were made: the PCR amplicons were pooled and were eluted in nuclease free water instead of elution buffer NE.

#### 2.2.3.4 Agarose gel electrophoresis

Agarose gel electrophoresis was used for analysis of PCR reactions, restriction digestions and gel extraction of the desired amplicons. Procedures were carried out as described in the literature (Sambrook & Russell, 2001). Unless otherwise stated, a 1% agarose gel was prepared using 1 x TAE buffer (40 mM Tris-acetate pH 8.5, 1 mM EDTA). Ethidium bromide was added to a final concentration of 0.5 µg/ml, to facilitate visualisation of DNA on an ultraviolet light trans-illuminator.

Gels were loaded with samples mixed with one quarter volume of Fermentas orange loading dye. A DNA molecular size marker (O'GeneRuler$^{TM}$ DNA ladder mix) was always loaded in the first lane for the estimation of DNA sizes. All gels were electrophoresed for one hour at 80V using a Bio-Rad PowerPac system. Directly after electrophoresis gels were captured using the Syngene ChemiGenius Bio-imaging system and GeneSnap software.

#### 2.2.3.5 Analysis of DNA concentration and purity

DNA concentration and purity were determined spectrophotometricaly using a NanoDrop ND-1000 system (NanoDrop Technologies, Inc). One absorbance unit at 260nm corresponds to 50 ng/µl double-stranded DNA. Measuring the absorbance at 260nm the DNA concentrations of the sample were calculated. Measurement of DNA purity was calculated based on the $A_{260/280}$ ratio, since proteins absorb light at 280nm. The device was always blanked with the solution in which the DNA was solubilised, in this case nuclease free water.

### 2.2.3.6 Restriction endonuclease digestions

Restriction enzyme digestion reactions were performed in reaction volumes of 30 µl, unless otherwise stated. Each reaction mixture consisted of the applicable restriction enzymes and buffer (Table 2.3), approximately 0.5-1 µg/µl DNA and nuclease free water to a final volume of 30 µl. Reaction mixtures were incubated at 37°C for approximately 1-16 hours. The digested samples were analysed with agarose gel electrophoreses (section 2.2.3.4) and purified (section 2.2.3.7). Enzymes and buffers used in the reactions are indicated in Table 2.3.

**Table 2.3: Restriction enzymes and buffers used in this study**

| Genome segment | Restriction enzyme | Recognition sequence | Buffer used (Fermentas) | Conditions for double digestion |
|---|---|---|---|---|
| Bacterial optimised Genome segment 6 | NcoI EcoRI | C↓CATGG and G↓AATTC | 2 X Tango buffer | 5 units/µl NcoI 5 units/µl EcoRI |
| Bacterial optimised Genome segment 2 | BglII XmaJI | A↓GATCT and C↓CTAGG | 1 X Buffer O | 5 units/µl BglII 10 units/µl XmaJI |

### 2.2.3.7 Gel purification of desired DNA fragments or products

Following restriction enzyme digestion of the purified PCR amplicons and vectors, for cloning purposes, the DNA fragments were separated on a 0.8% agarose gel. The restriction digested PCR amplicon and vector was excised from the agarose gels using a sterile stainless-steel blade, weighed, transferred into a clean tube and purified with a Nucleospin Extract II kit, according to the manufacturer's protocol, with slight modifications.

The NucleoSpin extract II kit uses the spin column technology and selective binding properties of a silica membrane. The NT buffer (5 mM Tris-Cl, pH 8.5) is added to the excised gel, which facilitates with the dissolvement of agarose at 50°C. The solution is passed through a spin column with a silica membrane. The DNA binds in the presence of chaotropic salts (present in NT buffer) to the silica membrane. Contaminations like salts, ethidium bromide and soluble macromolecular components are removed by a simple washing step with ethanolic buffer NT3. The pure DNA was eluted in nuclease free water. Modifications that were made: the plasmid DNA was eluted in nuclease free water instead of the elution buffer NE.

### 2.2.3.8 Ligation reactions

The purified restriction digested genome segments and plasmids were combined in a 3:1 molar ratio, unless otherwise stated. The amount of DNA was calculated using the following formula:

$$\frac{150 \text{ ng vector x size of insert}}{\text{Size of vector}} \quad x \quad \frac{3}{1} = \text{ng insert}$$

Ligation reactions were performed in reaction volume of 30 µl. The mixture consists of 5 Weiss u/µl T4 DNA ligase (Fermentas), 1 x T4 DNA ligase buffer (Fermentas), the applicable amount of insert and plasmid and nuclease free water to a final volume of 30 µl. The reaction mixture was incubated at 4°C overnight, unless otherwise stated.

### 2.2.3.9 Preparation of chemical competent *Escherichia coli* cells

Nucleic acids do not enter bacteria freely, but require assistance traversing the outer and inner cell membranes and in reaching an intracellular site where they can be expressed and replicated (Sambrook and Russell, 2001). In this study, chemical competent cells were used. Transformation of chemically competent cells is achieved by suspending the cells and the target DNA in an ice-cold calcium chloride based buffer. Under these cold conditions, the $Ca^{2+}$ ions create pores in the membrane that assists the binding of the DNA to the cell membrane and mask the negative charge on the DNA, easing the DNA passage through the hydrophobic cell membrane. The transformation occurs when the DNA is forced into the cells by applying a short 42°C heat shock, which results in a thermal current that sweeps the DNA into the cells.

All chemical competent cells used in this study were prepared in the laboratory, using the Inoue method (Sambrook and Russell, 2001). The Inoue method differs from other procedures in that the bacterial culture is grown at 18ºC rather than the conventional 37°C and it also allows one to store the cells for more than one year at -80°C. Glycerol stocks of JM109, Origami and Origami pGro7 were streaked out on SOB (2 g/ml tryptone, 0.05 g/ml yeast extract, 0.5 g/ml NaCl, 250 mM KCl, 2 M $MgCl_2$) agar plates, containing no antibiotics and grown for 16 hours at 37°C. One 25 ml culture of SOB medium was inoculated with a single colony of each of the respective cell lines (known as the starter culture). The starter culture was grown for 6 - 8 hours at 37°C with vigorous shaking (250-300rpm). A 250 ml flask containing an amount of the starter culture (2 ml) and new SOB medium was incubated overnight at 18°C, while shaking. The optical density was read at 600nm. The culture was

grown untill the optical density was 0.55 at 600nm. All subsequent steps were performed at 4°C. The culture was then transferred to an ice-water bath for 10 minutes, after which the cells were harvested by centrifugation at 2500 x g for 10 minutes. The supernatant was discarded and the cell pellets re-suspended in 80 ml of ice cold Inoue transformation buffer (55 mM $MnCl_2 \cdot 4H_2O$, 15 mM $CaCl_2 \cdot 2H_2O$, 250 mM KCl and 10 mM PIPES). The centrifugation step was repeated and the cell pellets were re-suspended in 20 ml ice-cold Inoue transformation buffer and 1.5 ml DMSO was added followed by incubation on ice for 10 minutes. The suspension was dispensed in 50 µl aliquots into pre-chilled, sterile microfuge tubes. The cells were immediately snap-frozen in a bath of liquid nitrogen and stored at -80°C.

#### 2.2.3.10 Transformation of chemical competent *Escherichia coli* cells

Transformation reaction consisted of 25 ng transforming DNA per 50 µl competent cells (section 2.2.3.8). Frozen chemical competent cells were removed from storage at -80°C and thawed on ice. The ligation reaction was added to the competent cells and gently mixed. The cell mixture was incubated on ice for 30 minutes after which it was transferred to 42°C and heat shocked for 90 seconds. The cells were immediately cooled on ice for 2 minutes after which 800 µl of recovery medium (usually SOC medium - 2 g/ml tryptone, 0.05 g/ml yeast extract, 0.5 g/ml NaCl, 250 mM KCl, 2M $MgCl_2$, 1M Glucose) was added (to maximize the transformation efficiency of the competent cells). The cells were allowed to recover at 37°C with gentle shaking at 180rpm, for 1 hour. Usually, 200 µl of the cell mixture was spread out on super optimal broth (15g agar/1L SOB medium) agar plates, containing 20 mM $MgSO_4$ and the appropriate antibiotics. The plates were left for 15 minutes at room temperature (18°C -23°C) to absorb the liquid and then incubated upside-down at 37°C for 16 hours.

#### 2.2.3.11 Long term storage of bacterial colonies

Bacterial colonies of interest were prepared for long term storage by adding glycerol to a final concentration of 15%. This was done by combining, in a microfuge tube, 817 µl of the culture and 183 µl of 80% glycerol to obtain a final volume of 1 ml. The stocks were then stored at -80°C.

#### 2.2.3.12 DNA sequence determination

DNA sequencing was carried out to confirm that the recombinant plasmid, containing the ORF of genome segment 2 and genome segment 6 does not contain any mutations and that the cloning was successful. Sanger sequencing was used. Samples and appropriate sequencing primers (Table 2.4) were sent to the DNA sequencing facility laboratory of the Central Analytical Facility of the University of Stellenbosch. DNA sequence electropherograms were analysed using FinchTV version 1.40 and DNAman was used to align sequences to reference sequences.

An *in silico* sequence of every construct was designed with DNAMAN Version 6 (Lynnon Corporation, 2005), prior to cloning of the construct, to facilitate the screening of mutations etc. occurring in the construct. When the sequence data was received from the service provider, the sequences were analysed and edited with FinchTV and the assembly of the construct was confirmed by aligning it with the *in silico* sequence, using DNAMAN. The commercially available primers used for sequencing made it possible to see regions of importance namely the restriction enzymes used for cloning, the start and stop codons of the ORF encoding VP2 or VP6, the T7 promoter, T7 transcription start, the *lac* operator as well as the ribosome binding sequence.

**Table 2.4: Oligonucleotide primers used for sequencing**

| Primer name | Oligonucleotide sequence (5' → 3') | $T_m$ (°C) | Length (bp) |
|---|---|---|---|
| pET Upstream primer | ATG CGT CCG GCG TAG A | 59°C | 16 |
| DuetDOWN1 primer | GAT TAT GCG GCC GTG TAC AA | 60°C | 20 |
| DuetUP2 primer | TTG TAC ACG GCC GCA TAA TC | 60°C | 20 |
| T7 terminator primer | CCG CTG AGC AAT AAC TAG C | 56°C | 19 |
| InternalVP2(Bact)_F primer | CGA TTG CTT TAA AAC CC | 45.9°C | 17 |

*All primers were synthesized by Integrated DNA technologies (IDT)

### 2.2.4 Expression of rotavirus VP2 (genome segment 2) and VP6 (genome segment 6)

The general expression protocol was followed according to the pET system manual (Novagen). The positive control (VP2/6 lysate) used for the expression experiments was produced in insect cells and provided by Dr Khuzwayo Jere from our laboratory. The desired strain (which contains the recombinant plasmid for expression) was used to inoculate 2 ml LB Broth medium (10 mg/ml tryptone, 5 mg/ml yeast extract and 10 mg/ml NaCl), containing the appropriate antibiotics for the plasmid and host strain. The culture was grown at 37°C

with shaking at 180rpm until the $OD_{600}$ reached a density between 0.4-1 (0.6 recommended). The cultures were immediately split into two cultures: one contained 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) to induce the induction of the proteins expression (induced culture) and the other culture received no IPTG (uninduced culture). Incubation was continued, for both cultures, for another 2-3 hours.

### 2.2.5 Cell lysis using BugBuster protein extraction reagent

Cells were placed on ice for 5 minutes after which the cells were harvested by centrifugation, 5 000 x g for 5 minutes. The supernatant was discarded and the cells weighed. Novagen BugBuster protein extraction buffer and Lysonase were added according to the instruction of the manufacturer (5 ml Bugbuster per gram wet cell paste and 10 µl Lysonase per gram wet cell paste). The cells were resuspended by gentle vortexing. At this point 500 µl of each sample was aliquoted into a 1.5 ml microfuge tube which represented the total protein fraction. The insoluble material was then removed by centrifugation, 13 000 x g for 15 minutes. The supernatant was aliquoted into a 1.5 ml microfuge tube which presented the soluble protein fraction.

### 2.2.6 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SDS-PAGE was used for routine analysis of protein expression and purification procedures (Sambrook and Russell, 2001). This technique separates proteins according to their electrophoretic mobility (a function of the length of a polypeptide chain and its charge). SDS is an anionic detergent that linearizes the proteins and imparts a negative charge to linearized proteins. The binding of SDS to the polypeptide chain imparts an even distribution of charge per unit mass, thereby resulting in fractionation by approximate size during electrophoresis.

Separating gels had a final concentration of 15% acrylamide, unless otherwise stated. The composition of the separating gels was 30% acrylamide, 0.8% bis-acrylamide, 376 mM Tris-Cl (pH 8.8), 0.1% SDS, 0.1% APS (ammoniumperoxsidesulphate) and 0.008% TEMED (Tetramethylethylenediamine). The separating gel was prepared by mixing the components in an Erlenmeyer flask before the addition of the APS and TEMED. The seperating gel was then poured into a Bio-Rad gel casting apparatus. The seperating gel was left for polymerization for at least 1 hour at room temperature after isopropanol was added to the surface of the gel. The isopropanol was thoroughly removed using filter paper and the

stacking gel was added. The stacking gel consisted of 4% acrylamide, 0.1% bisacrylamide, 372 mM Tris-Cl (pH 6.8), 0.1% SDS, 0.1% APS and 0.008% TEMED. The stacking gel was polymerized in the presence of a 10-well comb, at room temperature.

Protein samples were prepared by combining 10 µl sample with 10 µl 4 x Laemmli protein loading buffer, 18 µl of water and 2 µl 20 x reducing agent. The samples were then mixed and boiled for 5 minutes at 98°C. Unless otherwise stated, 30 µl of the 40 µl mixture was loaded onto the gel. The PageRuler$^{TM}$ protein molecular size marker mixture was loaded in one lane. The loaded gel was electrophoresed in 1 x Tris Glycine SDS (25 mM Tris, 2 M Glycine and 0.1% SDS) buffer at a current of 35mA using a Bio-Rad PowerPac Basic system. Electrophoresis was carried out for about 40 minutes or until the pink dye front was at the bottom of the gel. The gels were removed from the glass plates and stained with Coomassie brilliant blue solution (0.4% (w/v) Coomassie Brilliant Blue R-250, 45% (v/v), Methanol, 10% (v/v) Glacial acetic acid), while shaking at 60rpm for 1 hour. After this incubation the gels were removed from the staining solution and destained, for 1 hour, with methanol-acetic acid gel solution (45% (v/v) Methanol and 10% (v/v) Glacial acetic acid). The gel was gently stirred with occasionally exchanging the destain solution until the gels were fully destained. The gel was documented by scanning (HP digital document scanner) and subsequently dried for 1 hour at 60°C.

### 2.3 Results and discussion

As mentioned previously this chapter focused on the cloning into the pETDuet vector and expression of the open reading frames (ORF) of the bacterial codon-optimised VP2 (genome segment 2) and VP6 (genome segment 6) of the South African GR10924 G9P[6] rotavirus strain. The cloning strategy used to generate the recombinant plasmids (pETDuet-1_VP2, pETDuet-1_VP6 and pETDuet-1_VP2/6) for expression is indicated in Figure 2.2.

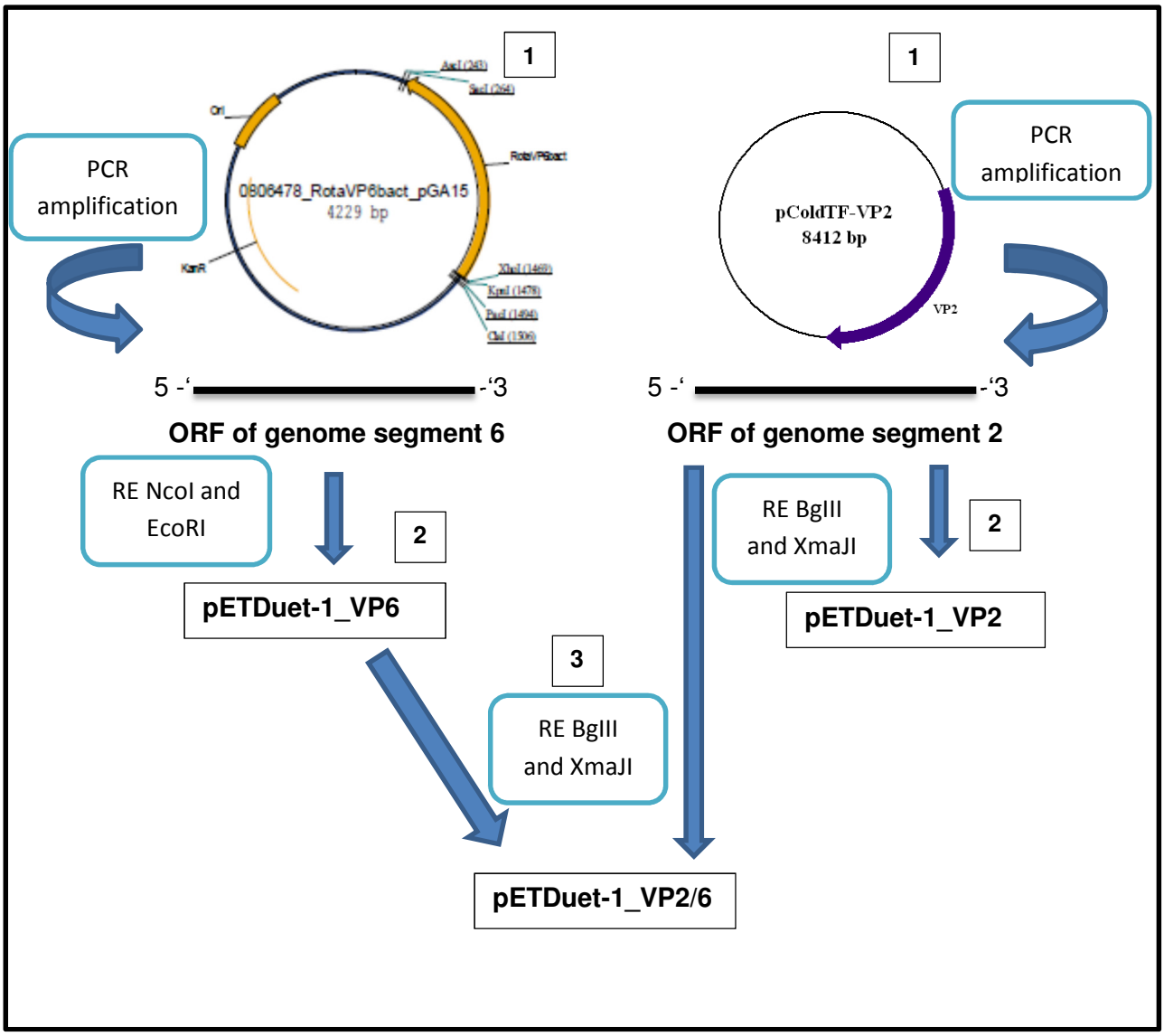


**Figure 2.2: A schematic illustration of the cloning strategy to generate recombinant plasmids for expression in bacteria.**

(1) Polymerase chain reaction was used to amplify the ORF region encoding VP6 (genome segment 6) and VP2 (genome segment 2) from 0806478_RotaVP6bact_pGA15 and pColdTF-VP2, respectively. (2) Thereafter, the coding regions of genome segment 2 and genome segment 6, as well as the expression plasmid, pETDuet-1, were digested with the appropriate restriction enzymes (genome segment 2: BglII and XmaJI; genome segment 6: EcoRI and NcoI). Ligation reactions were carried out which yielded two constructs for expression experiments namely, pETDuet-1_VP2, pETDuet-1_VP6. (3) A third restriction enzyme digestion was carried out on both the PCR amplified ORF of genome segment 2 from the pColdTF_VP2 vector and pETDuet-1_VP6 using the restriction enzymes BglII and XmaJI. The ligation reaction yielded the construct pETDuet-1_VP2/6.

### 2.3.1 Cloning of the open reading frame encoding VP6 (genome segment 6) into pETDuet-1

The bacterial codon optimised genome segment 6 was purchased from Geneart, which provided the gene in the plasmid pGA15VP6bact.

In order to clone the open reading frame encoding VP6 into the expression vector pETDuet-1 (Figure 2.1), it was PCR amplified from pGA15VP6bact using primers that were designed to contain the restriction enzyme sites, NcoI and EcoRI, to facilitate directional cloning (Figure 2.3). The bacterial codon optimised ORF of genome segment 6 did not contain these restriction enzyme sites. The enzyme sites are located in multiple cloning site 1 (MCS1) of the expression plasmid, pETDuet-1. The primers were designed using the bacterial codon optimised ORF sequence of genome segment 6 (VP6) (Figure 2.3).

Restriction enzyme ***NcoI***

Forward primer 5′CGAG**CCATGG**

```
1       CGAATTGGGG ATCCACTACT AGTTAATTAA TTAACGCCAT GTTAAGGTAC CTCGAGATGG
        GCTTAACCCC TAGGTGATGA TCAATTAATT AATTGCGGTA CAATTCCATG GAGCTCTACC

        ACGTGCTGTA TAG'3
61      ATGTGCTGTA TAGCCTGAGC AAAACCCTGA AAGATGCGCG TGATAAAATT GTGGAAGGCA
        TACACGACAT ATCGGACTCG TTTTGGGACT TTCTACGCGC ACTATTTTAA CACCTTCCGT

121     CCCTGTATAG CAACGTGAGC GATCTGATTC AGCAGTTTAA CCAGATGATC ATTACCATGA
        GGGACATATC GTTGCACTCG CTAGACTAAG TCGTCAAATT GGTCTACTAG TAATGGTACT

181     ACGGCAACGA ATTTCAGACC GGCGGCATTG GCAACCTGCC GATTCGTAAC TGGAACTTTG
        TGCCGTTGCT TAAAGTCTGG CCGCCGTAAC CGTTGGACGG CTAAGCATTG ACCTTGAAAC

241     ATTTTGGCCT GCTGGGCACC ACGCTGCTGA ACCTGGATGC GAACTATGTG GAAACCGCGC
        TAAAACCGGA CGACCCGTGG TGCGACGACT TGGACCTACG CTTGATACAC CTTTGGCGCG

301     GTAACACCAT TGATTACTTC GTGGATTTCG TTGATAACGT GTGCATGGAT GAAATGGTGC
        CATTGTGGTA ACTAATGAAG CACCTAAAGC AACTATTGCA CACGTACCTA CTTTACCACG

361     GTGAAAGCCA GCGTAACGGC ATCGCGCCGC AGAGCGATAG CCTGCGTAAA CTGAGCGGCA
        CACTTTCGGT CGCATTGCCG TAGCGCGGCG TCTCGCTATC GGACGCATTT GACTCGCCGT

421     TTAAATTCAA ACGCATCAAC TTCGATAACA GCAGCGAATA TATCGAAAAC TGGAACCTGC
        AATTTAAGTT TGCGTAGTTG AAGCTATTGT CGTCGCTTAT ATAGCTTTTG ACCTTGGACG

481     AGAACCGTCG TCAGCGTACC GGCTTTACCT TTCATAAACC GAACATCTTT CCGTATAGCG
        TCTTGGCAGC AGTCGCATGG CCGAAATGGA AAGTATTTGG CTTGTAGAAA GGCATATCGC

541     CGAGCTTTAC CCTGAACCGT AGCCAGCCGG CGCATGATAA CCTGATGGGC ACCATGTGGC
        GCTCGAAATG GGACTTGGCA TCGGTCGGCC GCGTACTATT GGACTACCCG TGGTACACCG

601     TGAACGCGGG CAGCGAAATT CAGGTGGCGG GCTTTGATTA TAGCTGCGCG ATTAACGCGC
        ACTTGCGCCC GTCGCTTTAA GTCCACCGCC CGAAACTAAT ATCGACGCGC TAATTGCGCG

661     CGGCGAACAC CCAGCAGTTT GAACATATTG TGCAGCTGCG TCGTGTGCTG ACCACCGCGA
        GCCGCTTGTG GGTCGTCAAA CTTGTATAAC ACGTCGACGC AGCACACGAC TGGTGGCGCT

721     CCATTACCCT GCTGCCGGAT GCGGAACGTT TTAGCTTTCC GCGTGTGATT AACAGCGCGG
        GGTAATGGGA CGACGGCCTA CGCCTTGCAA AATCGAAAGG CGCACACTAA TTGTCGCGCC

781     ATGGCGCGAC CACCTGGTAT TTTAATCCGG TGATTCTGCG TCCGAACAAC GTGGAAGTGG
        TACCGCGCTG GTGGACCATA AAATTAGGCC ACTAAGACGC AGGCTTGTTG CACCTTCACC

841     AATTTCTGCT GAACGGCCAG ATTATTAACA CCTATCAGGC GCGTTTTGGC ACCATTGTGG
        TTAAAGACGA CTTGCCGGTC TAATAATTGT GGATAGTCCG CGCAAAACCG TGGTAACACC

901     CGCGTAACTT TGATACCATT CGCCTGAGCT TTCAGCTGAT GCGTCCGCCG AACATGACCC
        GCGCATTGAA ACTATGGTAA GCGGACTCGA AAGTCGACTA CGCAGGCGGC TTGTACTGGG

961     CGAGCGTGGC GGCGCTGTTT CCGAACGCGC AGCCGTTTGA ACATCATGCG ACCGTTGGCC
        GCTCGCACCG CCGCGACAAA GGCTTGCGCG TCGGCAAACT TGTAGTACGC TGGCAACCGG

1021    TGACCCTGAA AATTGAAAGC GCGGTGTGCG AAAGCGTTCT GGCCGATGCG AGCGAAACCA
        ACTGGGACTT TTAACTTTCG CGCCACACGC TTTCGCAAGA CCGGCTACGC TCGCTTTGGT

1081    TGCTGGCCAA CGTGACCAGC GTGCGTCAGG AATATGCGAT TCCGGTTGGC CCGGTGTTTC
        ACGACCGGTT GCACTGGTCG CACGCAGTCC TTATACGCTA AGGCCAACCG GGCCACAAAG
```

```
1141    CGCCGGGTAT GAACTGGACC GATCTGATTA CCAACTATAG CCCGAGCCGT GAAGATAACC
        GCGGCCCATA CTTGACCTGG CTAGACTAAT GGTTGATATC GGGCTCGGCA CTTCTATTGG

1201    TGCAGCGTGT GTTTACCGTG GCGAGCATTC GTAGCATGCT GGTGAAATAA TAACATATGG
        ACGTCGCACA CAAATGGCAC CGCTCGTAAG CATCGTACGA CCACTTTATT ATTGTATACC
                                  Reverse primer 3'-CCACTTTATC ATTCTTAAGCC
                                                     Restriction enzyme   EcoRI
1261    AGCTCTTACA TCGCTGGCGC GCCCTAGTGG CGTAATCATG GTCATAGCTG TTTCCTGT
        TCGAGAATGT AGCGACCGCG CGGGATCACC GCATTAGTAC CAGTATCGAC AAAGGACA
        TC-5'
```

**Figure 2.3: Sequence of the ORF encoding VP6 (GR10924 G9P[6]) indicating the primers used for amplification.** The forward and reverse primers are indicated in purple and the restriction enzyme sites in black. The start codon is indicated in green and the stop codon in red.

In order to obtain an amplicon of the ORF encoding VP6, optimisation of PCR annealing temperature had to be done. A temperature gradient ranging from 50°C - 60°C, was used. A 5 µl aliquot of the annealing temperature optimisation reaction mixture was analysed using a 1% agarose gel, as shown in Figure 2.4

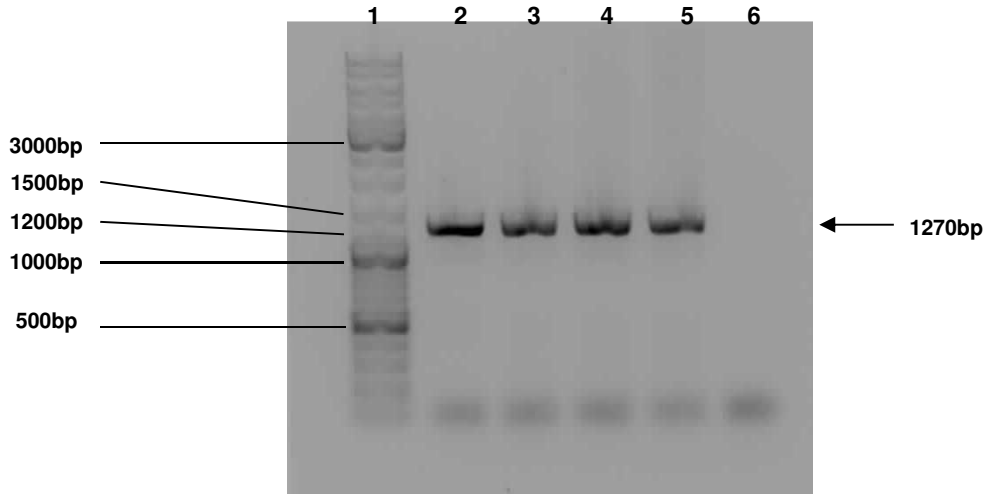


**Figure 2.4: Analysis by agarose gel electrophoresis of the PCR amplification gradient of the ORF encoding VP6.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl volume of each 50 µl PCR reaction was loaded as follows; lanes: 2) genome segment 6 amplification reaction at 50°C; 3) genome segment 6 amplification reaction at 55°C; 4) genome segment 6 amplification reaction at 57°C; 5) genome segment 6 amplification reaction at 60°C; 6) no template control.

The amplification of the ORF encoding VP6 from pGAVP6bact using the temperature gradient was successful as seen in Figure 2.4. Lanes 2 to 5 show a 1270bp amplicon as expected. The lack of a band in lane 6 which serves as the negative control (the negative control has no template DNA and is usually water instead of DNA), indicates that no contamination took place.

Another four PCR amplifications reaction was done as described in section 2.2.3.2, with an annealing temperature of 50°C, since the band at 50°C was the clearest during the PCR optimisation experiment (Figure 2.4). The PCR amplification was repeated and the PCR amplification reactions pooled and purified from oligonucleotides and other contaminations by means of PCR clean-up to ensure good quality DNA for cloning. This yielded approximately 30 ng/µl product of the ORF encoding VP6, which was stored at 4°C, for cloning purposes. A 5 µl sample of the purified ORF encoding VP6 was analysed by a 1% agarose gel electrophoresis, as shown in Figure 2.5.

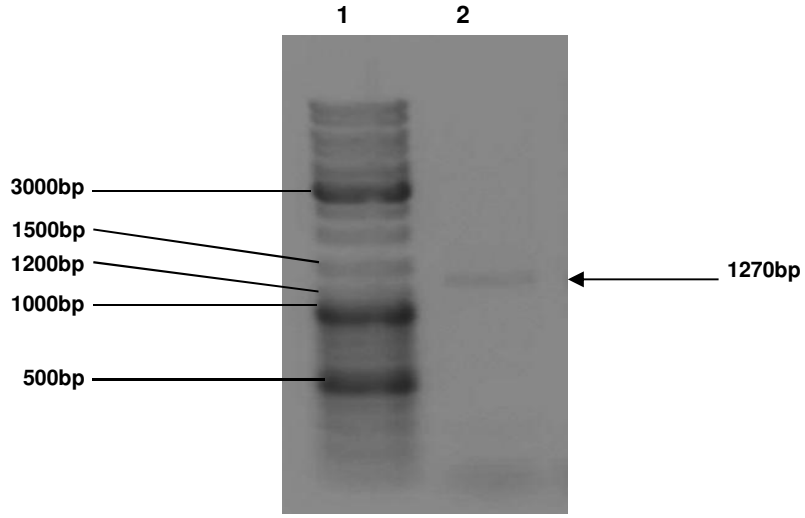


**Figure 2.5: Analysis by agarose gel electrophoresis following a PCR clean-up procedure of the amplified ORF encoding VP6.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl reaction of the 50 µl PCR clean-up reaction was loaded as follows; lanes: 2) ORF of genome segment 6.

Double restriction endonuclease digestion was done with NcoI and EcoRI, as described in section 2.2.3.6, for the amplified ORF encoding VP6 and pETDuet-1. These digests were analysed by loading a 5 µl sample of each reaction on a 1% agarose gel, as shown in Figure 2.6 (Only the results of pETDuet-1 is shown and not the results of the ORF encoding for

VP6, due to low yield of the ORF of genome segment 6). The expected results, for the restriction digestion were to see two bands for pETDuet-1 (5377bp and 43bp) and one band for the ORF encoding VP6 (1197bp), since the enzymes both only cuts at one site.

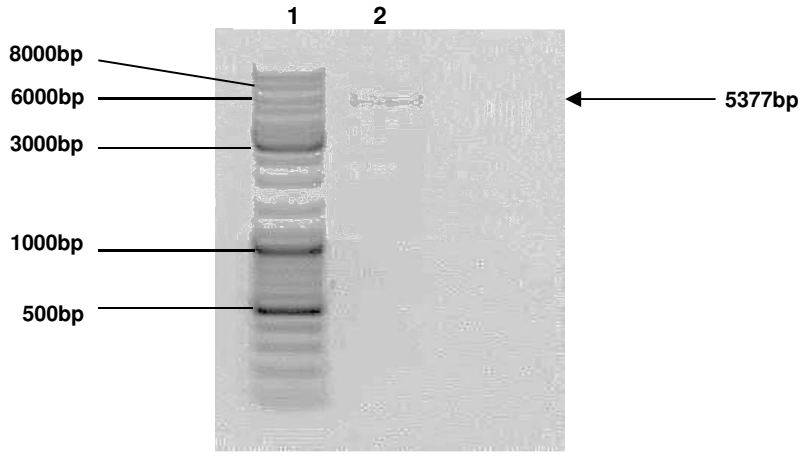


**Figure 2.6: Agarose gel electrophoresis analysis of NcoI and EcoRI double restriction enzyme digest of pETDuet-1.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl volume of 30 µl NcoI and EcoRI digestion reaction was loaded as follow; lane: 2) pETDuet-1.

The expected result of a band at 5377bp for the restriction digest was obtained for pETDuet-1 (Figure 2.6), but the 43bp band could not be observed due to the small size of the band. The yield of the ORF encoding VP6 was very low and thus not analysed on a gel. The samples of the double restriction digestion were analysed by a 0.8% gel and purified by means of a gel extraction as described in section 2.2.3.7. This yielded about 35 µl purified product of both the ORF encoding VP6 and pETDuet-1 (46.6 ng/µl ORF encoding VP6 and 66.6 ng/µl pETDuet-1). Ligation of the ORF encoding VP6 into pETDuet-1 was done as described in section 2.2.3.8. The background control contained no DNA insert. Transformation of competent JM109 cells, using above mentioned ligation reaction mixtures, was done as described in section 2.2.3.10. More than 100 colonies were obtained. Twenty-three colonies of pETDuet-1_VP6 were chosen for mini-preparation of plasmid DNA and screened by means of double restriction digestion, as described in section 2.2.3.6 (prepared in a total reaction volume of 25 µl). Mini-preparation was done on overnight LB medium cultures of the chosen colonies. Restriction analysis was done on the mini-preparation of plasmid DNA to see if the ORF of genome segment 6 was cloned into the pETDuet-1 vector.

The plasmid map that would result when the ORF encoding VP6 is cloned into pETDuet-1, is shown in Figure 2.7

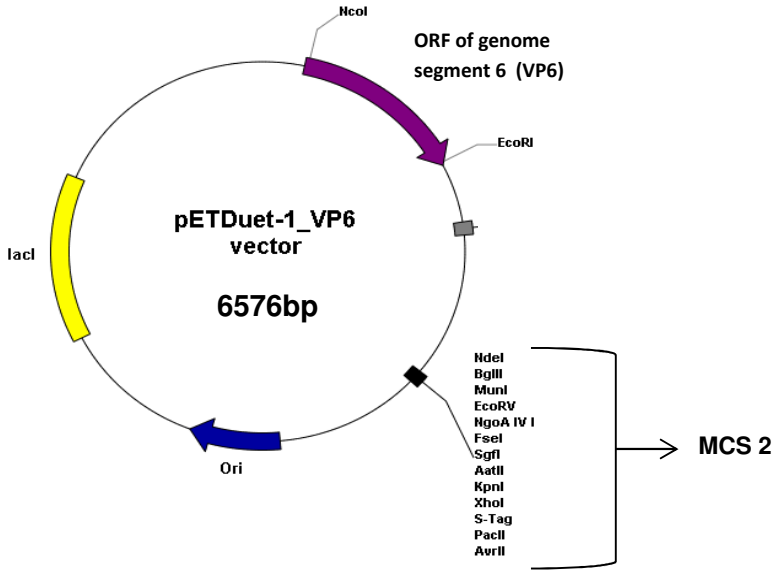


**Figure 2.7: Plasmid map that would result when the ORF of the bacterial codon optimised genome segment 6 is cloned into the pETDuet-1 vector.** The NcoI and EcoRI recognition sites indicate the enzymes used to clone the ORF of genome segment 6 (indicated in purple) into the pETDuet-1 vector, the Ori site is indicated in blue, the *lacI* in yellow and the multiple cloning site 2 in black.

Each sample was subjected to a double restriction endonuclease digestion with the restriction enzymes NcoI and EcoRI, since these were the enzymes used for cloning. A 10 µl sample of each 25 µl double restriction digestion was analysed on a 1% agarose gel, as shown in Figure 2.8. A colony was considered positive if two fragments (5377bp and 1197bp) were visible after the double digestion reaction.

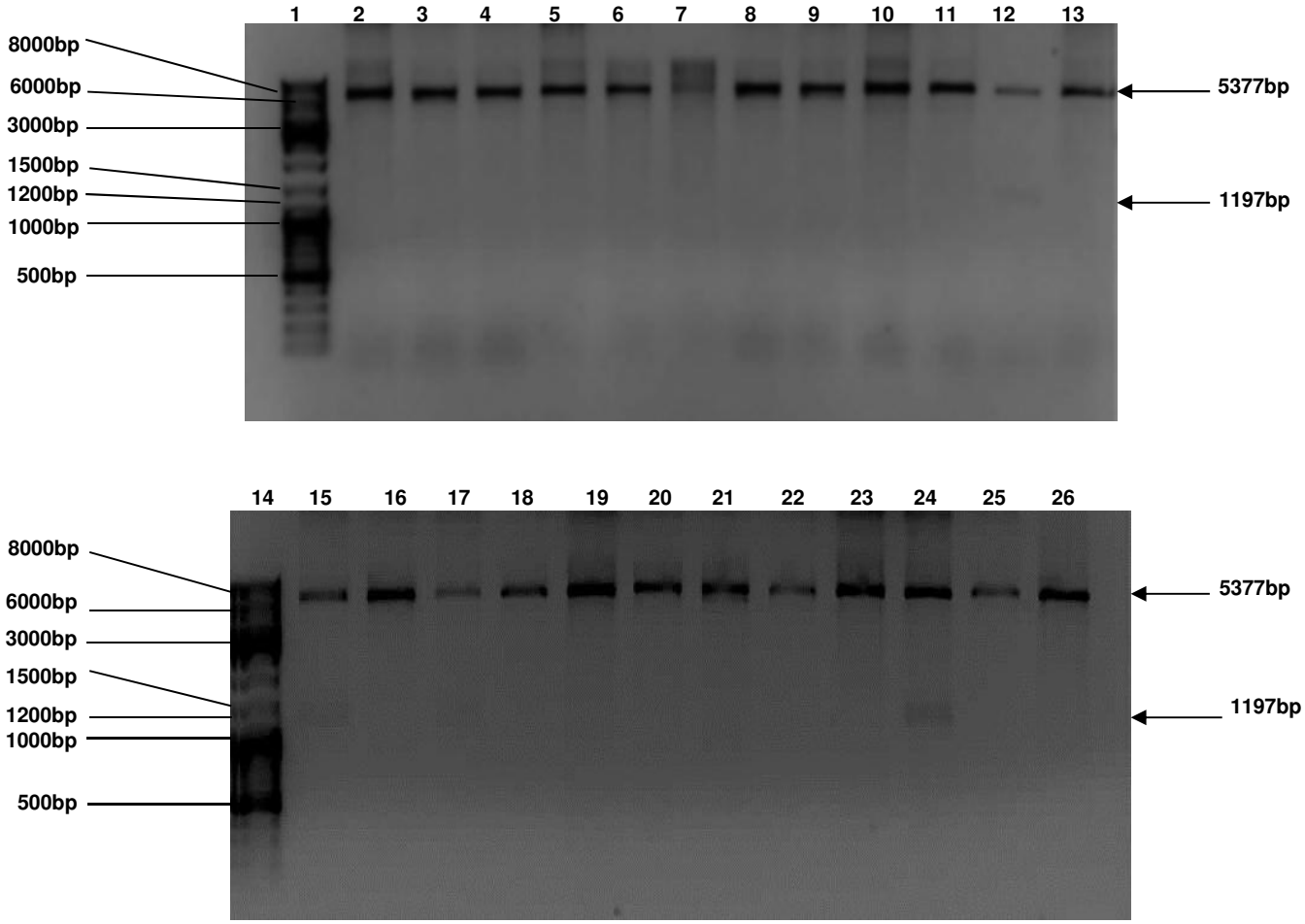


**Figure 2.8: Analysis by agarose gel electrophoresis of restriction enzyme digests of possible pETDuet-1_VP6 plasmids.** Lanes: 1 and 14) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 25 µl NcoI and EcoRI pETDuet-1_VP6 digestion reaction was loaded as follows; lanes: 2) colony 1; 3) colony 2; 4) colony 3; 5) colony 4; 6 colony 5; 7) colony 6; 8) colony 7; 9) colony 8; 10 colony 9; 11) colony 10; 12) colony 11; 13) colony 12; 15) colony 13: 16) colony 14; 17) colony 15; 18) colony 16; 19) colony 17; 20) colony 18; 21) colony 19; 22 colony 20; 23) colony 21; 24) colony 22; 25) colony 23; 26) Negative control, pETDuet-1 containing no genome segment 6.

Colonies 2-10, 12, 14, 16-21 and 23-25 showed no visible band at 5377bp and 1197bp with the restriction enzyme digestions. However four colonies (colonies 11, 13, 15 and 22) did give the expected results namely two bands one at 5377bp and one at 1197bp. Therefore,

all four colonies were used for further analysis. A midi-preparation of the plasmid DNA was done for all four as described in section 2.2.3.1. A 10 µl sample of each 400 µl plasmid preparation was analysed on a 1% agarose gel, as shown in Figure 2.9.

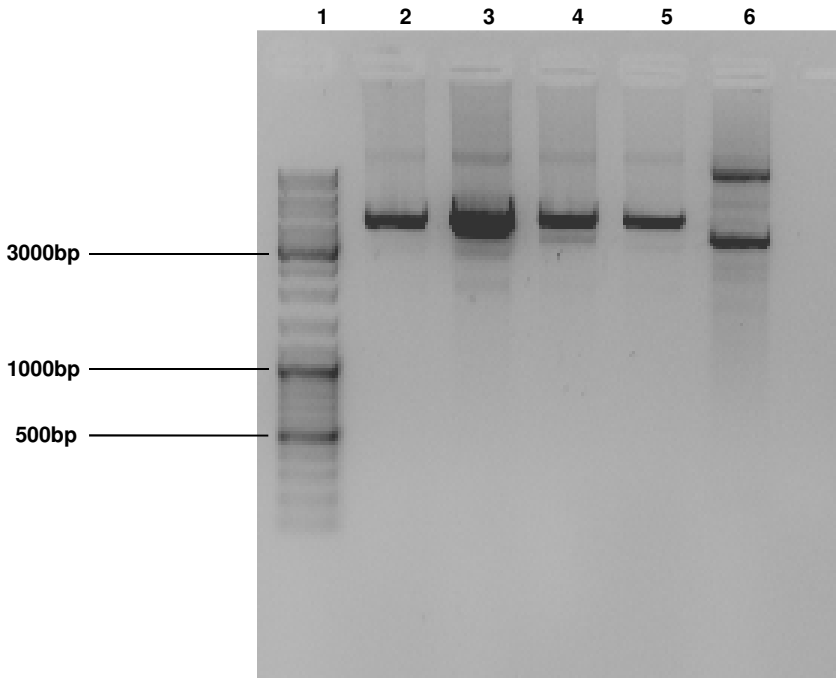


**Figure 2.9: Analysis by agarose gel electrophoresis of plasmid extraction of possible pETDuet-1_VP6 colonies.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of each 400 µl pETDuet-1_VP6 midi-preparation was loaded as follows; lanes: 2) colony 11; 3) colony 13; 4) colony 15; 5) colony 22; 6) pETDuet-1 containing no insert.

The results in Figure 2.9 showed that plasmid DNA was obtained for all four pETDuet-1_VP6 colonies. These results were expected but sequencing was still necessary to ensure that the plasmid contained the coding region of interest. The plasmid DNA of two colonies (colony 11 - lane 2 and colony 13 - lane 3) was sent for sequencing, to ensure that the ORF encoding VP6 was cloned into pETDuet-1 and that no mutations had occurred.

Commercial available primers (Table 2.4) were used for sequencing of pETDuet-1_VP6 clones, as described in section 2.2.3.12. The sequence alignment is indicated in Figure 2.10.

```
pETDuet-1_VP6     ....................................GG      2
pETDuet-1_VP6_F   ggcatacctgatccggcgcattaatacgactcactatagG     40
pETDuet-1_VP6_R   ....................................      0


pETDuet-1_VP6     GGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATT     42
pETDuet-1_VP6_F   GGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATT     80
pETDuet-1_VP6_R   ....................................      0


pETDuet-1_VP6     TTGTTTAACTTTAAGAAGGAGATATACCATGGATGTGCTG     82
pETDuet-1_VP6_F   TTGTTTAACTTTAAGAAGGAGATATACCATGGATGTGCTG    120
pETDuet-1_VP6_R   ....................................      0


pETDuet-1_VP6     TATAGCCTGAGCAAAACCCTGAAAGATGCGCGTGATAAAA    122
pETDuet-1_VP6_F   TATAGCCTGAGCAAAACCCTGAAAGATGCGCGTGATAAAA    160
pETDuet-1_VP6_R   ....................................      0


pETDuet-1_VP6     TTGTGGAAGGCACCCTGTATAGCAACGTGAGCGATCTGAT    162
pETDuet-1_VP6_F   TTGTGGAAGGCACCCTGTATAGCAACGTGAGCGATCTGAT    200
pETDuet-1_VP6_R   ....................................AT     2
                                                       **

pETDuet-1_VP6     TCAGCAGTTTAACCAGATGATCATTACCATGAACGGCAAC    202
pETDuet-1_VP6_F   TCAGCAGTTTAACCAGATGATCATTACCATGAACGGCAAC    240
pETDuet-1_VP6_R   TCAGCAGTTTAACCAGATGATCATTACCATGAACGGCAAC     42
                  ************************************

pETDuet-1_VP6     GAATTTCAGACCGGCGGCATTGGCAACCTGCCGATTCGTA    242
pETDuet-1_VP6_F   GAATTTCAGACCGGCGGCATTGGCAACCTGCCGATTCGTA    280
pETDuet-1_VP6_R   GAATTTCAGACCGGCGGCATTGGCAACCTGCCGATTCGTA     82
                  ************************************

pETDuet-1_VP6     ACTGGAACTTTGATTTTGGCCTGCTGGGCACCACGCTGCT    282
pETDuet-1_VP6_F   ACTGGAACTTTGATTTTGGCCTGCTGGGCACCACGCTGCT    320
pETDuet-1_VP6_R   ACTGGAACTTTGATTTTGGCCTGCTGGGCACCACGCTGCT    122
                  ************************************

pETDuet-1_VP6     GAACCTGGATGCGAACTATGTGGAAACCGCGCGTAACACC    322
pETDuet-1_VP6_F   GAACCTGGATGCGAACTATGTGGAAACCGCGCGTAACACC    360
pETDuet-1_VP6_R   GAACCTGGATGCGAACTATGTGGAAACCGCGCGTAACACC    162
                  ************************************

pETDuet-1_VP6     ATTGATTACTTCGTGGATTTCGTTGATAACGTGTGCATGG    362
pETDuet-1_VP6_F   ATTGATTACTTCGTGGATTTCGTTGATAACGTGTGCATGG    400
pETDuet-1_VP6_R   ATTGATTACTTCGTGGATTTCGTTGATAACGTGTGCATGG    202
                  ************************************

pETDuet-1_VP6     ATGAAATGGTGCGTGAAAGCCAGCGTAACGGCATCGCGCC    402
pETDuet-1_VP6_F   ATGAAATGGTGCGTGAAAGCCAGCGTAACGGCATCGCGCC    440
pETDuet-1_VP6_R   ATGAAATGGTGCGTGAAAGCCAGCGTAACGGCATCGCGCC    242
                  ************************************

pETDuet-1_VP6     GCAGAGCGATAGCCTGCGTAAACTGAGCGGCATTAAATTC    442
pETDuet-1_VP6_F   GCAGAGCGATAGCCTGCGTAAACTGAGCGGCATTAAATTC    480
pETDuet-1_VP6_R   GCAGAGCGATAGCCTGCGTAAACTGAGCGGCATTAAATTC    282
                  ************************************
```

```
pETDuet-1_VP6    AAACGCATCAACTTCGATAACAGCAGCGAATATATCGAAA    482
pETDuet-1_VP6_F  AAACGCATCAACTTCGATAACAGCAGCGAATATATCGAAA    520
pETDuet-1_VP6_R  AAACGCATCAACTTCGATAACAGCAGCGAATATATCGAAA    322
                 ****************************************

pETDuet-1_VP6    ACTGGAACCTGCAGAACCGTCGTCAGCGTACCGGCTTTAC    522
pETDuet-1_VP6_F  ACTGGAACCTGCAGAACCGTCGTCAGCGTACCGGCTTTAC    560
pETDuet-1_VP6_R  ACTGGAACCTGCAGAACCGTCGTCAGCGTACCGGCTTTAC    362
                 ****************************************

pETDuet-1_VP6    CTTTCATAAACCGAACATCTTTCCGTATAGCGCGAGCTTT    562
pETDuet-1_VP6_F  CTTTCATAAACCGAACATCTTTCCGTATAGCGCGAGCTTT    600
pETDuet-1_VP6_R  CTTTCATAAACCGAACATCTTTCCGTATAGCGCGAGCTTT    402
                 ****************************************

pETDuet-1_VP6    ACCCTGAACCGTAGCCAGCCGGCGCATGATAACCTGATGG    602
pETDuet-1_VP6_F  ACCCTGAACCGTAGCCAGCCGGCGCATGATAACCTGATGG    640
pETDuet-1_VP6_R  ACCCTGAACCGTAGCCAGCCGGCGCATGATAACCTGATGG    442
                 ****************************************

pETDuet-1_VP6    GCACCATGTGGCTGAACGCGGGCAGCGAAATTCAGGTGGC    642
pETDuet-1_VP6_F  GCACCATGTGGCTGAACGCGGGCAGCGAAATTCAGGTGGC    680
pETDuet-1_VP6_R  GCACCATGTGGCTGAACGCGGGCAGCGAAATTCAGGTGGC    482
                 ****************************************

pETDuet-1_VP6    GGGCTTTGATTATAGCTGCGCGATTAACGCGCCGGCGAAC    682
pETDuet-1_VP6_F  GGGCTTTGATTATAGCTGCGCGATTAACGCGCCGGCGAAC    720
pETDuet-1_VP6_R  GGGCTTTGATTATAGCTGCGCGATTAACGCGCCGGCGAAC    522
                 ****************************************

pETDuet-1_VP6    ACCCAGCAGTTTGAACATATTGTGCAGCTGCGTCGTGTGC    722
pETDuet-1_VP6_F  ACCCAGCAGTTTGAACATATTGTGCAGCTGCGTCGTGTGC    760
pETDuet-1_VP6_R  ACCCAGCAGTTTGAACATATTGTGCAGCTGCGTCGTGTGC    562
                 ****************************************

pETDuet-1_VP6    TGACCACCGCGACCATTACCCTGCTGCCGGATGCGGAACG    762
pETDuet-1_VP6_F  TGACCACCGCGACCATTACCCTGCTGCCGGATGCGGAACG    800
pETDuet-1_VP6_R  TGACCACCGCGACCATTACCCTGCTGCCGGATGCGGAACG    602
                 ****************************************

pETDuet-1_VP6    TTTTAGCTTTCCGCGTGTGATTAACAGCGCGGATGGCGCG    802
pETDuet-1_VP6_F  TTTTAGCTTTCCGCGTGTGATTAACAGCGCGGATGGCGCG    840
pETDuet-1_VP6_R  TTTTAGCTTTCCGCGTGTGATTAACAGCGCGGATGGCGCG    642
                 ****************************************

pETDuet-1_VP6    ACCACCTGGTATTTTAATCCGGTGATTCTGCGTCCGAACA    842
pETDuet-1_VP6_F  ACCACCTGGTATTTTAATCCGGTGATTCTGCGTCCGAACA    880
pETDuet-1_VP6_R  ACCACCTGGTATTTTAATCCGGTGATTCTGCGTCCGAACA    682
                 ****************************************

pETDuet-1_VP6    ACGTGGAAGTGGAATTTCTGCTGAACGGCCAGATTATTAA    882
pETDuet-1_VP6_F  ACGTGGAAGTGGAATTTCTGCTGAACGGCCAGATTATTAA    920
pETDuet-1_VP6_R  ACGTGGAAGTGGAATTTCTGCTGAACGGCCAGATTATTAA    722
                 ****************************************

pETDuet-1_VP6    CACCTATCAGGCGCGTTTTGGCACCATTGTGGCGCGTAAC    922
pETDuet-1_VP6_F  CACCTATCAGGCGCGTTTTGGCACCATTGTGGCGCGTAAC    960
pETDuet-1_VP6_R  CACCTATCAGGCGCGTTTTGGCACCATTGTGGCGCGTAAC    762
                 ****************************************
```

```
pETDuet-1_VP6    TTTGATACCATTCGCCTGAGCTTTCAGCTGATGCGTCCGC    962
pETDuet-1_VP6_F  TTTGATACCATTCGCCTGAGCTTTCAGCTGATGCGTCCGC   1000
pETDuet-1_VP6_R  TTTGATACCATTCGCCTGAGCTTTCAGCTGATGCGTCCGC    802
                 ************************************

pETDuet-1_VP6    CGAACATGACCCCGAGCGTGGCGGCGCTGTTTCCGAACGC   1002
pETDuet-1_VP6_F  CGAACATGACCCCGAGCGTGGCGGCGCTGTTTCCGAACGC   1040
pETDuet-1_VP6_R  CGAACATGACCCCGAGCGTGGCGGCGCTGTTTCCGAACGC    842
                 ************************************

pETDuet-1_VP6    GCAGCCGTTTGAACATCATGCGACCGTTGGCCTGACCCTG   1042
pETDuet-1_VP6_F  GCAGCCGTTTGAACATCATGCGACCGTTGGCCTGACCCTG   1080
pETDuet-1_VP6_R  GCAGCCGTTTGAACATCATGCGACCGTTGGCCTGACCCTG    882
                 ************************************

pETDuet-1_VP6    AAAATTGAAAGCGCGGTGTGCGAAAGCGTTCTGGCCGATG   1082
pETDuet-1_VP6_F  AAAATTGAAAGCGCGGTGTGCGAAAGCGTTCTGGCCGATG   1120
pETDuet-1_VP6_R  AAAATTGAAAGCGCGGTGTGCGAAAGCGTTCTGGCCGATG    922
                 ************************************

pETDuet-1_VP6    CGAGCGAAACCATGCTGGCCAACGTGACCAGCGTGCGTCA   1122
pETDuet-1_VP6_F  CGAGCGAAACCATGCTGGCCAACGTGACCAGCGTGCGTCA   1160
pETDuet-1_VP6_R  CGAGCGAAACCATGCTGGCCAACGTGACCAGCGTGCGTCA    962
                 ************************************

pETDuet-1_VP6    GGAATATGCGATTCCGGTTGGCCCGGTGTTTCCGCCGGGT   1162
pETDuet-1_VP6_F  GGAATATGCGATTCCGGTTGGCCCGGTGT...........   1189
pETDuet-1_VP6_R  GGAATATGCGATTCCGGTTGGCCCGGTGTTTCCGCCGGGT   1002
                 ****************************

pETDuet-1_VP6    ATGAACTGGACCGATCTGATTACCAACTATAGCCCGAGCC   1202
pETDuet-1_VP6_F  ....................................      1189
pETDuet-1_VP6_R  ATGAACTGGACCGATCTGATTACCAACTATAGCCCGAGCC   1042

pETDuet-1_VP6    GTGAAGATAACCTGCAGCGTGTGTTTACCGTGGCGAGCAT   1242
pETDuet-1_VP6_F  ....................................      1189
pETDuet-1_VP6_R  GTGAAGATAACCTGCAGCGTGTGTTTACCGTGGCGAGCAT   1082

pETDuet-1_VP6    TCGTAGCATGCTGGTGAAATAGTAAGAATTCGAGCTCGGC   1282
pETDuet-1_VP6_F  ....................................      1189
pETDuet-1_VP6_R  TCGTAGCATGCTGGTGAAATAGTAAGAATTCGAGCTCGGC   1122
```

**Figure 2.10: Nucleotide sequence alignments of the recombinant pETDuet-1 vector containing the ORF of the bacterial codon optimised genome segment 6 (VP6) with the *in silico* clone of pETDuet-1_VP6.** The *in silico* clone of pETDuet-1_VP6 is indicated as pETDuet-1_VP6, the forward sequence is indicated as pETDuet1_VP6_F nucleotide sequence and the reverse sequence is indicated as pETDuet1_VP6_R nucleotide sequence. The restriction enzymes used for the amplification of genome segment 6 are indicated in blue. The start codon of genome segment 6 is indicated in green and the stop codon is indicated in red. The T7 promoter-1 is indicated in purple, the T7 transcription start-1 is indicated in dark blue and the *lac* operator in yellow.

Both colony 11 and colony 13 gave the expected results, however, only the sequences results of colony 13 are shown. It was clear from Figure 2.10 that the ORF encoding VP6 was cloned into pETDuet-1. The commercially available primers used for sequencing made it possible to see parts of the sequences namely the restriction enzymes used for cloning, the start and stop codon of the bacterial codon optimised ORF of genome segment 6 as well as the T7 promoter-1, the T7 transcription start-1, the *lac* operator as well as the ribosome binding sequence. These results were promising making it possible to go to the next step of the project namely the individual expression of VP6 in bacteria (discussed in section 2.3.3).

### 2.3.2 Cloning of open reading frame encoding VP2 into pETDuet-1 and pETDuet-1_VP6

The synthetic ORF of bacterial codon optimised genome segment 2 (VP2) was obtained from Dr AC Potgieter (Deltamune (Pty) Ltd R&D, South Africa), who provided genome segment 2 in the pColdTF vector. Dr AC Potgieter found that it was possible to express VP2 in bacteria when using the expression protocol of the cold shock vectors.

In order to clone the ORF encoding VP2 into the expression vector pETDuet-1 (Figure 2.1.), it was PCR amplified from pColdTF_VP2 with primers that were designed to contain the restriction enzymes; BglII and XmaJI, to facilitate directional cloning as illustrated in Figure 2.11. The ORFs of genome segment 2 and 6 should, therefore, not contain recognition sequences for these two restriction enzymes, to ensure that the restriction enzymes do not cut within either the ORF of genome segment 6 or genome segment 2. These enzymes sites are present in multiple cloning site 2 (MCS2) of pETDuet-1. The primers were designed using the bacterial codon optimised sequence of genome segment 2 (VP2) (Figure 2.11).

Restriction enzyme *BglII*
Forward primer 5'CCAG**AGATCT**ATGGCGTA

```
1       CGAATTGGGG ATCGATCCAC TAGTTAATTA ACGCGATGTA AGGTACCTCG AGATGGCGTA
        GCTTAACCCC TAGCTAGGTG ATCAATTAAT TGCGCTACAT TCCATGGAGC TCTACCGCAT

        TCG'3
61      TCGTAAACGT GGCGCGCGTC GTGAAGCGAA CCTGAACAAC AACGATCGTA TGCAGGAAAA
        AGCATTTGCA CCGCGCGCAG CACTTCGCTT GGACTTGTTG TTGCTAGCAT ACGTCCTTTT
121     AATCGACGAA AAACAGGATA GCAACAAAAT TCAGCTGTCT GATAAAGTTC TGAGCAAAAA
        TTAGCTGCTT TTTGTCCTAT CGTTGTTTTA AGTCGACAGA CTATTTCAAG ACTCGTTTTT

181     AGAAGAAATC GTTACCGATA GCCACGAAGA AGTGAAAGTT ACCGACGAAC TGAAAAAAAG
        TCTTCTTTAG CAATGGCTAT CGGTGCTTCT TCACTTTCAA TGGCTGCTTG ACTTTTTTTC

241     CACCAAAGAA GAAAGCAAAC AGCTGCTGGA AGTGCTGAAA ACGAAAGAAG AGCACCAGAA
        GTGGTTTCTT CTTTCGTTTG TCGACGACCT TCACGACTTT TGCTTTCTTC TCGTGGTCTT

301     AGAAATCCAG TATGAAATCC TGCAGAAAAC CATTCCGACC TTTGAACCGA AAGAAACCAT
        TCTTTAGGTC ATACTTTAGG ACGTCTTTTG GTAAGGCTGG AAACTTGGCT TTCTTTGGTA

361     TCTGCGCAAA CTGGAAGATA TTCAGCCGGA ACTGGCCAAA AAACAGACCA AACTGTTCCG
        AGACGCGTTT GACCTTCTAT AAGTCGGCCT TGACCGGTTT TTTGTCTGGT TTGACAAGGC

421     TATTTTCGAA CCGAAACAGC TGCCGATTTA TCGTGCGAAC GGCGAACGTG AACTGCGTAA
        ATAAAAGCTT GGCTTTGTCG ACGGCTAAAT AGCACGCTTG CCGCTTGCAC TTGACGCATT

481     CCGTTGGTAC TGGAAACTGA AAAAAGATAC CCTGCCGGAT GGCGATTATG ATGTGCGCGA
        GGCAACCATG ACCTTTGACT TTTTTCTATG GGACGGCCTA CCGCTAATAC TACACGCGCT

541     ATATTTCCTG AACCTGTATG ATCAGGTGCT GACCGAAATG CCGGATTATC TGCTGCTGAA
        TATAAAGGAC TTGGACATAC TAGTCCACGA CTGGCTTTAC GGCCTAATAG ACGACGACTT

601     AGATATGGCG GTGGAAAATA AAAACAGCCG TGATGCGGGC AAAGTGGTGG ATAGCGAAAC
        TCTATACCGC CACCTTTTAT TTTTGTCGGC ACTACGCCCG TTTCACCACC TATCGCTTTG

661     CGCGAGCATT TGTGATGCGA TCTTCCAGGA TGAAGAAACC GAAGGCGCGG TGCGTCGTTT
        GCGCTCGTAA ACACTACGCT AGAAGGTCCT ACTTCTTTGG CTTCCGCGCC ACGCAGCAAA

721     TATTGCGGAA ATGCGTCAGC GTGTGCAGGC GGATCGTAAC GTGGTGAACT ATCCGAGCAT
        ATAACGCCTT TACGCAGTCG CACACGTCCG CCTAGCATTG CACCACTTGA TAGGCTCGTA

781     TCTGCATCCG ATTGATTATG CGTTCAACGA ATACTTTCTG CAGCATCAGC TGGTTGAACC
        AGACGTAGGC TAACTAATAC GCAAGTTGCT TATGAAAGAC GTCGTAGTCG ACCAACTTGG

841     GCTGAACAAC GATATCATCT TCAACTATAT CCCGGAACGT ATTCGTAACG ATGTGAACTA
        CGACTTGTTG CTATAGTAGA AGTTGATATA GGGCCTTGCA TAAGCATTGC TACACTTGAT

901     CATCCTGAAC ATGGATCGCA ACCTGCCGAG CACCGCGCGT TATATCCGTC CGAACCTGCT
        GTAGGACTTG TACCTAGCGT TGGACGGCTC GTGGCGCGCA ATATAGGCAG GCTTGGACGA

961     GCAGGATCGT CTGAACCTGC ATGATAACTT CGAAAGCCTG TGGGATACCA TTACCACGAG
        CGTCCTAGCA GACTTGGACG TACTATTGAA GCTTTCGGAC ACCCTATGGT AATGGTGCTC

1021    CAACTATATT CTGGCCCGTA GCGTGGTGCC GGATCTGAAA GAACTGGTGA GCACCGAAGC
        GTTGATATAA GACCGGGCAT CGCACCACGG CCTAGACTTT CTTGACCACT CGTGGCTTCG

1081    GCAGATTCAG AAAATGAGCC AGGATCTGCA GCTGGAAGCG CTGACCATTC AGAGCGAAAC
        CGTCTAAGTC TTTTACTCGG TCCTAGACGT CGACCTTCGC GACTGGTAAG TCTCGCTTTG
```

```
1141    CCAGTTTCTG ACCGGCATTA ACAGCCAGGC GGCGAACGAT TGCTTTAAAA CCCTGATTGC
        GGTCAAAGAC TGGCCGTAAT TGTCGGTCCG CCGCTTGCTA ACGAAATTTT GGGACTAACG

1201    GGCCATGCTG TCTCAGCGTA CCATGAGCCT GGATTTTGTG ACCACCAACT ATATGAGCCT
        CCGGTACGAC AGAGTCGCAT GGTACTCGGA CCTAAAACAC TGGTGGTTGA TATACTCGGA

1261    GATTAGCGGC ATGTGGCTGC TGACCGTGGT GCCGAACGAC ATGTTTATTC GTGAAAGCCT
        CTAATCGCCG TACACCGACG ACTGGCACCA CGGCTTGCTG TACAAATAAG CACTTTCGGA

1321    GGTGGCGTGC CAGCTGGCCA TTGTGAACAC CATTATCTAT CCGGCGTTTG GCATGCAGCG
        CCACCGCACG GTCGACCGGT AACACTTGTG GTAATAGATA GGCCGCAAAC CGTACGTCGC

1381    TATGCATTAT CGTAACGGCG ATCCGCAGAC CCCGTTTCAG ATTGCGGAAC AGCAGATCCA
        ATACGTAATA GCATTGCCGC TAGGCGTCTG GGGCAAAGTC TAACGCCTTG TCGTCTAGGT

1441    GAACTTTCAG GTGGCGAACT GGCTGCATTT TGTGAACAAC AACCAGTTTC GTCAGGCGGT
        CTTGAAAGTC CACCGCTTGA CCGACGTAAA ACACTTGTTG TTGGTCAAAG CAGTCCGCCA

1501    GATTGATGGC GTGCTGAACC AGGTGCTGAA CGATAACATT CGTAACGGCC ATGTGATTAA
        CTAACTACCG CACGACTTGG TCCACGACTT GCTATTGTAA GCATTGCCGG TACACTAATT

1561    CCAACTGATG GAAGCCCTGA TGCAGCTGTC TCGTCAGCAG TTTCCGACCA TGCCGATCGA
        GGTTGACTAC CTTCGGGACT ACGTCGACAG AGCAGTCGTC AAAGGCTGGT ACGGCTAGCT

1621    TTATAAACGT AGCATTCAGC GTGGCATTCT GCTGCTGTCT AACCGTCTGG GCCAGCTGGT
        AATATTTGCA TCGTAAGTCG CACCGTAAGA CGACGACAGA TTGGCAGACC CGGTCGACCA

1681    TGATCTGACC CGTCTGCTGG CCTATAACTA TGAAACGCTG ATGGCATGCA TCACGATGAA
        ACTAGACTGG GCAGACGACC GGATATTGAT ACTTTGCGAC TACCGTACGT AGTGCTACTT

1741    CATGCAGCAT GTGCAGACCC TGACCACCGA AAAACTGCAG CTGACCAGCG TGACCAGCCT
        GTACGTCGTA CACGTCTGGG ACTGGTGGCT TTTTGACGTC GACTGGTCGC ACTGGTCGGA

1801    GTGCATGCTG ATTGGCAACG CGACCGTGAT TCCGAGCCCG CAGACCCTGT TCCATTATTA
        CACGTACGAC TAACCGTTGC GCTGGCACTA AGGCTCGGGC GTCTGGGACA AGGTAATAAT

1861    TAACGTGAAC GTGAACTTTC ACAGCAACTA TAACGAACGT ATTAACGATG CGGTGGCGAT
        ATTGCACTTG CACTTGAAAG TGTCGTTGAT ATTGCTTGCA TAATTGCTAC GCCACCGCTA

1921    TATTACCGCG GCGAACCGTC TGAATCTGTA TCAGAAAAAA ATGAAAGCGA TCGTGGAAGA
        ATAATGGCGC CGCTTGGCAG ACTTAGACAT AGTCTTTTTT TACTTTCGCT AGCACCTTCT

1981    TTTTCTGAAA CGCCTGTACA TTTTTGATGT GAGCCGTGTG CCGGATGATC AGATGTATCG
        AAAAGACTTT GCGGACATGT AAAAACTACA CTCGGCACAC GGCCTACTAG TCTACATAGC

2041    TCTGCGTGAT CGTCTGCGTC TGCTGCCGGT GGAAATTCGT CGTCTGGATA TCTTTAACCT
        AGACGCACTA GCAGACGCAG ACGACGGCCA CCTTTAAGCA GCAGACCTAT AGAAATTGGA

2101    GATCCTGATG AATATGGATC AGATCGAACG TGCGAGCGAT AAAATTGCGC AGGGCGTGAT
        CTAGGACTAC TTATACCTAG TCTAGCTTGC ACGCTCGCTA TTTTAACGCG TCCCGCACTA

2161    TATTGCGTAT CGTGATATGC ATCTGGAACG TGATGAAATG TATGGCTACG TGAACATTGC
        ATAACGCATA GCACTATACG TAGACCTTGC ACTACTTTAC ATACCGATGC ACTTGTAACG

2221    GCGTAACCTG GAAGGCTTTC AGCAGATTAA CCTGGAAGAA CTGATGCGTA GCGGCGATTA
        CGCATTGGAC CTTCCGAAAG TCGTCTAATT GGACCTTCTT GACTACGCAT CGCCGCTAAT

2281    TGCGCAGATT ACCAACATGC TGCTGAACAA CCAGCCGGTG GCGCTGGTTG GTGCGCTGCC
        ACGCGTCTAA TGGTTGTACG ACGACTTGTT GGTCGGCCAC CGCGACCAAC CACGCGACGG
```

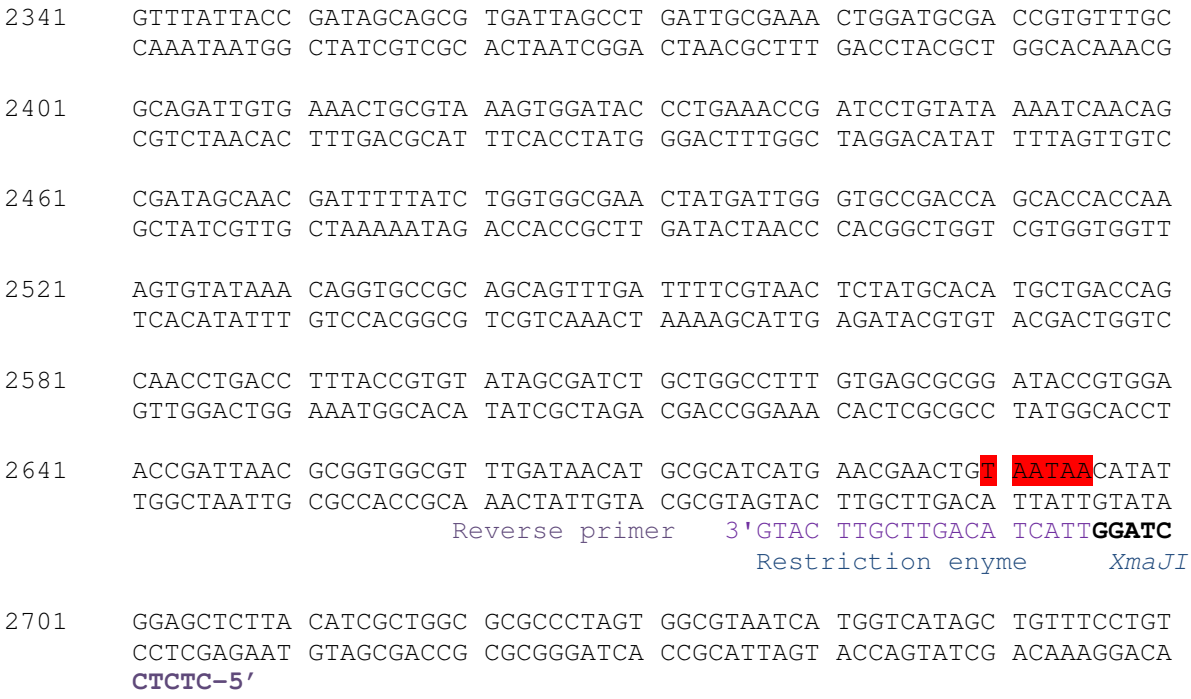


**Figure 2.11: Sequence of the ORF encoding VP2 (GR10924 G9P[6]) indicating the primers used for amplification.** The forward and reverse primers are indicated in purple and the restriction enzymes sites in black. The start codon is indicated in green and the stop codon in red.

PCR amplification was performed at an annealing temperature of 57°C, as described in section 2.2.3.2. A 5 µl sample of the amplification reaction was analysed using a 1% agarose gel (Figure 2.12).

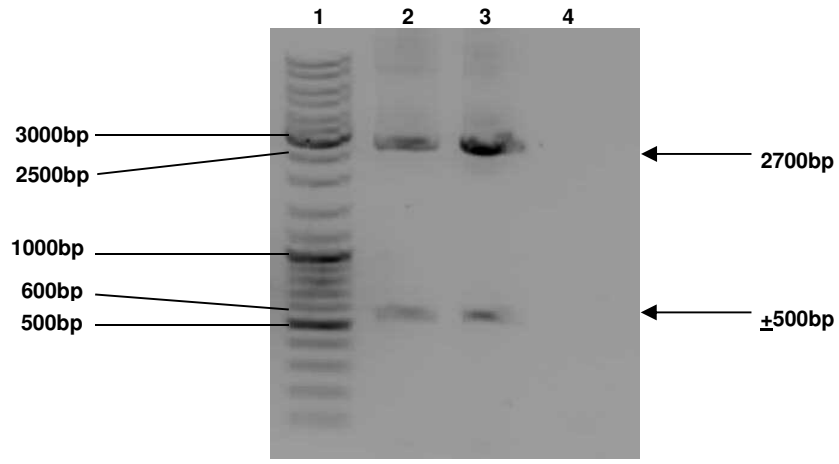


**Figure 2.12: Analysis by agarose gel electrophoresis of the PCR amplification of the ORF encoding VP2 from the pColdTF vector.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl of each 50 µl PCR reaction was loaded as follows; lanes: 2-3) genome segment 2 amplification reactions (same reaction done in duplicate); 4) negative control.

The PCR amplification of the ORF encoding VP2, at 57°C annealing temperature, was successful but there was a non-specific ± 500bp amplicon visible on the agarose gel (Figure 2.12). Lanes 2 and 3 show a 2700bp amplicon as expected (the same reaction done in duplicate). The lack of a band in lane 4, which served as the negative control (the negative control has no template DNA and is usually water instead of DNA), indicated that no contamination took place. The band at 2700bp was excized before silica extraction was performed. The PCR amplified ORF encoding VP2 was purified from oligonucleotides, other contaminants and the non-specific amplicon by means of gel extraction to ensure good quality DNA for cloning. Gel extraction was done as described in section 2.2.3.7. This yielded about 50 ng/µl product of the ORF encoding VP2, which was stored at 4°C for cloning. A 5 µl sample of the purified ORF encoding VP2 was analysed using a 1% agarose gel, as shown in Figure 2.13.

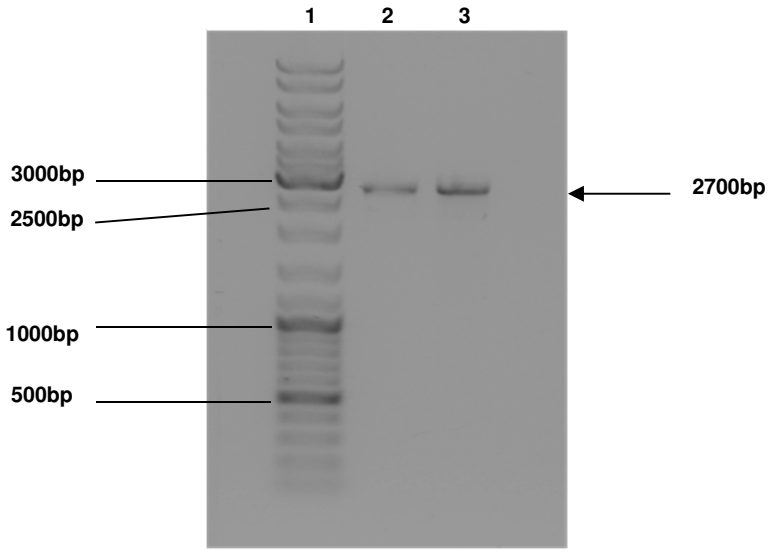


**Figure 2.13: Analysis by agarose gel electrophoresis of the amplified ORF encoding VP2 after gel extraction.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl reaction of each 50 µl gel extracted reaction was loaded as follows; lanes: 2-3) duplicate purified genome segment 2

The gel extraction purification of the ORF encoding VP2 was successful (Figure 2.13). Lanes 2 and 3 show the expected fragment at 2700bp and illustrate that the non-specific amplicon was successfully removed by gel excision of the ORF of genome segment 2.

To evaluate the individual expression of VP2 as well as the co-expression of VP2/6 in bacteria, two types of constructs had to be made namely i) pETDuet-1_VP2 and ii) pETDuet-1_VP2/6. Double restriction endonuclease digestion was done with the restriction enzymes BglII and XmaJI for the amplified ORF encoding VP2, pETDuet-1 and pETDuet-1_VP6, as described in section 2.2.3.6. These digestions were analysed on a 1% agarose gel, as shown in Figure 2.14. The expected result for the restriction digestion of the ORF encoding VP2 was to linearize and two bands to be visible for pETDuet-1 (5292bp and 128bp) and pETDuet-1_VP6 (6448bp and 128bp) since the enzymes only have one recognition site in all three of them. However, the second band for pETDuet-1_VP6 and pETDuet-1 would not be visible on the agarose gel since it is too small.

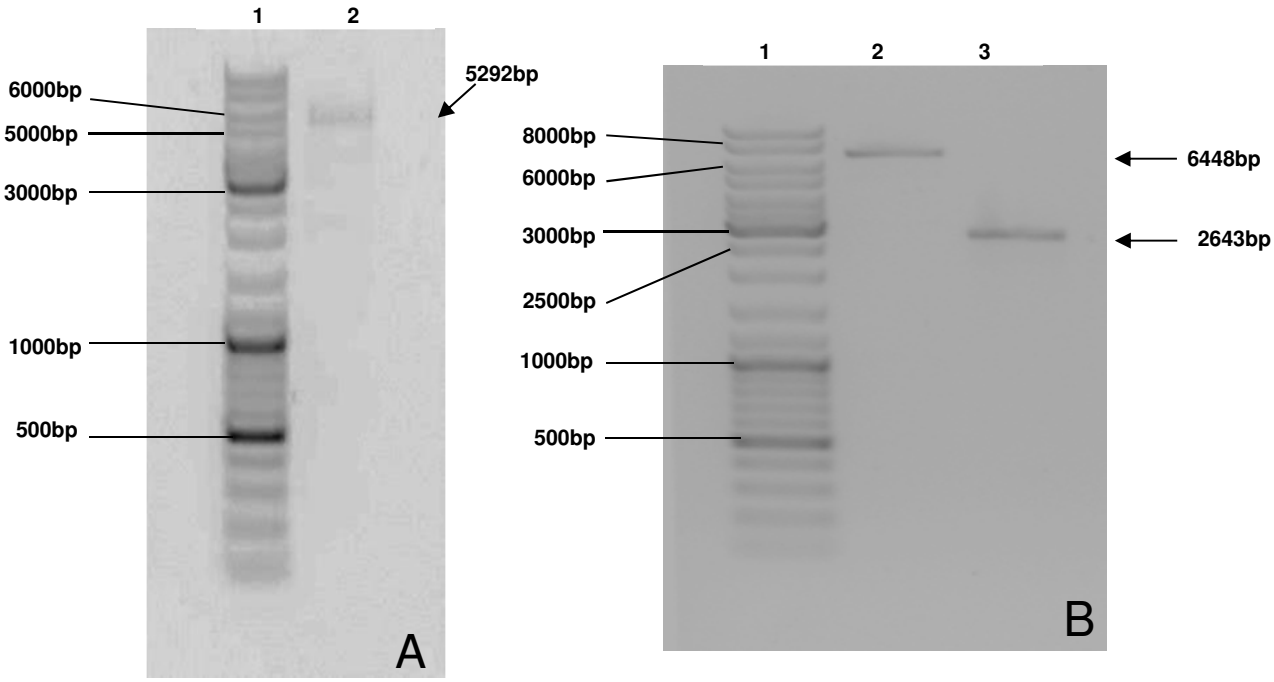


**Figure 2.14: Analysis by agarose gel electrophoresis of restriction enzyme analysis with BglII and AvrII of the ORF encoding VP2, pETDuet-1 and pETDuet-1_VP6. (A)** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl volume of 30 µl BglII and AvrII digestion reactions were loaded as follows; lane: 2) pETDuet-1; **(B)** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl volume of 30 µl BglII and AvrII digestion reactions were loaded as follows; lanes: 2) digested pETDuet-1_VP6; 3) digested genome segment 2.

The expected results for the restriction digestion were obtained for all three reactions (pETDuet-1, ORF encoding VP2 and pETduet-1_VP6) (Figure 2.14). The remaining sample of the double restriction digestions reactions were analysed by a 0.8% gel and purified by means of a gel extraction as described in section 2.2.3.7. This gel extraction yielded 22.7 ng/µl ORF encoding VP2, 15.4 ng/µl pETDuet-1 and 17.6 ng/µl pETDuet-1_VP6. Ligation reaction of the ORF encoding VP2 into pETDuet-1 and pETDuet-1_VP6 was done as described in section 2.2.3.8. The ligation background control contained no DNA insert. The transformation of JM109 competent cells, using the above mentioned ligation reactions, was done as described in section 2.2.3.10. Twenty colonies of pETDuet-1_VP2 and 20 colonies of pETDuet-1_VP2/6 were chosen for mini preparation of plasmid DNA and screening by means of double restriction digestion, as described in section 2.2.3.6. Mini-plasmid preparations were done on overnight LB medium cultures of the chosen colonies. Restriction analysis was done on plasmid DNA to see if the cloning of the ORF encoding VP2 into the

pETDuet-1 and pETDuet-1_VP6 was successful. The restriction enzymes chosen were BglII and AvrII, since these were the enzymes used for cloning. A cloning overview and the plasmid map that would result when the ORF encoding VP2 is cloned into pETDuet-1 and pETDuet-1_VP6 is illustrated in Figure 2.15.

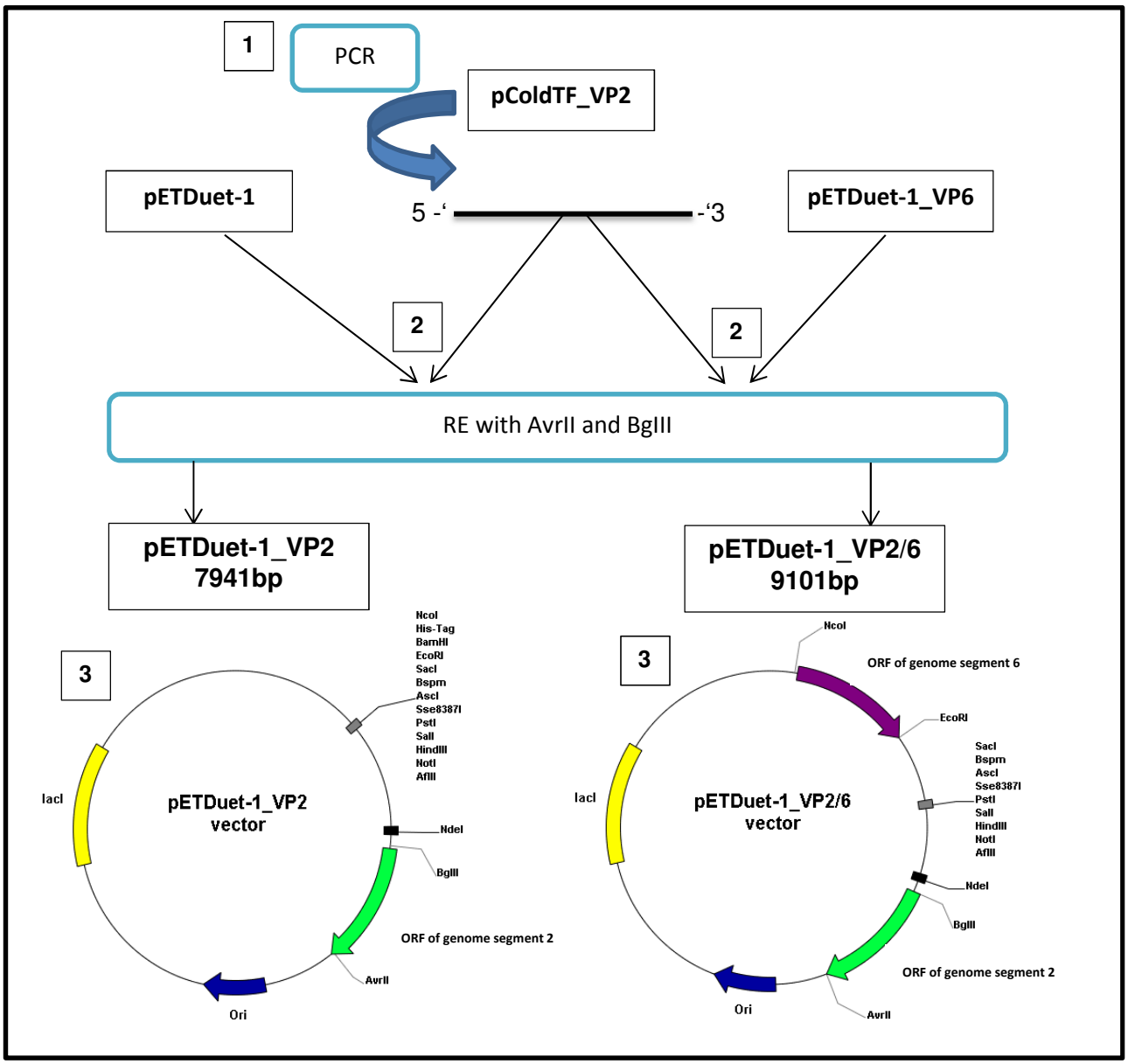


**Figure 2.15: Cloning overview and plasmid map that would result when the ORF encoding VP2 is cloned into pETDuet-1 and pETDuet-1_VP6. (1)** ORF encoding VP2 was PCR amplified from pColdTF. **(2)** RE digestion of the ORF of the bacterial codon optimised genome segment 2, pETDuet-1 and pETDuet-1_VP6 was done with BglII and AvrII. **(3)** The ligation reactions conducted resulted in two constructs: pETDuet-1_VP2 (7941bp) and pETDuet-1_VP2/6 (9101bp).

Only two possible pETDuet-1_VP2 (colonies 3 and 4) and two possible pETDuet-1_VP2/6 (colonies 6 and 16) plasmids were identified based on agarose gel electrophoresis of the mini preparation of plasmid DNA (results not shown). Therefore, only these possible pETDuet-1_VP2 and pETDuet-1_VP2/6 plasmids were subjected to a double restriction endonuclease digestion with the enzymes BglII and AvrII, since these were the enzymes used for cloning. A 10 µl sample of each 25 µl double restriction digestion was analysed by 1% agarose gel electrophoresis, as shown in Figure 2.16. A plasmid was considered positive if two fragments (5292bp and 2649bp) were visible for pETDuet-1_VP2 and two fragments (6452bp and 2649bp) for pETDuet-1_VP2/6, after the double restriction enzyme digestion.

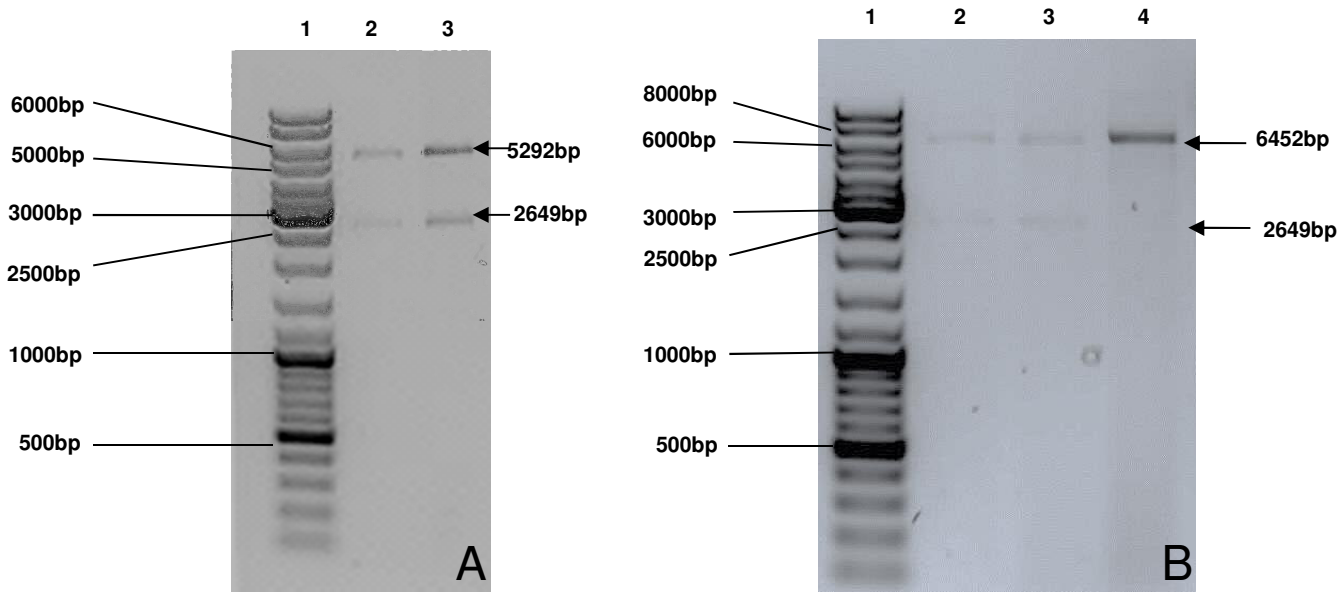


**Figure 2.16: Analysis by agarose gel electrophoresis of restriction enzyme analysis of pETDuet_VP2 and pETDuet_VP2/6 plasmids, from the cloning experiment. (A)** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 25 µl BglII and AvrII digestion reaction was loaded as follows; lanes: 2) pETDuet_VP2 colony 3; 3) pETDuet_VP2 colony 4. (**B)** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 25 µl BglII and AvrII digestion reaction was loaded as follows; lanes: 2) pETDuet_VP2/6 colony 6; 3) pETDuet_VP2/6 colony 16; 4) Digested pETDuet-1_VP6.

Both colonies from the restriction digestion of both pETDuet-1_VP2 (colonies 3 and 4) and pETDuet-1_VP2/6 (colony 6 and 16) gave the expected results (Figure 2.16A and B). A midi plasmid preparation of plasmid DNA was done for both positive colonies of both pETDuet-1_VP2 and pETDuet-1_VP2/6, as described in section 2.2.3.1. A 10 µl sample of each 400 µl plasmid preparation of both pETDuet-1_VP2 and pETDuet-1_VP2/6 colonies was analysed by a 1% agarose gel, as shown in Figure 2.17.

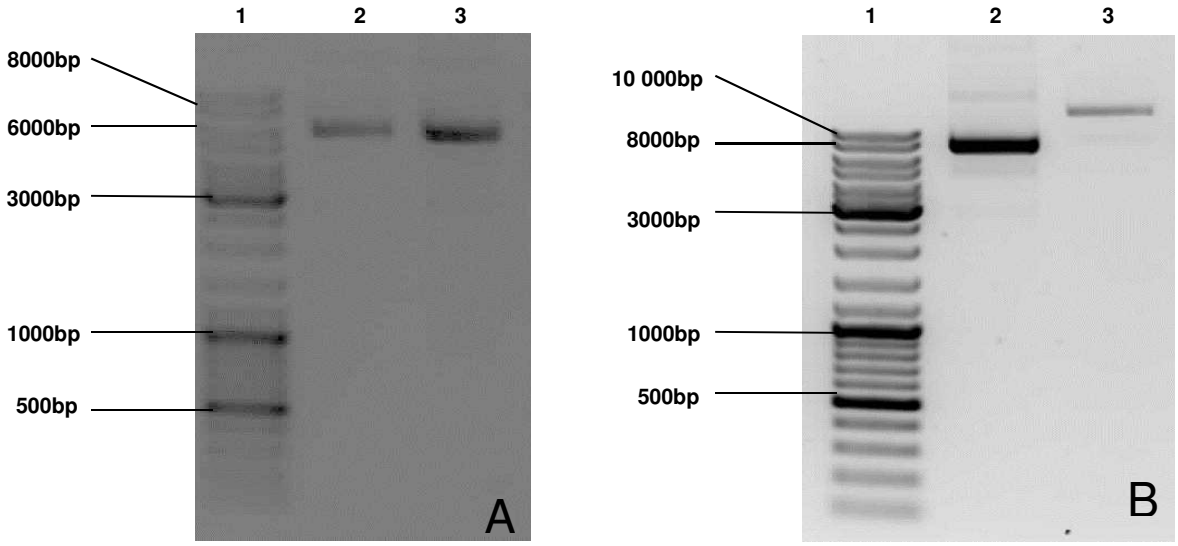


**Figure 2.17: Analysis by agarose gel electrophoresis of midi plasmid preparation of pETDuet-1_VP2 and pETDuet-1_VP2/6 colonies.** A 10 µl volume of 400 µl plasmid extraction was loaded. (**A)** Lanes: 1) 10 000bp O'Generuler DNA marker (Fermentas); 2) pETDuet-1_VP2 colony 3; 3) pETDuet-1_VP2 colony 4. (**B)** Lanes: 1) 10 000bp O'Generuler DNA marker (Fermentas); 2) pETDuet-1_VP2/6 colony 6; 3) pETDuet-1_VP2/6 colony 16.

The results in Figure 2.17 showed that plasmid DNA was obtained for both pETDuet-1_VP2 (Colony 3 and 4) and pETDuet-1_VP2/6 colonies (Colonies 6 and 16). The difference in the fragments obtained for the pETDuet-1_VP2/6 colonies can possibly be due to the fact that a mutation has occurred during the cloning procedure. However, to make sure that the midi-plasmid preparation of pETDuet-1_VP2/6 was successful both colonies for pETDuet-1_VP2/6, were subjected to a restriction digestion with AfeI. This enzyme was chosen to determine whether or not the ORF encoding VP2 was cloned into the expression vector pETDuet-1_VP6. A 10 µl sample of each 25 µl double restriction digestion was analysed on

a 1% agarose gel (Figure 2.18). Three fragments (1713bp, 2700bp and 4688bp) were expected for pETDuet-1_VP2/6, after restriction digestion.

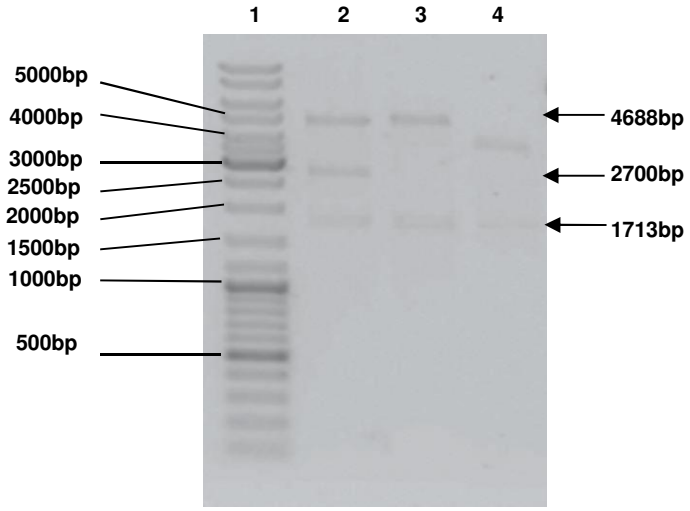


**Figure 2.18: Analysis by agarose gel electrophoresis of restriction enzyme digestion of pETDuet-1_VP2/6 plasmids, from midi plasmid preparation.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 25 µl AfeI digestion reaction was loaded as follows; lanes: 2) pETDuet-1_VP2/6 colony 6; 3) pETDuet-1_VP2/6 colony 16; 4) Digested pETDuet-1_VP6.

Figure 2.18 shows that the expected results were obtained for only one pETDuet-1_VP2/6 colony 6 (lane 2) namely three fragments, at 4688bp, 2700bp and other at 1713bp. This indicates that the ORF encoding VP2 was cloned into pETDuet-1_VP6. The pETDuet-1_VP2/6 colony 16 only showed two bands indicating that the ORF encoding VP2 was probably cloned into pETDuet-1_VP6, but a mutation could have occurred at the AfeI restriction site. These results were encouraging for the pETDuet-1_VP2/6 construct, but sequencing was still necessary to ensure that the plasmid contained the coding regions of interest. The midi plasmid preparation of both pETDuet-1_VP2 colonies (colony 3 and colony 4) and pETDuet-_VP2/6 colony 6 were sent for sequencing, to ensure that the ORF encoding VP2 was cloned correctly into both pETDuet-1_VP2 and pETDuet-1_VP2/6 and that no mutations had occurred during cloning procedures. Only the sequence results of pETDuet-1_VP2 colony 4 are shown in Figure 2.19 and of pETDuet-1_VP2/6 colony 6 in Figure 2.20.

```
pETDuet-1_VP2      GCATAATCGAAATTAATACGACTCACTATAGGGGAATTGT    240
pETDuet-1_VP2_F    ............TAATACGACTCACTATAGGGGAATTGT     27
pETDuet-1_VP2_Int  ........................................      0
pETDuet-1_VP2_R    ........................................      0

pETDuet-1_VP2      GAGCGGATAACAATTCCCCATCTTAGTATATTAGTTAAGT    280
pETDuet-1_VP2_F    GAGCGGATAACAATTCCCCATCTTAGTATATTAGTTAAGT     67
pETDuet-1_VP2_Int  ........................................      0
pETDuet-1_VP2_R    ........................................      0

pETDuet-1_VP2      ATAAGAAGGAGATATACATATGGCAGATCTATGGCGTATC    320
pETDuet-1_VP2_F    ATAAGAAGGAGATATACATATGGCAGATCTATGGCGTATC    107
pETDuet-1_VP2_Int  ........................................      0
pETDuet-1_VP2_R    ........................................      0

pETDuet-1_VP2      GTAAACGTGGCGCGCGTCGTGAAGCGAACCTGAACAACAA    360
pETDuet-1_VP2_F    GTAAACGTGGCGCGCGTCGTGAAGCGAACCTGAACAACAA    147
pETDuet-1_VP2_Int  ........................................      0
pETDuet-1_VP2_R    ........................................      0

pETDuet-1_VP2      CGATCGTATGCAGGAAAAAATCGACGAAAAACAGGATAGC    400
pETDuet-1_VP2_F    CGATCGTATGCAGGAAAAAATCGACGAAAAACAGGATAGC    187
pETDuet-1_VP2_Int  ........................................      0
pETDuet-1_VP2_R    ........................................      0

pETDuet-1_VP2      AACAAAATTCAGCTGTCTGATAAAGTTCTGAGCAAAAAAG    440
pETDuet-1_VP2_F    AACAAAATTCAGCTGTCTGATAAAGTTCTGAGCAAAAAAG    227
pETDuet-1_VP2_Int  ........................................      0
pETDuet-1_VP2_R    ........................................      0

pETDuet-1_VP2      AAGAAATCGTTACCGATAGCCACGAAGAAGTGAAAGTTAC    480
pETDuet-1_VP2_F    AAGAAATCGTTACCGATAGCCACGAAGAAGTGAAAGTTAC    267
pETDuet-1_VP2_Int  ........................................      0
pETDuet-1_VP2_R    ........................................      0

pETDuet-1_VP2      CGACGAACTGAAAAAAAGCACCAAAGAAGAAAGCAAACAG    520
pETDuet-1_VP2_F    CGACGAACTGAAAAAAAGCACCAAAGAAGAAAGCAAACAG    307
pETDuet-1_VP2_Int  ........................................      0
pETDuet-1_VP2_R    ........................................      0

pETDuet-1_VP2      CTGCTGGAAGTGCTGAAAACGAAAGAAGAGCACCAGAAAG    560
pETDuet-1_VP2_F    CTGCTGGAAGTGCTGAAAACGAAAGAAGAGCACCAGAAAG    347
pETDuet-1_VP2_Int  ........................................      0
pETDuet-1_VP2_R    ........................................      0

pETDuet-1_VP2      AAATCCAGTATGAAATCCTGCAGAAAACCATTCCGACCTT    600
pETDuet-1_VP2_F    AAATCCAGTATGAAATCCTGCAGAAAACCATTCCGACCTT    387
pETDuet-1_VP2_Int  ........................................      0
pETDuet-1_VP2_R    ........................................      0
```

```
pETDuet-1_VP2      TGAACCGAAAGAAACCATTCTGCGCAAACTGGAAGATATT    640
pETDuet-1_VP2_F    TGAACCGAAAGAAACCATTCTGCGCAAACTGGAAGATATT    427
pETDuet-1_VP2_Int ....................................        0
pETDuet-1_VP2_R    ....................................        0

pETDuet-1_VP2      CAGCCGGAACTGGCCAAAAAACAGACCAAACTGTTCCGTA    680
pETDuet-1_VP2_F    CAGCCGGAACTGGCCAAAAAACAGACCAAACTGTTCCGTA    467
pETDuet-1_VP2_Int ....................................        0
pETDuet-1_VP2_R    ....................................        0

pETDuet-1_VP2      TTTTCGAACCGAAACAGCTGCCGATTTATCGTGCGAACGG    720
pETDuet-1_VP2_F    TTTTCGAACCGAAACAGCTGCCGATTTATCGTGCGAACGG    507
pETDuet-1_VP2_Int ....................................        0
pETDuet-1_VP2_R    ....................................        0

pETDuet-1_VP2      CGAACGTGAACTGCGTAACCGTTGGTACTGGAAACTGAAA    760
pETDuet-1_VP2_F    CGAACGTGAACTGCGTAACCGTTGGTACTGGAAACTGAAA    547
pETDuet-1_VP2_Int ....................................        0
pETDuet-1_VP2_R    ....................................        0

pETDuet-1_VP2      AAAGATACCCTGCCGGATGGCGATTATGATGTGCGCGAAT    800
pETDuet-1_VP2_F    AAAGATACCCTGCCGGATGGCGATTATGATGTGCGCGAAT    587
pETDuet-1_VP2_Int ....................................        0
pETDuet-1_VP2_R    ....................................        0

pETDuet-1_VP2      ATTTCCTGAACCTGTATGATCAGGTGCTGACCGAAATGCC    840
pETDuet-1_VP2_F    ATTTCCTGAACCTGTATGATCAGGTGCTGACCGAAATGCC    627
pETDuet-1_VP2_Int ....................................        0
pETDuet-1_VP2_R    ....................................        0

pETDuet-1_VP2      GGATTATCTGCTGCTGAAAGATATGGCGGTGGAAAATAAA    880
pETDuet-1_VP2_F    GGATTATCTGCTGCTGAAAGATATGGCGGTGGAAAATAAA    667
pETDuet-1_VP2_Int ....................................        0
pETDuet-1_VP2_R    ....................................        0

pETDuet-1_VP2      AACAGCCGTGATGCGGGCAAAGTGGTGGATAGCGAAACCG    920
pETDuet-1_VP2_F    AACAGCCGTGATGCGGGCAAAGTGGTGGATAGCGAAACCG    707
pETDuet-1_VP2_Int ....................................        0
pETDuet-1_VP2_R    ....................................        0

pETDuet-1_VP2      CGAGCATTTGTGATGCGATCTTCCAGGATGAAGAAACCGA    960
pETDuet-1_VP2_F    CGAGCATTTGTGATGCGATCTTCCAGGATGAAGAAACCGA    747
pETDuet-1_VP2_Int ....................................        0
pETDuet-1_VP2_R    ....................................        0

pETDuet-1_VP2      AGGCGCGGTGCGTCGTTTTATTGCGGAAATGCGTCAGCGT   1000
pETDuet-1_VP2_F    AGGCGCGGTGCGTCGTTTTATTGCGGAAATGCGTCAGCGT    787
pETDuet-1_VP2_Int ....................................        0
pETDuet-1_VP2_R    ....................................        0
```

```
pETDuet-1_VP2      GTGCAGGCGGATCGTAACGTGGTGAACTATCCGAGCATTC   1040
pETDuet-1_VP2_F    GTGCAGGCGGATCGTAACGTGGTGAACTATCCGAGCATTC    827
pETDuet-1_VP2_Int  ....................................      0
pETDuet-1_VP2_R    ....................................      0


pETDuet-1_VP2      TGCATCCGATTGATTATGCGTTCAACGAATACTTTCTGCA   1080
pETDuet-1_VP2_F    TGCATCCGATTGATTATGCGTTCAACGAATACTTTCTGCA    867
pETDuet-1_VP2_Int  ....................................      0
pETDuet-1_VP2_R    ....................................      0


pETDuet-1_VP2      GCATCAGCTGGTTGAACCGCTGAACAACGATATCATCTTC   1120
pETDuet-1_VP2_F    GCATCAGCTGGTTGAACCGCTGAACAACGATATCATCTTC    907
pETDuet-1_VP2_Int  ....................................      0
pETDuet-1_VP2_R    ....................................      0


pETDuet-1_VP2      AACTATATCCCGGAACGTATTCGTAACGATGTGAACTACA   1160
pETDuet-1_VP2_F    AACTATATCCCGGAACGTATTCGTAACGATGTGAACTACA    947
pETDuet-1_VP2_Int  ....................................      0
pETDuet-1_VP2_R    ....................................      0



pETDuet-1_VP2      TCCTGAACATGGATCGCAACCTGCCGAGCACCGCGCGTTA   1200
pETDuet-1_VP2_F    TCCTGAACATGGATCGCAACCTGCCGAGCACCGCGCGTTA    987
pETDuet-1_VP2_Int  ....................................      0
pETDuet-1_VP2_R    ....................................      0


pETDuet-1_VP2      TATCCGTCCGAACCTGCTGCAGGATCGTCTGAACCTGCAT   1240
pETDuet-1_VP2_F    TATCCGTCCGAACCTGCTGCAGGATCGTCTGAACCTGCAT   1027
pETDuet-1_VP2_Int  ....................................      0
pETDuet-1_VP2_R    ....................................      0


pETDuet-1_VP2      GATAACTTCGAAAGCCTGTGGGATACCATTACCACGAGCA   1280
pETDuet-1_VP2_F    GATAACTTCGAAAGCCTGTGGGATACCATTACCACGAGCA   1067
pETDuet-1_VP2_Int  ....................................      0
pETDuet-1_VP2_R    ....................................      0


pETDuet-1_VP2      ACTATATTCTGGCCCGTAGCGTGGTGCCGGATCTGAAAGA   1320
pETDuet-1_VP2_F    ACTATATTCTGGCCCGTAGCGTGGTGCCGGATCTGAAAGA   1107
pETDuet-1_VP2_Int  ....................................      0
pETDuet-1_VP2_R    ....................................      0


pETDuet-1_VP2      ACTGGTGAGCACCGAAGCGCAGATTCAGAAAATGAGCCAG   1360
pETDuet-1_VP2_F    ACTGGTGAGCACCGAAGCGCAGATTCAGAAAATGAGCCAG   1147
pETDuet-1_VP2_Int  ....................................      0
pETDuet-1_VP2_R    ....................................      0


pETDuet-1_VP2      GATCTGCAGCTGGAAGCGCTGACCATTCAGAGCGAAACCC   1400
pETDuet-1_VP2_F    GATCTGCAGCTGGAAGCGCTGACCATTCAGAGCGAAACCC   1186
pETDuet-1_VP2_Int  ....................................      0
pETDuet-1_VP2_R    ....................................      0
```

```
pETDuet-1_VP2     AGTTTCTGACCGGCATTAACAGCCAGGCGGCGAACGATTG   1440
pETDuet-1_VP2_F   AGTTTCTGACCGGCATTAACAGCCAGGCGGCGAACGATTG   1225
pETDuet-1_VP2_Int ....................................       0
pETDuet-1_VP2_R   ....................................       0

pETDuet-1_VP2     CTTTAAAACCCTGATTGCGGCCATGCTGTCTCAGCGTACC   1480
pETDuet-1_VP2_F   CTTTAAAACCCTGATTGCG.....................   1244
pETDuet-1_VP2_Int ..................GGCCATGCTGTCTCAGCGTACC     22
pETDuet-1_VP2_R   ....................................       0

pETDuet-1_VP2     ATGAGCCTGGATTTTGTGACCACCAACTATATGAGCCTGA   1520
pETDuet-1_VP2_F   ....................................       1244
pETDuet-1_VP2_Int ATGAGCCTGGATTTTGTGACCACCAACTATATGAGCCTGA     62
pETDuet-1_VP2_R   ....................................       0

pETDuet-1_VP2     TTAGCGGCATGTGGCTGCTGACCGTGGTGCCGAACGACAT   1560
pETDuet-1_VP2_F   ....................................       1244
pETDuet-1_VP2_Int TTAGCGGCATGTGGCTGCTGACCGTGGTGCCGAACGACAT    102
pETDuet-1_VP2_R   ....................................       0

pETDuet-1_VP2     GTTTATTCGTGAAAGCCTGGTGGCGTGCCAGCTGGCCATT   1600
pETDuet-1_VP2_F   ....................................       1244
pETDuet-1_VP2_Int GTTTATTCGTGAAAGCCTGGTGGCGTGCCAGCTGGCCATT    142
pETDuet-1_VP2_R   ....................................       0

pETDuet-1_VP2     GTGAACACCATTATCTATCCGGCGTTTGGCATGCAGCGTA   1640
pETDuet-1_VP2_F   ....................................       1244
pETDuet-1_VP2_Int GTGAACACCATTATCTATCCGGCGTTTGGCATGCAGCGTA    182
pETDuet-1_VP2_R   ....................................       0

pETDuet-1_VP2     TGCATTATCGTAACGGCGATCCGCAGACCCCGTTTCAGAT   1680
pETDuet-1_VP2_F   ....................................       1244
pETDuet-1_VP2_Int TGCATTATCGTAACGGCGATCCGCAGACCCCGTTTCAGAT    222
pETDuet-1_VP2_R   ....................................       0

pETDuet-1_VP2     TGCGGAACAGCAGATCCAGAACTTTCAGGTGGCGAACTGG   1720
pETDuet-1_VP2_F   ....................................       1244
pETDuet-1_VP2_Int TGCGGAACAGCAGATCCAGAACTTTCAGGTGGCGAACTGG    262
pETDuet-1_VP2_R   ....................................       0

pETDuet-1_VP2     CTGCATTTTGTGAACAACAACCAGTTTCGTCAGGCGGTGA   1760
pETDuet-1_VP2_F   ....................................       1244
pETDuet-1_VP2_Int CTGCATTTTGTGAACAACAACCAGTTTCGTCAGGCGGTGA    302
pETDuet-1_VP2_R   ....................................       0

pETDuet-1_VP2     TTGATGGCGTGCTGAACCAGGTGCTGAACGATAACATTCG   1800
pETDuet-1_VP2_F   ....................................       1244
pETDuet-1_VP2_Int TTGATGGCGTGCTGAACCAGGTGCTGAACGATAACATTCG    342
pETDuet-1_VP2_R   ....................................       0
```

```
pETDuet-1_VP2     TAACGGCCATGTGATTAACCAACTGATGGAAGCCCTGATG   1840
pETDuet-1_VP2_F   .....................................   1244
pETDuet-1_VP2_Int TAACGGCCATGTGATTAACCAACTGATGGAAGCCCTGATG    382
pETDuet-1_VP2_R   .....................................      0


pETDuet-1_VP2     CAGCTGTCTCGTCAGCAGTTTCCGACCATGCCGATCGATT   1880
pETDuet-1_VP2_F   .....................................   1244
pETDuet-1_VP2_Int CAGCTGTCTCGTCAGCAGTTTCCGACCATGCCGATCGATT    422
pETDuet-1_VP2_R   .....................................      0


pETDuet-1_VP2     ATAAACGTAGCATTCAGCGTGGCATTCTGCTGCTGTCTAA   1920
pETDuet-1_VP2_F   .....................................   1244
pETDuet-1_VP2_Int ATAAACGTAGCATTCAGCGTGGCATTCTGCTGCTGTCTAA    462
pETDuet-1_VP2_R   .....................................      0


pETDuet-1_VP2     CCGTCTGGGCCAGCTGGTTGATCTGACCCGTCTGCTGGCC   1960
pETDuet-1_VP2_F   .....................................   1244
pETDuet-1_VP2_Int CCGTCTGGGCCAGCTGGTTGATCTGACCCGTCTGCTGGCC    502
pETDuet-1_VP2_R   .....................................      0


pETDuet-1_VP2     TATAACTATGAAACGCTGATGGCATGCATCACGATGAACA   2000
pETDuet-1_VP2_F   .....................................   1244
pETDuet-1_VP2_Int TATAACTATGAAACGCTGATGGCATGCATCACGATGAACA    542
pETDuet-1_VP2_R   ...................................ACA      3


pETDuet-1_VP2     TGCAGCATGTGCAGACCCTGACCACCGAAAAACTGCAGCT   2040
pETDuet-1_VP2_F   .....................................   1244
pETDuet-1_VP2_Int TGCAGCATGTGCAGACCCTGACCACCGAAAAACTGCAGCT    582
pETDuet-1_VP2_R   TGCAGCATGTGCAGACCCTGACCACCGAAAAACTGCAGCT     43


pETDuet-1_VP2     GACCAGCGTGACCAGCCTGTGCATGCTGATTGGCAACGCG   2080
pETDuet-1_VP2_F   .....................................   1244
pETDuet-1_VP2_Int GACCAGCGTGACCAGCCTGTGCATGCTGATTGGCAACGCG    622
pETDuet-1_VP2_R   GACCAGCGTGACCAGCCTGTGCATGCTGATTGGCAACGCG     83


pETDuet-1_VP2     ACCGTGATTCCGAGCCCGCAGACCCTGTTCCATTATTATA   2120
pETDuet-1_VP2_F   .....................................   1244
pETDuet-1_VP2_Int ACCGTGATTCCGAGCCCGCAGACCCTGTTCCATTATTATA    662
pETDuet-1_VP2_R   ACCGTGATTCCGAGCCCGCAGACCCTGTTCCATTATTATA    123


pETDuet-1_VP2     ACGTGAACGTGAACTTTCACAGCAACTATAACGAACGTAT   2160
pETDuet-1_VP2_F   .....................................   1244
pETDuet-1_VP2_Int ACGTGAACGTGAACTTTCACAGCAACTATAACGAACGTAT    702
pETDuet-1_VP2_R   ACGTGAACGTGAACTTTCACAGCAACTATAACGAACGTAT    163


pETDuet-1_VP2     TAACGATGCGGTGGCGATTATTACCGCGGCGAACCGTCTG   2200
pETDuet-1_VP2_F   .....................................   1244
pETDuet-1_VP2_Int TAACGATGCGGTGGCGATTATTACCGCGGCGAACCGTCTG    742
pETDuet-1_VP2_R   TAACGATGCGGTGGCGATTATTACCGCGGCGAACCGTCTG    203
```

```
pETDuet-1_VP2     AATCTGTATCAGAAAAAAATGAAAGCGATCGTGGAAGATT   2240
pETDuet-1_VP2_F   ....................................   1244
pETDuet-1_VP2_Int AATCTGTATCAGAAAAAAATGAAAGCGATCGTGGAAGATT    782
pETDuet-1_VP2_R   AATCTGTATCAGAAAAAAATGAAAGCGATCGTGGAAGATT    243


pETDuet-1_VP2     TTCTGAAACGCCTGTACATTTTTGATGTGAGCCGTGTGCC   2280
pETDuet-1_VP2_F   ....................................   1244
pETDuet-1_VP2_Int TTCTGAAACGCCTGTACATTTTTGATGTGAGCCGTGTGCC    822
pETDuet-1_VP2_R   TTCTGAAACGCCTGTACATTTTTGATGTGAGCCGTGTGCC    283


pETDuet-1_VP2     GGATGATCAGATGTATCGTCTGCGTGATCGTCTGCGTCTG   2320
pETDuet-1_VP2_F   ....................................   1244
pETDuet-1_VP2_Int GGATGATCAGATGTATCGTCTGCGTGATCGTCTGCGTCTG    862
pETDuet-1_VP2_R   GGATGATCAGATGTATCGTCTGCGTGATCGTCTGCGTCTG    323


pETDuet-1_VP2     CTGCCGGTGGAAATTCGTCGTCTGGATATCTTTAACCTGA   2360
pETDuet-1_VP2_F   ....................................   1244
pETDuet-1_VP2_Int CTGCCGGTGGAAATTCGTCGTCTGGATATCTTTAACCTGA    902
pETDuet-1_VP2_R   CTGCCGGTGGAAATTCGTCGTCTGGATATCTTTAACCTGA    363


pETDuet-1_VP2     TCCTGATGAATATGGATCAGATCGAACGTGCGAGCGATAA   2400
pETDuet-1_VP2_F   ....................................   1244
pETDuet-1_VP2_Int TCCTGATGAATATGGATCAGATCGAACGTGCGAGCGATAA    942
pETDuet-1_VP2_R   TCCTGATGAATATGGATCAGATCGAACGTGCGAGCGATAA    403


pETDuet-1_VP2     AATTGCGCAGGGCGTGATTATTGCGTATCGTGATATGCAT   2440
pETDuet-1_VP2_F   ....................................   1244
pETDuet-1_VP2_Int AATTGCGCAGGGCGTGATTATTGCGTATCGTGATATGCAT    982
pETDuet-1_VP2_R   AATTGCGCAGGGCGTGATTATTGCGTATCGTGATATGCAT    443


pETDuet-1_VP2     CTGGAACGTGATGAAATGTATGGCTACGTGAACATTGCGC   2480
pETDuet-1_VP2_F   ....................................   1244
pETDuet-1_VP2_Int CTGGAACGTGATGAAATGTATGGCTACGTGAACATTGCGC   1022
pETDuet-1_VP2_R   CTGGAACGTGATGAAATGTATGGCTACGTGAACATTGCGC    483


pETDuet-1_VP2     GTAACCTGGAAGGCTTTCAGCAGATTAACCTGGAAGAACT   2520
pETDuet-1_VP2_F   ....................................   1244
pETDuet-1_VP2_Int GTAACCTGGAAGGCTTTCAGCAGATTAACCTGGAAGAACT   1062
pETDuet-1_VP2_R   GTAACCTGGAAGGCTTTCAGCAGATTAACCTGGAAGAACT    523


pETDuet-1_VP2     GATGCGTAGCGGCGATTATGCGCAGATTACCAACATGCTG   2560
pETDuet-1_VP2_F   ....................................   1244
pETDuet-1_VP2_Int GATGCGTAGCGGCGATTATGCGCAGATTACCAACATGCTG   1102
pETDuet-1_VP2_R   GATGCGTAGCGGCGATTATGCGCAGATTACCAACATGCTG    563


pETDuet-1_VP2     CTGAACAACCAGCCGGTGGCGCTGGTTGGTGCGCTGCCGT   2600
pETDuet-1_VP2_F   ....................................   1244
pETDuet-1_VP2_Int CTGAACAACCAGCCGGTGGCGCTGGTTGGTGCGCTGCCGT   1142
pETDuet-1_VP2_R   CTGAACAACCAGCCGGTGGCGCTGGTTGGTGCGCTGCCGT    603
```

```
pETDuet-1_VP2      TTATTACCGATAGCAGCGTGATTAGCCTGATTGCGAAACT   2640
pETDuet-1_VP2_F    ......................................    1244
pETDuet-1_VP2_Int  TTATT................................    1147
pETDuet-1_VP2_R    TTATTACCGATAGCAGCGTGATTAGCCTGATTGCGAAACT    643

pETDuet-1_VP2      GGATGCGACCGTGTTTGCGCAGATTGTGAAACTGCGTAAA   2680
pETDuet-1_VP2_F    ......................................    1244
pETDuet-1_VP2_Int  ......................................    1147
pETDuet-1_VP2_R    GGATGCGACCGTGTTTGCGCAGATTGTGAAACTGCGTAAA    683

pETDuet-1_VP2      GTGGATACCCTGAAACCGATCCTGTATAAAATCAACAGCG   2720
pETDuet-1_VP2_F    ......................................    1244
pETDuet-1_VP2_Int  ......................................    1147
pETDuet-1_VP2_R    GTGGATACCCTGAAACCGATCCTGTATAAAATCAACAGCG    723

pETDuet-1_VP2      ATAGCAACGATTTTTATCTGGTGGCGAACTATGATTGGGT   2760
pETDuet-1_VP2_F    ......................................    1244
pETDuet-1_VP2_Int  ......................................    1147
pETDuet-1_VP2_R    ATAGCAACGATTTTTATCTGGTGGCGAACTATGATTGGGT    763

pETDuet-1_VP2      GCCGACCAGCACCACCAAAGTGTATAAACAGGTGCCGCAG   2800
pETDuet-1_VP2_F    ......................................    1244
pETDuet-1_VP2_Int  ......................................    1147
pETDuet-1_VP2_R    GCCGACCAGCACCACCAAAGTGTATAAACAGGTGCCGCAG    803

pETDuet-1_VP2      CAGTTTGATTTTCGTAACTCTATGCACATGCTGACCAGCA   2840
pETDuet-1_VP2_F    ......................................    1244
pETDuet-1_VP2_Int  ......................................    1147
pETDuet-1_VP2_R    CAGTTTGATTTTCGTAACTCTATGCACATGCTGACCAGCA    843

pETDuet-1_VP2      ACCTGACCTTTACCGTGTATAGCGATCTGCTGGCCTTTGT   2880
pETDuet-1_VP2_F    ......................................    1244
pETDuet-1_VP2_Int  ......................................    1147
pETDuet-1_VP2_R    ACCTGACCTTTACCGTGTATAGCGATCTGCTGGCCTTTGT    883

pETDuet-1_VP2      GAGCGCGGATACCGTGGAACCGATTAACGCGGTGGCGTTT   2920
pETDuet-1_VP2_F    ......................................    1244
pETDuet-1_VP2_Int  ......................................    1147
pETDuet-1_VP2_R    GAGCGCGGATACCGTGGAACCGATTAACGCGGTGGCGTTT    923

pETDuet-1_VP2      GATAACATGCGCATCATGAACGAACTGTAATAACCTAGGC   2960
pETDuet-1_VP2_F    ......................................    1244
pETDuet-1_VP2_Int  ......................................    1147
pETDuet-1_VP2_R    GATAACATGCGCATCATGAACGAACTGTAATAACCTAGGC    963

pETDuet-1_VP2      TGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGG   3000
pETDuet-1_VP2_F    ......................................    1244
pETDuet-1_VP2_Int  ......................................    1147
pETDuet-1_VP2_R    TGCTG....................................    968
```

```
pETDuet-1_VP2     GCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAG   3040
pETDuet-1_VP2_F   ..........................................   1244
pETDuet-1_VP2_Int ..........................................   1147
pETDuet-1_VP2_R   ..........................................    968
```

**Figure 2.19: Nucleotide sequence alignments of the recombinant pETDuet-1 vector containing the ORF encoding VP2 with the *in silico* clone of pETDuet-1_VP2.** The *in silico* clone of pETDuet-1_VP2 is indicated as pETDuet-1_VP2, the forward sequence is indicated as pETDuet1_VP2_F nucleotide sequence, the reverse sequence is indicated as pETDuet1_VP2_R nucleotide sequence and the internal sequence is indicated as pETDuet1_VP2_Int nucleotide sequence. The restriction enzymes used for the amplification of genome segment 2 are indicated in blue. The start codon of genome segment 2 is indicated in green and the stop codon is indicated in red. The T7 promoter-1 is indicated in purple, the T7 transcription start-1 is indicated in dark blue and the *lac* operator in yellow.

```
pETDuet1_VP2/6    GCATAATCGAAATTAATACGACTCACTATAGGGGAATTGT   1400
pETDuet1_VP2/6_F  ..............ATACGACTCACTATAGGGGAATTGT     25
pETDuet-1_VP2_Int ........................................      0
pETDuet1_VP2/6_R  ........................................      0

pETDuet1_VP2/6    GAGCGGATAACAATTCCCCATCTTAGTATATTAGTTAAGT   1440
pETDuet1_VP2/6_F  GAGCGGATAACAATTCCCCATCTTAGTATATTAGTTAAGT     65
pETDuet-1_VP2_Int ........................................      0
pETDuet1_VP2/6_R  ........................................      0

pETDuet1_VP2/6    ATAAGAAGGAGATATACATATGGCAGATCTATGGCGTATC   1480
pETDuet1_VP2/6_F  ATAAGAAGGAGATATACATATGGCAGATCTATGGCGTATC    105
pETDuet-1_VP2_Int ........................................      0
pETDuet1_VP2/6_R  ........................................      0

pETDuet1_VP2/6    GTAAACGTGGCGCGCGTCGTGAAGCGAACCTGAACAACAA   1520
pETDuet1_VP2/6_F  GTAAACGTGGCGCGCGTCGTGAAGCGAACCTGAACAACAA    145
pETDuet-1_VP2_Int ........................................      0
pETDuet1_VP2/6_R  ........................................      0

pETDuet1_VP2/6    CGATCGTATGCAGGAAAAAATCGACGAAAAACAGGATAGC   1560
pETDuet1_VP2/6_F  CGATCGTATGCAGGAAAAAATCGACGAAAAACAGGATAGC    185
pETDuet-1_VP2_Int ........................................      0
pETDuet1_VP2/6_R  ........................................      0

pETDuet1_VP2/6    AACAAAATTCAGCTGTCTGATAAAGTTCTGAGCAAAAAAG   1600
pETDuet1_VP2/6_F  AACAAAATTCAGCTGTCTGATAAAGTTCTGAGCAAAAAAG    225
pETDuet-1_VP2_Int ........................................      0
pETDuet1_VP2/6_R  ........................................      0

pETDuet1_VP2/6    AAGAAATCGTTACCGATAGCCACGAAGAAGTGAAAGTTAC   1640
pETDuet1_VP2/6_F  AAGAAATCGTTACCGATAGCCACGAAGAAGTGAAAGTTAC    265
pETDuet-1_VP2_Int ........................................      0
pETDuet1_VP2/6_R  ........................................      0

pETDuet1_VP2/6    CGACGAACTGAAAAAAAGCACCAAAGAAGAAAGCAAACAG   1680
pETDuet1_VP2/6_F  CGACGAACTGAAAAAAAGCACCAAAGAAGAAAGCAAACAG    305
pETDuet-1_VP2_Int ........................................      0
pETDuet1_VP2/6_R  ........................................      0

pETDuet1_VP2/6    CTGCTGGAAGTGCTGAAAACGAAAGAAGAGCACCAGAAAG   1720
pETDuet1_VP2/6_F  CTGCTGGAAGTGCTGAAAACGAAAGAAGAGCACCAGAAAG    345
pETDuet-1_VP2_Int ........................................      0
pETDuet1_VP2/6_R  ........................................      0

pETDuet1_VP2/6    AAATCCAGTATGAAATCCTGCAGAAAACCATTCCGACCTT   1760
pETDuet1_VP2/6_F  AAATCCAGTATGAAATCCTGCAGAAAACCATTCCGACCTT    385
pETDuet-1_VP2_Int ........................................      0
pETDuet1_VP2/6_R  ........................................      0
```

```
pETDuet1_VP2/6    TGAACCGAAAGAAACCATTCTGCGCAAACTGGAAGATATT   1800
pETDuet1_VP2/6_F  TGAACCGAAAGAAACCATTCTGCGCAAACTGGAAGATATT    425
pETDuet-1_VP2_Int ....................................       0
pETDuet1_VP2/6_R  ....................................       0


pETDuet1_VP2/6    CAGCCGGAACTGGCCAAAAAACAGACCAAACTGTTCCGTA   1840
pETDuet1_VP2/6_F  CAGCCGGAACTGGCCAAAAAACAGACCAAACTGTTCCGTA    465
pETDuet-1_VP2_Int ....................................       0
pETDuet1_VP2/6_R  ....................................       0


pETDuet1_VP2/6    TTTTCGAACCGAAACAGCTGCCGATTTATCGTGCGAACGG   1880
pETDuet1_VP2/6_F  TTTTCGAACCGAAACAGCTGCCGATTTATCGTGCGAACGG    505
pETDuet-1_VP2_Int ....................................       0
pETDuet1_VP2/6_R  ....................................       0


pETDuet1_VP2/6    CGAACGTGAACTGCGTAACCGTTGGTACTGGAAACTGAAA   1920
pETDuet1_VP2/6_F  CGAACGTGAACTGCGTAACCGTTGGTACTGGAAACTGAAA    545
pETDuet-1_VP2_Int ....................................       0
pETDuet1_VP2/6_R  ....................................       0


pETDuet1_VP2/6    AAAGATACCCTGCCGGATGGCGATTATGATGTGCGCGAAT   1960
pETDuet1_VP2/6_F  AAAGATACCCTGCCGGATGGCGATTATGATGTGCGCGAAT    585
pETDuet-1_VP2_Int ....................................       0
pETDuet1_VP2/6_R  ....................................       0


pETDuet1_VP2/6    ATTTCCTGAACCTGTATGATCAGGTGCTGACCGAAATGCC   2000
pETDuet1_VP2/6_F  ATTTCCTGAACCTGTATGATCAGGTGCTGACCGAAATGCC    625
pETDuet-1_VP2_Int ....................................       0
pETDuet1_VP2/6_R  ....................................       0


pETDuet1_VP2/6    GGATTATCTGCTGCTGAAAGATATGGCGGTGGAAAATAAA   2040
pETDuet1_VP2/6_F  GGATTATCTGCTGCTGAAAGATATGGCGGTGGAAAATAAA    665
pETDuet-1_VP2_Int ....................................       0
pETDuet1_VP2/6_R  ....................................       0


pETDuet1_VP2/6    AACAGCCGTGATGCGGGCAAAGTGGTGGATAGCGAAACCG   2080
pETDuet1_VP2/6_F  AACAGCCGTGATGCGGGCAAAGTGGTGGATAGCGAAACCG    705
pETDuet-1_VP2_Int ....................................       0
pETDuet1_VP2/6_R  ....................................       0


pETDuet1_VP2/6    CGAGCATTTGTGATGCGATCTTCCAGGATGAAGAAACCGA   2120
pETDuet1_VP2/6_F  CGAGCATTTGTGATGCGATCTTCCAGGATGAAGAAACCGA    745
pETDuet-1_VP2_Int ....................................       0
pETDuet1_VP2/6_R  ....................................       0


pETDuet1_VP2/6    AGGCGCGGTGCGTCGTTTTATTGCGGAAATGCGTCAGCGT   2160
pETDuet1_VP2/6_F  AGGCGCGGTGCGTCGTTTTATTGCGGAAATGCGTCAGCGT    785
pETDuet-1_VP2_Int ....................................       0
pETDuet1_VP2/6_R  ....................................       0
```

```
pETDuet1_VP2/6    GTGCAGGCGGATCGTAACGTGGTGAACTATCCGAGCATTC   2200
pETDuet1_VP2/6_F  GTGCAGGCGGATCGTAACGTGGTGAACTATCCGAGCATTC    825
pETDuet-1_VP2_Int .....................................      0
pETDuet1_VP2/6_R  .....................................      0

pETDuet1_VP2/6    TGCATCCGATTGATTATGCGTTCAACGAATACTTTCTGCA   2240
pETDuet1_VP2/6_F  TGCATCCGATTGATTATGCGTTCAACGAATACTTTCTGCA    865
pETDuet-1_VP2_Int .....................................      0
pETDuet1_VP2/6_R  .....................................      0

pETDuet1_VP2/6    GCATCAGCTGGTTGAACCGCTGAACAACGATATCATCTTC   2280
pETDuet1_VP2/6_F  GCATCAGCTGGTTGAACCGCTGAACAACGATATCATCTTC    905
pETDuet-1_VP2_Int .....................................      0
pETDuet1_VP2/6_R  .....................................      0

pETDuet1_VP2/6    AACTATATCCCGGAACGTATTCGTAACGATGTGAACTACA   2320
pETDuet1_VP2/6_F  AACTATATCCCGGAACGTATTCGTAACGATGTGAACTACA    945
pETDuet-1_VP2_Int .....................................      0
pETDuet1_VP2/6_R  .....................................      0

pETDuet1_VP2/6    TCCTGAACATGGATCGCAACCTGCCGAGCACCGCGCGTTA   2360
pETDuet1_VP2/6_F  TCCTGAACATGGATCGCAACCTGCCGAGCACCGCGCGTTA    985
pETDuet-1_VP2_Int .....................................      0
pETDuet1_VP2/6_R  .....................................      0

pETDuet1_VP2/6    TATCCGTCCGAACCTGCTGCAGGATCGTCTGAACCTGCAT   2400
pETDuet1_VP2/6_F  TATCCGTCCGAACCTGCTGCAGGATCGTCTGAACCTGCAT   1025
pETDuet-1_VP2_Int .....................................      0
pETDuet1_VP2/6_R  .....................................      0

pETDuet1_VP2/6    GATAACTTCGAAAGCCTGTGGGATACCATTACCACGAGCA   2440
pETDuet1_VP2/6_F  GATAACTTCGAAAGCCTGTGGGATACCATTACCACGAGCA   1065
pETDuet-1_VP2_Int .....................................      0
pETDuet1_VP2/6_R  .....................................      0

pETDuet1_VP2/6    ACTATATTCTGGCCCGTAGCGTGGTGCCGGATCTGAAAGA   2480
pETDuet1_VP2/6_F  ACTATATTCTGGCCCGTAGCGTGGTGCCGGATCTGAAAGA   1103
pETDuet-1_VP2_Int .....................................      0
pETDuet1_VP2/6_R  .....................................      0

pETDuet1_VP2/6    ACTGGTGAGCACCGAAGCGCAGATTCAGAAAATGAGCCAG   2520
pETDuet1_VP2/6_F  ACTGGTGAGCACCGAAGCGCAGATTCAGAAAATGAGCCAG   1143
pETDuet-1_VP2_Int .....................................      0
pETDuet1_VP2/6_R  .....................................      0

pETDuet1_VP2/6    GATCTGCAGCTGGAAGCGCTGACCATTCAGAGCGAAACCC   2560
pETDuet1_VP2/6_F  GATCTGCAGCTGGAAGCGCTGACCATTCAGAGCGAAACCC   1183
pETDuet-1_VP2_Int .....................................      0
pETDuet1_VP2/6_R  .....................................      0
```

```
pETDuet1_VP2/6    AGTTTCTGACCGGCATTAACAGCCAGGCGGCGAACGATTG   2600
pETDuet1_VP2/6_F  AGTTTCTGACCGGCATTAACAGCCAGGCGGCGAACGATTG   1223
pETDuet-1_VP2_Int ....................................        0
pETDuet1_VP2/6_R  ....................................        0


pETDuet1_VP2/6    CTTTAAAACCCTGATTGCGGCCATGCTGTCTCAGCGTACC   2640
pETDuet1_VP2/6_F  CTTTAAAACCCTGATTGCGGCCATGCTGTCTCAGCGTACC   1263
pETDuet-1_VP2_Int ..................GGCCATGCTGTCTCAGCGTACC     22
pETDuet1_VP2/6_R  ....................................        0


pETDuet1_VP2/6    ATGAGCCTGGATTTTGTGACCACCAACTATATGAGCCTGA   2680
pETDuet1_VP2/6_F  ATGAGCCTGGA.........................      1274
pETDuet-1_VP2_Int ATGAGCCTGGATTTTGTGACCACCAACTATATGAGCCTGA     62
pETDuet1_VP2/6_R  ....................................        0


pETDuet1_VP2/6    TTAGCGGCATGTGGCTGCTGACCGTGGTGCCGAACGACAT   2720
pETDuet1_VP2/6_F  ....................................      1274
pETDuet-1_VP2_Int TTAGCGGCATGTGGCTGCTGACCGTGGTGCCGAACGACAT    102
pETDuet1_VP2/6_R  ....................................        0


pETDuet1_VP2/6    GTTTATTCGTGAAAGCCTGGTGGCGTGCCAGCTGGCCATT   2760
pETDuet1_VP2/6_F  ....................................      1274
pETDuet-1_VP2_Int GTTTATTCGTGAAAGCCTGGTGGCGTGCCAGCTGGCCATT    142
pETDuet1_VP2/6_R  ....................................        0


pETDuet1_VP2/6    GTGAACACCATTATCTATCCGGCGTTTGGCATGCAGCGTA   2800
pETDuet1_VP2/6_F  ....................................      1274
pETDuet-1_VP2_Int GTGAACACCATTATCTATCCGGCGTTTGGCATGCAGCGTA    182
pETDuet1_VP2/6_R  ....................................        0


pETDuet1_VP2/6    TGCATTATCGTAACGGCGATCCGCAGACCCCGTTTCAGAT   2840
pETDuet1_VP2/6_F  ....................................      1274
pETDuet-1_VP2_Int TGCATTATCGTAACGGCGATCCGCAGACCCCGTTTCAGAT    222
pETDuet1_VP2/6_R  ....................................        0


pETDuet1_VP2/6    TGCGGAACAGCAGATCCAGAACTTTCAGGTGGCGAACTGG   2880
pETDuet1_VP2/6_F  ....................................      1274
pETDuet-1_VP2_Int TGCGGAACAGCAGATCCAGAACTTTCAGGTGGCGAACTGG    262
pETDuet1_VP2/6_R  ....................................        0


pETDuet1_VP2/6    CTGCATTTTGTGAACAACAACCAGTTTCGTCAGGCGGTGA   2920
pETDuet1_VP2/6_F  ....................................      1274
pETDuet-1_VP2_Int CTGCATTTTGTGAACAACAACCAGTTTCGTCAGGCGGTGA    302
pETDuet1_VP2/6_R  ....................................        0


pETDuet1_VP2/6    TTGATGGCGTGCTGAACCAGGTGCTGAACGATAACATTCG   2960
pETDuet1_VP2/6_F  ....................................      1274
pETDuet-1_VP2_Int TTGATGGCGTGCTGAACCAGGTGCTGAACGATAACATTCG    342
pETDuet1_VP2/6_R  ....................................        0
```

```
pETDuet1_VP2/6    TAACGGCCATGTGATTAACCAACTGATGGAAGCCCTGATG   3000
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int TAACGGCCATGTGATTAACCAACTGATGGAAGCCCTGATG    382
pETDuet1_VP2/6_R  ....................................      0


pETDuet1_VP2/6    CAGCTGTCTCGTCAGCAGTTTCCGACCATGCCGATCGATT   3040
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int CAGCTGTCTCGTCAGCAGTTTCCGACCATGCCGATCGATT    422
pETDuet1_VP2/6_R  ....................................      0


pETDuet1_VP2/6    ATAAACGTAGCATTCAGCGTGGCATTCTGCTGCTGTCTAA   3080
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int ATAAACGTAGCATTCAGCGTGGCATTCTGCTGCTGTCTAA    462
pETDuet1_VP2/6_R .....................TTCTGCTGCTGTCTAA     16


pETDuet1_VP2/6    CCGTCTGGGCCAGCTGGTTGATCTGACCCGTCTGCTGGCC   3120
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int CCGTCTGGGCCAGCTGGTTGATCTGACCCGTCTGCTGGCC    502
pETDuet1_VP2/6_R  CCGTCTGGGCCAGCTGGTTGATCTGACCCGTCTGCTGGCC     56


pETDuet1_VP2/6    TATAACTATGAAACGCTGATGGCATGCATCACGATGAACA   3160
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int TATAACTATGAAACGCTGATGGCATGCATCACGATGAACA    542
pETDuet1_VP2/6_R  TATAACTATGAAACGCTGATGGCATGCATCACGATGAACA     96


pETDuet1_VP2/6    TGCAGCATGTGCAGACCCTGACCACCGAAAAACTGCAGCT   3200
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int TGCAGCATGTGCAGACCCTGACCACCGAAAAACTGCAGCT    582
pETDuet1_VP2/6_R  TGCAGCATGTGCAGACCCTGACCACCGAAAAACTGCAGCT    136


pETDuet1_VP2/6    GACCAGCGTGACCAGCCTGTGCATGCTGATTGGCAACGCG   3240
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int GACCAGCGTGACCAGCCTGTGCATGCTGATTGGCAACGCG    622
pETDuet1_VP2/6_R  GACCAGCGTGACCAGCCTGTGCATGCTGATTGGCAACGCG    176


pETDuet1_VP2/6    ACCGTGATTCCGAGCCCGCAGACCCTGTTCCATTATTATA   3280
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int ACCGTGATTCCGAGCCCGCAGACCCTGTTCCATTATTATA    662
pETDuet1_VP2/6_R  ACCGTGATTCCGAGCCCGCAGACCCTGTTCCATTATTATA    216


pETDuet1_VP2/6    ACGTGAACGTGAACTTTCACAGCAACTATAACGAACGTAT   3320
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int ACGTGAACGTGAACTTTCACAGCAACTATAACGAACGTAT    702
pETDuet1_VP2/6_R  ACGTGAACGTGAACTTTCACAGCAACTATAACGAACGTAT    256


pETDuet1_VP2/6    TAACGATGCGGTGGCGATTATTACCGCGGCGAACCGTCTG   3360
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int TAACGATGCGGTGGCGATTATTACCGCGGCGAACCGTCTG    742
pETDuet1_VP2/6_R  TAACGATGCGGTGGCGATTATTACCGCGGCGAACCGTCTG    296
```

```
pETDuet1_VP2/6    AATCTGTATCAGAAAAAAATGAAAGCGATCGTGGAAGATT   3400
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int AATCTGTATCAGAAAAAAATGAAAGCGATCGTGGAAGATT    782
pETDuet1_VP2/6_R  AATCTGTATCAGAAAAAAATGAAAGCGATCGTGGAAGATT    336


pETDuet1_VP2/6    TTCTGAAACGCCTGTACATTTTTGATGTGAGCCGTGTGCC   3440
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int TTCTGAAACGCCTGTACATTTTTGATGTGAGCCGTGTGCC    822
pETDuet1_VP2/6_R  TTCTGAAACGCCTGTACATTTTTGATGTGAGCCGTGTGCC    376


pETDuet1_VP2/6    GGATGATCAGATGTATCGTCTGCGTGATCGTCTGCGTCTG   3480
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int GGATGATCAGATGTATCGTCTGCGTGATCGTCTGCGTCTG    862
pETDuet1_VP2/6_R  GGATGATCAGATGTATCGTCTGCGTGATCGTCTGCGTCTG    416


pETDuet1_VP2/6    CTGCCGGTGGAAATTCGTCGTCTGGATATCTTTAACCTGA   3520
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int CTGCCGGTGGAAATTCGTCGTCTGGATATCTTTAACCTGA    902
pETDuet1_VP2/6_R  CTGCCGGTGGAAATTCGTCGTCTGGATATCTTTAACCTGA    456


pETDuet1_VP2/6    TCCTGATGAATATGGATCAGATCGAACGTGCGAGCGATAA   3560
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int TCCTGATGAATATGGATCAGATCGAACGTGCGAGCGATAA    942
pETDuet1_VP2/6_R  TCCTGATGAATATGGATCAGATCGAACGTGCGAGCGATAA    496


pETDuet1_VP2/6    AATTGCGCAGGGCGTGATTATTGCGTATCGTGATATGCAT   3600
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int AATTGCGCAGGGCGTGATTATTGCGTATCGTGATATGCAT    982
pETDuet1_VP2/6_R  AATTGCGCAGGGCGTGATTATTGCGTATCGTGATATGCAT    536


pETDuet1_VP2/6    CTGGAACGTGATGAAATGTATGGCTACGTGAACATTGCGC   3640
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int CTGGAACGTGATGAAATGTATGGCTACGTGAACATTGCGC   1022
pETDuet1_VP2/6_R  CTGGAACGTGATGAAATGTATGGCTACGTGAACATTGCGC    576


pETDuet1_VP2/6    GTAACCTGGAAGGCTTTCAGCAGATTAACCTGGAAGAACT   3680
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int GTAACCTGGAAGGCTTTCAGCAGATTAACCTGGAAGAACT   1062
pETDuet1_VP2/6_R  GTAACCTGGAAGGCTTTCAGCAGATTAACCTGGAAGAACT    616


pETDuet1_VP2/6    GATGCGTAGCGGCGATTATGCGCAGATTACCAACATGCTG   3720
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int GATGCGTAGCGGCGATTATGCGCAGATTACCAACATGCTG   1102
pETDuet1_VP2/6_R  GATGCGTAGCGGCGATTATGCGCAGATTACCAACATGCTG    656


pETDuet1_VP2/6    CTGAACAACCAGCCGGTGGCGCTGGTTGGTGCGCTGCCGT   3760
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int CTGAACAACCAGCCGGTGGCGCTGGTTGGTGCGCTGCCGT   1142
pETDuet1_VP2/6_R  CTGAACAACCAGCCGGTGGCGCTGGTTGGTGCGCTGCCGT    696
```

```
pETDuet1_VP2/6    TTATTACCGATAGCAGCGTGATTAGCCTGATTGCGAAACT   3800
pETDuet1_VP2/6_F  ........................................   1274
pETDuet-1_VP2_Int TTATTACCGATAGCAGCGTGATTAGCCTGATTGCGAAACT   1182
pETDuet1_VP2/6_R  TTATTACCGATAGCAGCGTGATTAGCCTGATTGCGAAACT    736


pETDuet1_VP2/6    GGATGCGACCGTGTTTGCGCAGATTGTGAAACTGCGTAAA   3840
pETDuet1_VP2/6_F  ........................................   1274
pETDuet-1_VP2_Int GGATGCGACCGTGTTTGCGCAGATTGTGAAACTGCGTAAA   1222
pETDuet1_VP2/6_R  GGATGCGACCGTGTTTGCGCAGATTGTGAAACTGCGTAAA    776


pETDuet1_VP2/6    GTGGATACCCTGAAACCGATCCTGTATAAAATCAACAGCG   3880
pETDuet1_VP2/6_F  ........................................   1274
pETDuet-1_VP2_Int GTG.....................................   1225
pETDuet1_VP2/6_R  GTGGATACCCTGAAACCGATCCTGTATAAAATCAACAGCG    816


pETDuet1_VP2/6    ATAGCAACGATTTTTATCTGGTGGCGAACTATGATTGGGT   3920
pETDuet1_VP2/6_F  ........................................   1274
pETDuet-1_VP2_Int ........................................   1225
pETDuet1_VP2/6_R  ATAGCAACGATTTTTATCTGGTGGCGAACTATGATTGGGT    856


pETDuet1_VP2/6    GCCGACCAGCACCACCAAAGTGTATAAACAGGTGCCGCAG   3960
pETDuet1_VP2/6_F  ........................................   1274
pETDuet-1_VP2_Int ........................................   1225
pETDuet1_VP2/6_R  GCCGACCAGCACCACCAAAGTGTATAAACAGGTGCCGCAG    896


pETDuet1_VP2/6    CAGTTTGATTTTCGTAACTCTATGCACATGCTGACCAGCA   4000
pETDuet1_VP2/6_F  ........................................   1274
pETDuet-1_VP2_Int ........................................   1225
pETDuet1_VP2/6_R  CAGTTTGATTTTCGTAACTCTATGCACATGCTGACCAGCA    936


pETDuet1_VP2/6    ACCTGACCTTTACCGTGTATAGCGATCTGCTGGCCTTTGT   4040
pETDuet1_VP2/6_F  ........................................   1274
pETDuet-1_VP2_Int ........................................   1225
pETDuet1_VP2/6_R  ACCTGACCTTTACCGTGTATAGCGATCTGCTGGCCTTTGT    976


pETDuet1_VP2/6    GAGCGCGGATACCGTGGAACCGATTAACGCGGTGGCGTTT   4080
pETDuet1_VP2/6_F  ........................................   1274
pETDuet-1_VP2_Int ........................................   1225
pETDuet1_VP2/6_R  GAGCGCGGATACCGTGGAACCGATTAACGCGGTGGCGTTT   1016


pETDuet1_VP2/6    GATAACATGCGCATCATGAACGAACTGTAATAACCTAGGC   4120
pETDuet1_VP2/6_F  ........................................   1274
pETDuet-1_VP2_Int ........................................   1225
pETDuet1_VP2/6_R  GATAACATGCGCATCATGAACGAACTGTAATAACCTAGG.   1054


pETDuet1_VP2/6    TGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGG   4160
pETDuet1_VP2/6_F  ........................................   1274
pETDuet-1_VP2_Int ........................................   1225
pETDuet1_VP2/6_R  ........................................   1054
```

```
pETDuet1_VP2/6    GCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAG   4200
pETDuet1_VP2/6_F  ....................................   1274
pETDuet-1_VP2_Int ....................................   1225
pETDuet1_VP2/6_R  ....................................   1054
```

**Figure 2.20: Nucleotide sequence alignments of the recombinant pETDuet-1_VP6 vector containing the ORF encoding VP2 with the *in silico* clone of pETDuet-1_VP2/6.** The in *silico clone* of pETDuet-1_VP2/6 is indicated as pETDuet-1_VP2/6, the forward sequence is indicated as pETDuet1_VP2/6_F nucleotide sequence, the reverse sequence is indicated as pETDuet1_VP2/6_R nucleotide sequence and the internal sequence is indicated as pETDuet-1_VP2/6_Int nucleotide sequence. The restriction enzymes used for the amplification of genome segment 6 are indicated in blue. The start codon of genome segment 6 is indicated in green and the stop codon is indicated in red. The T7 promoter-1 is indicated in purple, the T7 transcription start-1 is indicated in dark blue and the *lac* operator in yellow.

Commercial available primers (see Figure 2.1) for pETDuet_1 were used for sequencing, as described in section 2.2.3.12. The sequence alignment is indicated for pETDuet-1_VP2 in Figure 2.19 and for pETDuet-1_VP2/6 in Figure 2.20.

It is clear from Figure 2.19 and Figure 2.20 that the ORF encoding VP2 was cloned into both pETDuet-1 and pETDuet-1_VP6, respectively and that no mutations had occurred during cloning procedures. A part of the ORF encoding VP2 could not be sequenced (1459bp – 1997bp for pETDuet-1 and 2651bp – 3064bp for pETDuet-1_VP2/6) with the commercially available primers. Therefore, an internal primer had to be designed to sequence this part, which is also indicated in Figure 2.19 and Figure 2.20. The results of the internal primer indicated that the entire open reading frame encoding VP2 was cloned into both pETDuet_1_VP2 and pETDuet-1_VP6 respectively. Based on these results the next step of the project namely the individual and co-expression of VP2/6 in bacteria (discussed in section 2.3.3) was started.

### 2.3.3 Expression of the rotavirus bacterial codon optimised proteins VP2 and VP6

In order to evaluate the individual and co-expression of the bacterial codon optimised proteins, all constructs (pETDuet-1_VP6 colony 13, pETDuet-1_VP2 colony 4 and pETDuet-1_VP2/6 colony 6) were transformed into the bacterial Origami cells. The transformation of the Origami cells was done as described in section 2.2.3.10. Expression and analysis of protein expression were done as described in section 2.2.4, 2.2.5 and 2.2.6.

***Expression of rotavirus VP6 from pETDuet-1_VP6:***

Expression of the recombinant protein, VP6 (Figure 2.21) was done to determine firstly, if it is expressed in bacteria and if so secondly, if the VP6 was in a soluble form. A band at 48kDa will indicate that expression of VP6 did occur. If a band at 48kDa was visible in the soluble fraction sample it would be an indication that the VP6 may be soluble in bacteria.

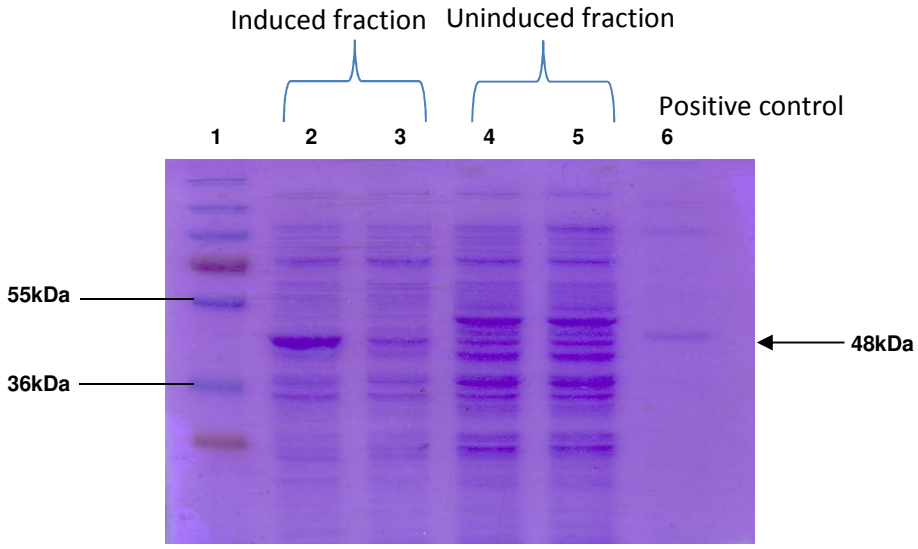


**Figure 2.21: SDS-PAGE analysis of VP6 expression in bacteria.** . Lane: 1) Page Ruler protein marker (Fermentas). A 10 µl volume of each of the prepared 40 µl protein samples was loaded as follows; lanes; 2) induced total fraction; 3) induced soluble fraction; 4) uninduced total fraction; 5) uninduced soluble fraction; 6) positive control, insect cell-produced VP2/6 DLP lysate.

The expected result was obtained as visualized by SDS-PAGE (Figure 2.21). After induction the band at 48kDa indicated that the recombinant bacterial codon optimised VP6 was expressed in bacteria (Figure 2.21, lane 2). However, the band was only visible in the total

fraction and not the soluble fraction (Figure 2.21, lane 3) indicating that the expressed protein was not soluble in bacteria. Lane 6 is the positive control, double-layered particles VP2/6 insect cell lysate obtained from Dr. Jere.

To improve the solubility of the bacterial codon optimised VP6 in bacteria two steps were taken. Firstly, the recombinant construct was transformed into Origami cells containing plasmids expressing the pGro 7 chaperone and secondly, three lyses buffers were used. The solubility of the proteins was tested using three different lyses buffers. Cells were harvested by centrifugation at 5 000 x g for 5 minutes. The lyses buffers that were used were (i) Phosphate buffer solution (PBS) containing 0.5% nonyl phenoxyl-polyethoxyl-ethanol (NP40), (ii) 10 mM Tris containing 0.1 mM EDTA, 1% dissolved organic carbon (DOC), 0.1% SDS and (iii) 10 mM Tris containing 0.1 mM EDTA, 1% SDS. To prevent the degradation of the proteins, all three lyses buffers contained a Complete Mini EDTA-free protease inhibitor cocktail solution (Roche) that contains serine and protease inhibitors. The cells were re-suspended in 100 µl of the lyses buffers. The soluble and total protein fractions were separated through centrifugation at 13 000 x g for 15 minutes. Samples re-suspended in buffers (i) and (ii) were centrifuged immediately. The sample re-suspended in buffer (iii) was incubated at room temperature for 30 minutes prior to centrifugation. The insoluble fractions were discarded and the presence of the proteins in the soluble fraction was evaluated using SDS-PAGE (section 2.2.6) (Figure 2.22).

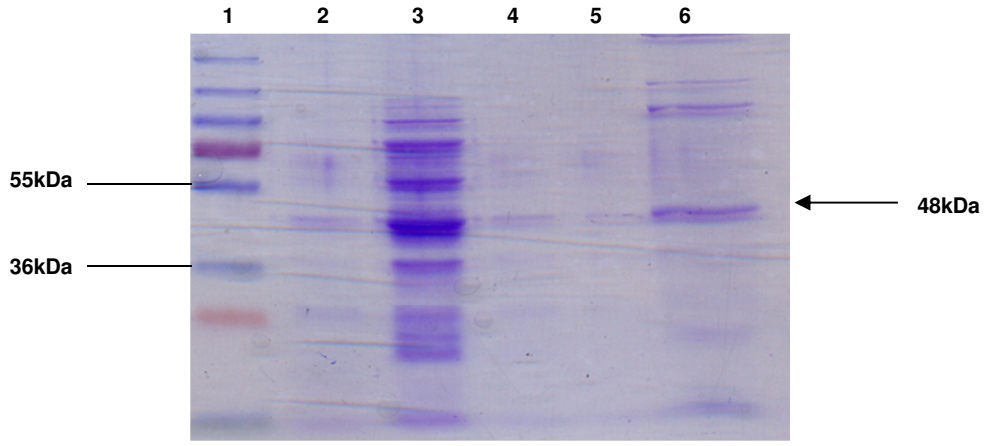


**Figure 2.22: SDS-PAGE analysis of VP6 solubility in bacteria using three different lyses buffers.** Lane: 1) Page Ruler protein marker (Fermentas). A 10 µl volume of each of the prepared soluble fraction of the protein samples were loaded as follow; lanes: 2) Lyses buffer 1: Phosphate buffer solution (PBS) containing 0.5% nonyl phenoxyl-polyethoxyl-ethanol (NP40); 3) Lyses buffer 2: 10 mM Tris containing 0.1 mM EDTA, 1% dissolved organic carbon (DOC), 0.1% SDS and; 4) Lyses buffer 3: 10 mM Tris containing 0.1 mM EDTA, 1% SDS; 5) positive control, insect cell-produced DLP VP2/6 lysate.

The same quantity of proteins was loaded for all three buffers. A band at 48kDa would indicate that bacterial codon optimised VP6 may possibly be soluble in bacteria. A band of the expected size (48kDa) could be seen in Figure 2.22, lane 2 - 4, but the best solubility seemed to be in lane 2 as more soluble protein (including possibly VP6) was visible. Lane 5 contains the positive control, insect cell-produced DLP VP2/6 lysate.

***Co-expression of rotavirus VP2/6 proteins in bacteria***

Co-expression of the bacterial codon optimised recombinant proteins, VP2 and VP6 were done to determine firstly if they do co-express in bacteria and if so, secondly, if the two proteins (VP2 and VP6) were soluble in bacteria. We expected to see two bands namely at 48kDa (VP6) and at 102kDa (VP2) in order to have a positive result for the co-expression of the two proteins. Bands at 48kDa and 102kDa in the soluble protein sample will indicate that these two proteins (VP2 and VP6) are soluble in bacteria.

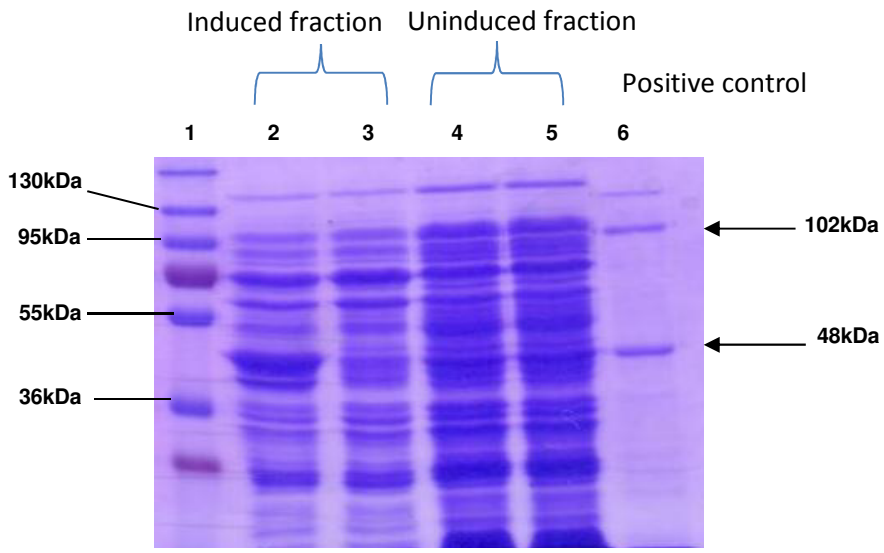


**Figure 2.23: SDS-PAGE analysis of co-expression of bacterial codon optimised VP2/6 in bacteria**. Lane: 1) Page Ruler protein marker (Fermentas). A 10 µl volume of each of the prepared 40 µl protein samples were loaded as follows; lanes: 2) Induced total fraction; 3) Induced soluble fraction; 4) Uninduced total fraction; 5) Uninduced soluble fraction; 6) positive control, insect cell-produced DLP VP2/6 lysate.

After induction a band at 48kDa indicated that the bacterial codon optimised VP6 protein was possibly expressed (Figure 2.23). The absence of an additional band at 102kDa after induction indicated that VP2 was probably not expressed. However, to confirm with certainty that VP2 was not expressed other factors have to be looked at that could possibly confirm the results, these factors include changing gel precentages, overloading of proteins and gel separation conditions. The next step was to try to co-express the proteins in the presence of chaperones. Only the induced total and induced soluble fraction were loaded on the SDS-PAGE gel.

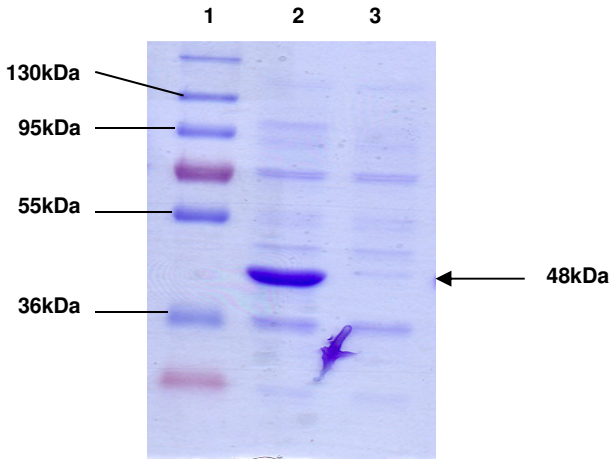


**Figure 2.24: SDS-PAGE analysis of co-expression of bacterial codon optimised VP2/6 in Origami cells in the presence of pGro7 chaperones.** Lane: 1) Page Ruler protein marker (Fermentas). A 10 µl volume of each of the prepared 40 µl protein samples were loaded as follows; lanes: 2) induced total fraction; 3) induced soluble fraction.

When VP2 and VP6 was expressed in Origami cells, in the presence of the pGro7 chaperones (Figure 2.24), VP6 was expressed at high levels but no band of the expected size (102kDa, VP2) was present. Therefore, VP2 could also not be expressed in the presence of chaperones. The next step was to investigate if VP2 could be expressed when it was expressed on its own.

***Expression of VP2 in bacteria***

Individual expression of the bacterial codon optimised recombinant protein VP2 in the pETDuet-1 plasmid, was done to see whether or not VP2 can be expressed. The expected result would be to see a band at 102kDa. No control was included with the samples.

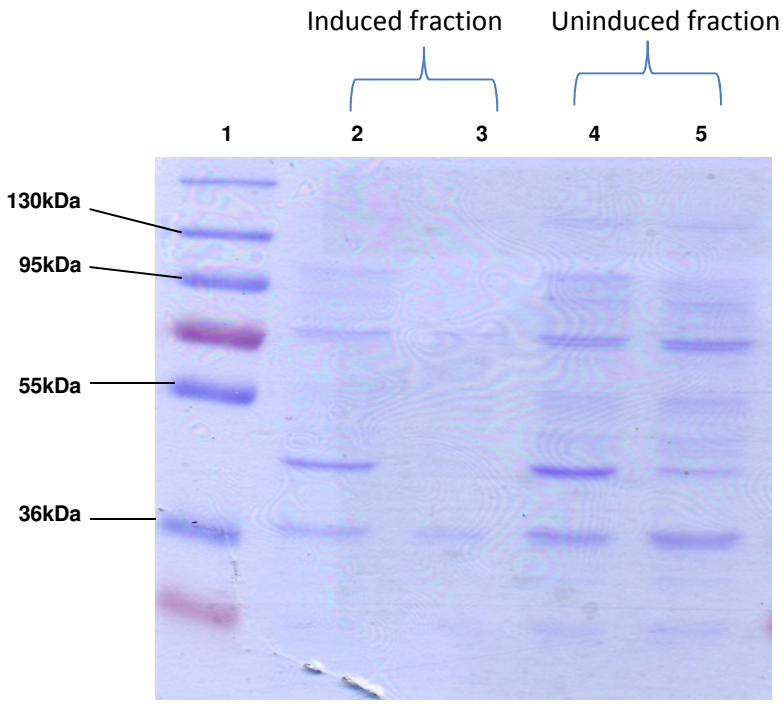


**Figure 2.25: SDS-PAGE analysis of individual expression of bacterial codon optimised VP2 in the expression plasmid pETDuet-1.** Lane: 1) Page Ruler protein marker (Fermentas). A 10 µl volume of each of the prepared 40 µl protein samples were loaded as follows; lanes: 2) induced total fraction; 3) induced soluble fraction; 4) uninduced total fraction; 5) uninduced soluble fraction.

The expected results were also not obtained (Figure 2.25), namely VP2 could not express (lane 2). To determine if the problem was with VP2 itself or with the T7 promoter (since the T7 promoter utilizes the T7 RNA polymerase to catalyse the formation of RNA in a 5 '-3' direction and to drive expression) in the second multiple cloning site of the expression plasmid pETDuet-1, the ORF encoding VP2 was expressed in its original plasmid, pColdTF.

***Expression of VP2 using the cold-shock vector pColdTF***

The expression of VP2 using pColdTF was done as described in the pCold DNA manual. The pColdTF_VP2 construct was transformed into Origami cells. The analysis of the VP2 expression was analysed as described in section 2.2.6. We expected to obtain a band at ±154kDa VP2 fusion protein to verify that VP2 can be expressed in bacteria, VP2 is 102kDa and the trigger factor 52kDa.

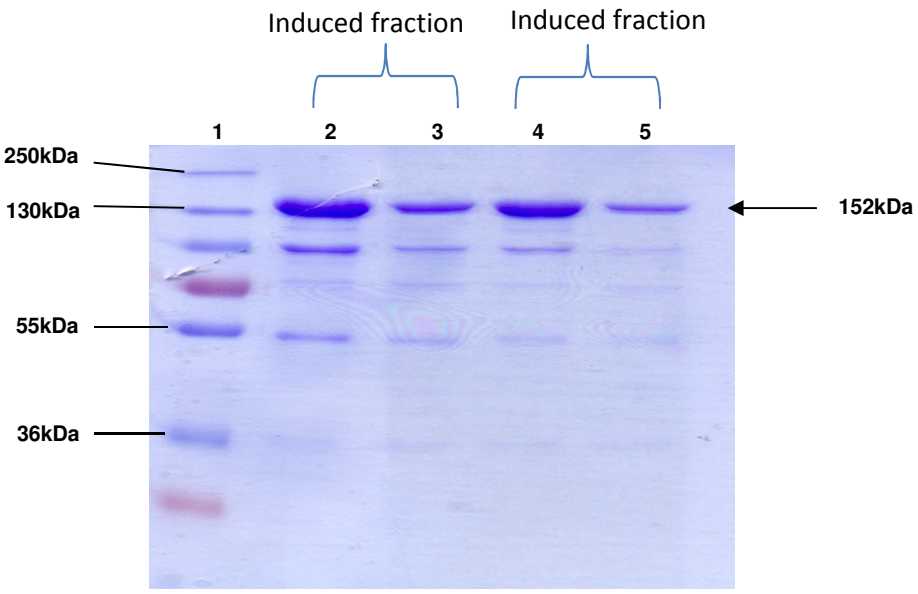


**Figure 2.26: SDS-PAGE analysis of expression of VP2 expressed in the pColdTF vector, in bacteria.** Lane: 1) Page Ruler protein marker (Fermentas). A 10 µl volume of each of the prepared protein samples were loaded as follows; lanes: 2) induced total fraction colony 1; 3) induced soluble fraction colony 1; 4) induced total fraction colony 2; 5) induced soluble fraction colony 2.

No control and only the induced samples were loaded since we only wanted to determine if VP2 did express. A band of approximately 154kDa (VP2 and the trigger factor) was visible (Figure 2.26) indicating that VP2 can be expressed in bacteria. Therefore, the problem was not with VP2 but could somehow be with the pETDuet-1 vector used. However, further studies will have to be conducted to determine for sure why VP2 did not express. Western blot analysis could also be done to confirm the expression of VP2 in bacteria.

### 2.4 Summary

This part of the study had two main objectives. Firstly, to clone the ORFs of the bacterial codon optimised genome segments 2 and 6 of the human rotavirus GR10942 G9P[6] strain into the expression plasmid pETDuet-1. Secondly, to investigate the possible co-expression of the ORF encoding the bacterial codon optimised rotavirus structural proteins VP2 and VP6 in bacteria.

Three constructs were produced in the expression vector namely pETDuet-1_VP2, pETDuet-1_VP6 and pETDuet-1_VP2/6. The expression of the bacterial codon optimised ORF encoding VP6 resulted in insoluble proteins (Figure 2.21). When VP6 was expressed in Origami cells containing plasmids expressing pGro7 chaperones and analysed in three different lysis buffers ((i) Phosphate buffer solution (PBS) containing 0.5% nonyl phenoxyl-polyethoxyl-ethanol (NP40), (ii) 10 mM Tris containing 0.1 mM EDTA, 1% dissolved organic carbon (DOC), 0.1% SDS and (iii) 10 mM Tris containing 0.1 mM EDTA, 1% SDS) soluble VP6 was obtained (Figure 2.22). However the co-expression of the bacterial codon optimised ORF encoding VP2 and VP6 could not be achieved since only VP6 could be expressed (Figure 2.23) even in the presence of Origami cells expressing pGro7 chaperones (Figure 2.24). VP2 could also not be expressed on its own in the pETDuet-1 vector (Figure 2.25). Since the ORF encoding VP2 could not be expressed individually or co-expressed with VP6, even in the presence of pGro 7 chaperones, VP2 was expressed in bacteria in the expression plasmid it was originallly received in namely pColdTF. This was done to determine if the problem was with the ORF encoding VP2 or possibly with the expression vector, pETDuet-1, that was used. Expression of VP2 in bacteria was possible in the pColdTF vector, however this was achieved in the presence of a trigger factor and with another expression protocol (Figure 2.26). Western blot analysis can also be done to determine if VP2 can express in bacteria. However, if the expected result is not obtained with Western blot analysis further studies will have to be conducted to determine exactly what is causing VP2 not to express in bacteria.

In conclusion; the ORF encoding VP6 could be expressed in bacteria and it was shown that soluble VP6 can be produced in bacteria. The co-expression of the ORF encoding VP2 and VP6 only showed the expression of VP6 but not VP2, even in the presence of chaperones. The expresson of VP2 expression vector pColdTF was possible, in bacteria. In terms of our long term objectives of developing double-layered particles in bacteria, this was an important first step. The next step will be to conduct studies to determine how VP2 can express in bacteria in the expression plasmid pETDuet-1. These studies include (i) western blot analysis to determine if VP2 is expressed at very low levels in bacteria, (ii) expressing VP2 when it is cloned into the first multiple clonning site of pETDuet-1, to determine if the problem is not with the second multiple cloning sites T7 promoter of pETDuet-1, and (iii) to express VP2 in the presence of a trigger factor.

# Chapter 3: Cloning of rotavirus structural proteins VP2 and VP6 open reading frames of a South African G9P[6] strain, for expression in several yeast strains

## 3.1 Introduction

Due to the difficulty to obtain expression of VP2 and double-layered particles formation in bacteria (Chapter 2), it was decided to pursue expression of VP2 and VP6 in yeast as well. Therefore, this part of the study involves the construction and expression of the rotavirus structural proteins, VP2 and VP6, in several yeast strains. Yeast expression systems are attractive because of their ability to grow quickly to high cell densities on simple defined media, their ability to be adapted to fermentation and the fact that they are easier and less expensive to work with compared to mammalian and insect cells. In the literature, Rodriguez-Limas and co-workers (Rodrigues-Limas *et al.*, 2011) showed that the expression and production of rotavirus-like particles, specifically double-layered particles (DLPs), was possible in *Saccharomyces cerevisiae*. The three genome segments encoding VP2, VP6 and VP7 were cloned into the pSP-GM2 plasmid (variant of pSP-G2 plasmid developed by Partow *et al.,* 2010) to obtain the pWR267 plasmid used for expression experiments of all three proteins. This study showed the formation of triple-layered particles (tLRP) in yeast although the tLRP yield was very low in comparison to other expression platforms (Rodrigues-Limas *et al.*, 2011). In a second study that Rodriguez-Limas and co-workers conducted they showed that *Saccharomyces cerevisiae* yeast extract containing rotavirus proteins can efficiently induce protection against infection in adult mice models depending on the route of immunisation serum and antibodies used (Rodriguez-Limas *et al.,* 2014).

The yeast expression part of the study had two main objectives. Firstly, to clone the open reading frames (ORF) of the wild type rotavirus genome segment 2 and genome segment 6 into the yeast expression vector, pKM173 which allows for the simultaneous expression of both genome segments. Secondly, to evaluate the co-expression of the rotavirus structural proteins, VP2 and VP6, in eight different yeast strains namely *Kluyveromyces marxianus, Kluyveromyces lactis, Candida deformans, Saccharomyces cerevisiae, Yarrowia lipolytica, Arxula adeninivorans, Hansenula polymorpha* and *Debaryomyces hansenii*. The yeast expression was carried out in collaboration with Prof. Jacobus Albertyn, University of the Free State (UFS) who developed, together with Dr. Michel Labuschagne and Prof. Martie

Smit, a wide-range yeast expression system that allows for multiple protein expression in any yeast strain that can be transformed (Albertyn *et al*., 2011). The preparation of the construct was done at the North-West University (NWU) whereas the expression experiments were done at the University of the Free State (UFS).

## 3.2 Materials and methods

As mentioned previously in Chapter 2 (section 2.2) the names, suppliers, and catalogue numbers of the reagents used in this study, are included in Appendix A.

### 3.2.1 Virus, bacterial and yeast strains used in this project

The ORFs of the wild type genome segment 2 and 6 consensus sequence derived from the viral genome (Potgieter *et al*., 2009) of the human rotavirus strain GR10924 (G9P[6]) was used for this part of the study. The open reading frame of genome segment 2 and genome segment 6 was purchased from Genscript and provided in the pUC57 vector by the company.

JM109 cells are a strain of *Escherichia coli (E. coli),* which was used for molecular cloning (discussed in Chapter 2, section 2.2.1). Eight yeast strains, *Kluyveromyces marxianus, Kluyveromyces lactis, Yarrowia lipolytica, Debaryomyces hansenii, Candida deformans, Hansenula polymorpha, Arxula adeninivorans* and *Saccharomyces cerevisiae* were obtained from the UNESCO-MIRCEN yeast culture collection at the University of the Free State for expression experiments. *Saccharomyces cerevisiae* was included as reference since VP2 and VP6 were successfully expressed in this yeast previously (Rodrigues-Limas *et al.,* 2011)**.**

### 3.2.2 Cloning vector

The wide-range yeast expression system was developed at the University of the Free State and is not limited to a specific yeast strain but allows for possible heterologous expression in any yeast that can be transformed (Albertyn *et al*., 2011). A vector pair pKM173 and pKM177 (Figure 3.1) was designed previously to allow upon fusion a dual insert vector. The wide range yeast expression system contains the 18S rRNA target sequences of *Kluyveromyces marxianus* to allow genomic integration into yeast cells. The selective marker for yeast expression is the *hph* gene conferring hygromycin B resistance. The vectors also contain the *Yarrowia lipolytica* TEF promoter and the *Kluyveromyces marxianus*

inulinase terminator. To facilitate sub-cloning in *Escherichia coli* the vectors contain the kanamycin resistance gene. The vector pair, pKM173 and pKM177, also contains the I-SceI restriction enzyme site (pKM173 one I-SceI restriction site and pKM177 two I-SceI restriction sites). The I-SceI site allows for the cassette containing the genome segment of interest (genome segment 6 cloned into pKM177 vector) to be cloned into the pKM173 vector linearized with I-SceI resulting in the dual expression vector.

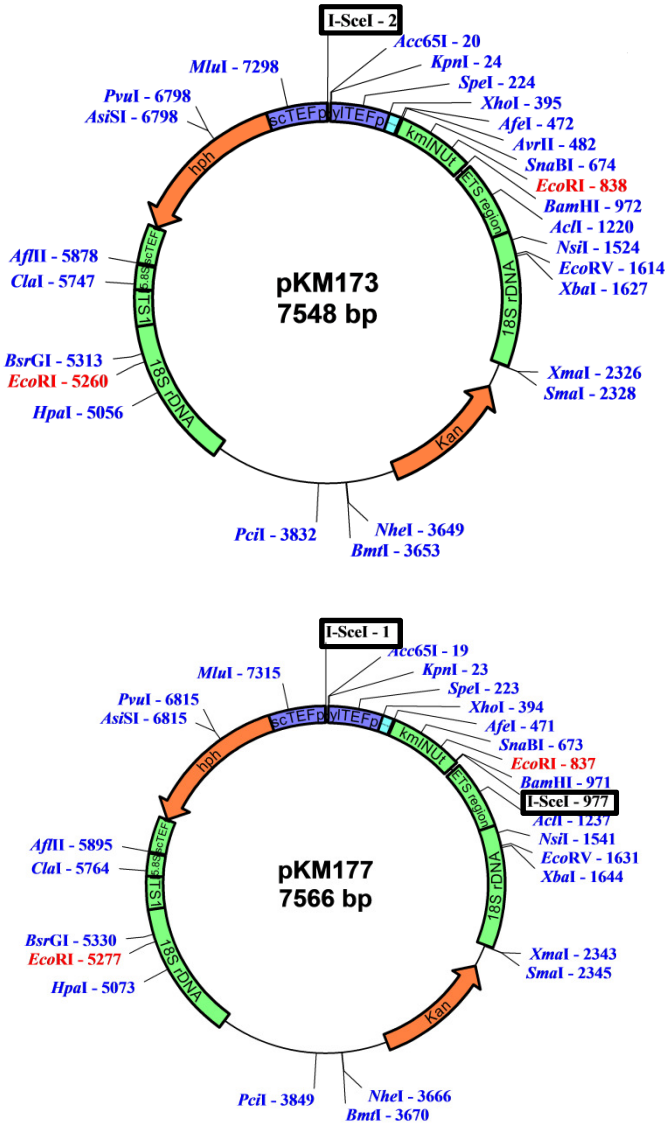


**Figure 3.1: The plasmid map of pKM173 and pKM177**. The plasmid maps illustrate the various properties of pKM173 and pKM177. The expression system contains the 18S rRNA target sequences of *Kluyveromyces marxianus* to allow genomic integration into yeast cells. The selective marker is the *hph* gene conferring hygromycin B resistance. The vectors also contain the *Yarrowia lipolytica* TEF promoter and the *Kluyveromyces marxianus* inulinase terminator.

### 3.2.3 DNA recombinant techniques

Recombinant DNA techniques used in this study, namely polymerase chain reaction (PCR), restriction endonuclease digestion, ligation and transformation were performed according to Sambrook and Russell (2001). Commercial kits were used for plasmid isolation, gel extraction and PCR-clean-up.

The basic principles and description of the following DNA recombinant techniques have been discussed in Chapter 2 (section 2.2.3) and will therefore not be discussed again: restriction endonuclease digestion, purification of PCR amplicons, agarose gel electrophoresis, analysis of DNA concentration and purity, gel purification of desired DNA fragments or product, ligation reactions, preparation of chemical competent *Escherichia coli* cells, transformation of chemical competent *Escherichia coli* cells and screening of colonies of transformed bacteria. However, some DNA techniques will be discussed below, namely plasmid isolation (will be discussed since a different plasmid isolation kit was used), PCR amplification of the coding sequences, dephosphorylation, long term storage of desired colonies, DNA sequencing determination, expression of proteins and analysis of protein expression.

#### 3.2.3.1 Plasmid isolation

Propagation of bacteria containing the plasmid of interest was inocculated into 50 ml LB medium (10 mg/ml kanamycin) and grown for 16 hours at 37°C, while shaking at 180rpm. The next day these cultures, containing the appropriate plasmid, were centrifuged at 5 000 x g for 10 minutes and the supernatants were discarded.

A commercial plasmid purification kit, QIAGEN® Plasmid Mini, Midi, Maxi, Mega and Giga, system, was used to do the plasmid purifications. The pellets of the cultures were used to isolate plasmids using the QIAGEN® Plasmid Mini, Midi, Maxi, Mega and Giga kit. The isolation was done according to the instructions of the manufacturer. This kit is designed to allow purification of ultrapure supercoiled plasmid DNA with high yields and uses an anion-exchange–based QIAGEN-tip to isolate plasmid DNA. The kit also incorporates LysesBlue which is a colour indicator that provides visual identification of optimum buffer mixing. This prevents common handling errors that lead to inefficient cell lysis and incomplete precipitation of SDS, genomic DNA, and cell debris.

The procedure consists of seven steps namely i) preparation and lysis of cell culture, ii) DNA purification by centrifugation, iii) filtration with an anion-exchange QIAGEN tip iv) medium

salt wash, v) precipitation of DNA, vi) wash and vii) elution. The alkaline conditions denature the chromosomal DNA and proteins. The lysis time ensures the optimum release of plasmid DNA without the release of chromosomal DNA and without exposing the plasmid DNA to denaturing conditions for too long. The lysate is then neutralized which causes denatured components to precipitate while only small plasmid DNA renaturates and stay in the solution. The plasmid DNA binds to the QIAGEN Anion-Exchange Resin under appropriate low-salt and pH conditions and RNA, proteins, dyes, and low-molecular-weight impurities are removed by a medium-salt wash. The plasmid DNA is eluted in a high-salt buffer and then concentrated and desalted by isopropanol precipitation. The second wash step removes precipitated salt and replaces isopropanol with the more volatile ethanol, making the DNA easier to re-dissolve. Plasmid DNA was re-dissolved in 100 µl nuclease free water.

#### 3.2.3.2 PCR amplification of the coding sequences

The amplification was performed with the use of a BioRad™ thermocycler. The coding regions were amplified in reaction mixtures containing 0.5 µM of both the forward and reverse primer (2.5 µl) (indicated in Table 3.1) and 0.5 µl template (57 ng), 10 x Takara Ex-Taq buffer (1X), Takara Ex-Taq (1.25 units/50µl), 2.5 mM dNTP and nuclease free water added to a final volume of 50µl. The negative controls contained no template.

**Table 3.1: Oligonucleotide primers used in this study for PCR amplification**

| Primer name | Oligonucleotide sequence (5' → 3') | $T_m$ (°C) | Length (bp) |
|---|---|---|---|
| VP6yeast_F | CAA C**CT CGA G**AT GGA TGT CCT GTA C | 58.6°C | 25 |
| VP6yeast_R | GTC C**AG CGC T**TT ATT TGA CAA GCA TGC T | 61.7°C | 28 |
| VP2yeast_F | CTC A**CT CGA G**AT GGC GTA CAG GAA AC | 60.7°C | 26 |
| VP2yeast_R | GCG T**CC TAG G**CT ACA ATT CGT TCA TGA T | 60.3ºC | 28 |

*All primers were synthesized by Integrated DNA technologies (IDT)

*Restriction enzymes in primers used for amplification are indicated in bold

The samples were denaturated at 98°C for 1.5 minutes followed by 30 cycles of amplification, unless otherwise stated. Each cycle consisted of 15 seconds denaturation at 95°C, 30 seconds of annealing (temperature depends on annealing temperature of primers, indicated in Table 3.2) and 2 minutes of extension at 72°C. The amplification cycle was

followed by one elongation cycle at 72°C for 10 minutes. PCR products were analysed by means of gel electrophoresis using a 1% agarose gel.

**Table 3.2: Annealing temperatures of different PCR primers**

| Genome segment amplified | Primers used | Annealing $T_m$ (°C) |
|---|---|---|
| Genome segment 2 | VP2yeast_F and VP2yeast_R | 55°C |
| Genome segment 6 | VP6yeast_F and VP6yeast_R | 55°C |

### 3.2.3.3 Restriction endonuclease digestions

Restriction enzyme digestion reactions were performed as explained in section 2.2.3.6, Chapter 2. The enzymes and buffers used for the cloning of the ORFs of the genome segments into the cloning vectors and expression vector are indicated in Table 3.3.

**Table 3.3: Restriction enzymes and buffers used in this study**

| Genome segment and the vector, the genome segment was cloned into | Restriction enzyme | Recognition sequence | Buffer used | Conditions for double digestion |
|---|---|---|---|---|
| Genome segment 2 pKM173 | XhoI<br>XmaJI | C↓TCGAG<br>C↓CTAGG | Buffer R | 10 units/µl XmaJI and 5 units/µl XhoI |
| Genome segment 6 pKM177 | XhoI<br>Eco47III | C↓TCGAG<br>AGC↓GCT | Buffer R | 5 units/µl XhoI and 5 units/µl Eco47III |
| Cassette containing genome segment 6 pKM173_VP2 | I-SceI | TAGGG↑ATAA↓<br>CAGGGTAAT | Tango Buffer | 20 units/µl I-SceI |

### 3.2.3.4 Dephosphorylation

Dephosphorylating reactions were performed on the restriction digestion reactions of pKM173_VP2 in volumes of 25 µl. Each reaction mixture consisted of 5 units of Antarctic phosphatase enzyme (Biolabs), 1/10 of 10 x Antarctic phosphatase reaction buffer (Biolabs) and approximately 1000 ng DNA cut with any restriction endonuclease in any buffer. Reaction mixtures were incubated at 37°C for approximately 15 minutes for 5' extensions or blunt-ends and 1 hour for 3' extensions. The digested samples were heat inactivated at

65°C for 5 minutes and analysed using agarose gel electrophoreses (section 2.2.3.4, Chapter 2) and purified (section 2.2.3.7, Chapter 2).

### 3.2.3.5 PCR colony screening

PCR reactions were performed in reaction volumes of 25 µl. These reaction mixtures consisted of the following namely 0.025 units GoTaq (Promega), 25 mM $MgCl_2$, 2.5 mM dNTPs, 1 x GoTaq green buffer, 1 µm of the appropriate forward and reverse primer and nuclease free water to a final volume of 25 µl. Colonies were picked from the overnight plates using sterile tips and each different one streaked out in a different block on the master plate grids. The colonies were also inoculated in their own 25 µl master mix (marked according to colony picked up with the particular toothpick) for a colony screening.

The final reaction mixtures of 25 µl was incubated in a Top-line PCR tube. Thermal cycling conditions were denaturation at 95°C for 2 minutes (the denaturation step breaks open the bacterial cell and releases the DNA for PCR) followed by 30 cycles of amplification, unless otherwise stated. Each cycle consisted of 30 seconds denaturation at 94°C, 30 seconds of annealing (temperature depends on annealing temperature of primers, indicated in Table 3.2) and 1.30 minutes of extension at 72°C. The amplification cycle was followed by one elongation cycle at 72°C for 10 minutes. The PCR products were analysed by means of gel electrophoresis using a 1% agarose gel. The master plate was incubated for 6 hours at 37°C.

### 3.2.3.6 Long term storage of desired colonies

Positive colonies for the constructs, transformed into the different yeast strains, were prepared for long time storage by adding glycerol to a final concentration of 15%. This was done by combining, in a cryotube, 500 µl of the culture and 500 µl of 30% glycerol to obtain a final volume of 1 ml. The stocks were then stored at -80°C.

### 3.2.3.7 DNA sequence determination

DNA sequencing was carried out to confirm that the recombinant plasmid, containing the genes of insert, does not contain any mutations and that cloning was successful. Sanger sequencing was used. Sequencing reactions were prepared using the BigDye terminator v. 3.1 cycle sequencing kit (Applied Biosystems) and analysed using the 3130*xl* Genetic Analyser from Applied Biosystems at the Department of Microbial, Biochemical and Food Biotechnology, University of the Free State. Good quality template DNA should be used for any sequencing reaction, low quantity and quality template DNA can be the major reasons for failed reactions. Plasmid DNA was used for sequencing and, therefore, 500 ng DNA was used for the reaction according to BigDye sequencing manual (published by Apllied Biosystmes). PCR cycle sequencing reaction were performed in a 1/16 reaction size (indicated in the manual).

The amplification was performed with the use of a BioRad$^{TM}$ thermocycler. The following reactions were done for both the forward and reverse reaction: the coding regions were amplified in reaction mixtures containing 1 µl of the forward/reverse primer (3.2 pmol.µl$^{-1}$) (Table 3.4) and 500 ng template, 0.5 µl reaction premix (supplied by the BigDye terminator v. 3.1 cycle sequencing kit), 2 µl dilution buffer and nuclease free water was added to a final volume of 10 µl. The samples were denaturated at 96°C for 1 minute followed by twenty five cycles of amplification. Each cycle consisted of 10 seconds denaturation at 96°C, 5 seconds of annealing (temperature depends on annealing temperature of primer, indicated in Table 3.2) and 4 minutes of extension at 60°C. The amplification cycle was paused at 4°C. The control reaction was set up according to the manual.

**Table 3.4: Oligonucleotide primers used for sequencing in this study**

| Primer name | Oligonucleotide sequence (5' → 3') | $T_m$ (°C) | Length (bp) |
|---|---|---|---|
| pKM173/177_F | GGT ATA AAA GAC CAC CGT CC | 52.7°C | 20 |
| pKM173/177_R | GAA CAG CTA GAG TGC GTT GG | 55.9°C | 20 |
| pKM173VP2/6_F1 | AAC GTT GAA GTG GAG TTT C | 50.4°C | 19 |
| pKM173VP2/6_F2 | GTA TTT ACA GTG GCT TCC | 54.4°C | 20 |
| pKM173VP2/6_R1 | GTG AGT CTG ATT GTG GTG C | 53.6°C | 19 |
| pKM173VP2/6_R2 | TGG TGG TCT CAT CAA CTG | 51.6°C | 18 |
| InternalVP2(WT)_F | GTT GTT GAC TCA GAA ACG G | 51.3°C | 19 |

*All primers were synthesized by Integrated DNA technologies (IDT)

The post reaction clean-up was done using the EDTA/ethanol precipitation protocol. The sequencing reaction was adjusted to 20 µl and transferred to an Eppendorf tube containing 125 mM EDTA and absolute ethanol. The sample was mixed by vortexing it for 5 seconds

and then precipitated at room temperature. To remove contaminants the solution was centrifuged at 20 000 x g (4°C), for 5 minutes. The supernatant was completely aspirated and the pellet was washed with 200 µl 70% ethanol, to remove salts and small organic molecules, and centrifuged at 20 000 x g (4°C) for 5 minutes. The supernatant was discarded and the pellet dried for 5 minutes using a Speed-Vac. Samples were stored at 4°C, in the dark (due to the photosensitive Big Dye), until the sequencing procedure was performed.

An *in silico* sequence of every construct was designed with DNAMAN Version 6 (Lynnon Corporation, 2005) computer software, prior to cloning of the construct, to facilitate the screening of mutations and other mistakes occurring in the construct. The sequences were analysed and edited with FinchTV and the assembly of the construct was confirmed by aligning it with the *in silico* sequence.

### 3.2.4 Expression of proteins

For the expression of the desired proteins, yeast strains were prepared as described in section 3.2.4.1 and 3.2.4.2, and the transformation of the competent cells was done as described in section 3.2.4.3. The desired strain (which contains the recombinant plasmid for expression) was used to inoculate 5 ml Yeast extract peptone dextrose (YPD – 10 mg/ml yeast extract, 20 mg/ml peptone, 20 mg/ml dextrose) medium containing the appropriate antibiotics for the plasmid and host strain. All cultures except for *Kluyveromyces marxianus* (grown at 37ºC) were grown at 30°C with shaking at 180rpm for 24 hours. The cells were harvested by centrifugation and screened for expression as described in section 3.2.4.4. The overnight cultures were also used for genomic DNA isolation (described in section 3.2.4.5) to see whether or not the genome segments of interest integrated into the different yeast strains.

#### 3.2.4.1 Preparation of *Kluyveromyces lactis, Candida deformans, Saccharomyces cerevisiae, Yarrowia lipolytica, Arxula adeninivorans, Hansenula polymorpha* and *Debaryomyces hansenii* competent cells using the Bicine method

Nucleic acids do not enter yeast cells freely, but require assistance traversing the outer and inner cell membranes and in reaching an intracellular site where they can be expressed and replicated (Sambrook and Russell, 2001). Therefore, transformation is the introduction of exogenous DNA into cells and the subsequent inheritance and expression of that DNA. The

Bicine method (Klebe *et al.,* 1983) has the advantage that yeast cells can be made competent for transformation in advance, allowing large batches of cells to be made and saving time. Yeast cells are made competent for transformation by treatment with ethylene glycol and dimethyl sulfoxide (DMSO) and then stored at -80°C.

All competent cells used in this study were prepared using the Bicine method, except for *Kluyveromyces marxianus*. Glycerol stocks of the several yeast strains were streaked out on YPD agar plates and grown for 16 hours at 30°C. Each yeast strain was inoculated into 100 ml of YPD medium, respectively. The cultures were grown until an OD of $A_{600}$ was between 0.6-0.8nm. The cells were harvested by centrifugation (in two 50 ml Falcon tubes) at 1000 x g for 5 minutes. The supernatant was discarded and the cell pellets of both tubes were combined and washed with 50 ml solution 1 (1 M sorbitol, 10 mM bicine pH 8.35, 3% ethylene glycol and 5% DMSO). The centrifugation step was repeated and the cell pellets were re-suspended in 2 ml solution 1. The suspension was dispensed in 200 µl aliquots into pre-chilled, sterile microfuge tubes. The cells were frozen for 4-8 hours at -20°C and then transferred and stored at -80°C. Slow freezing of cells results in good viability of cells.

### 3.2.4.2 Preparation of *Kluyveromyces marxianus* competent cells using the one step method

This method is faster, simpler and easier to perform than other competent cell methods. It incorporates four substances namely carrier DNA, Lithium acetate (LiAc), Polyethylene glycol (PEG) and Dithiothreitol (DTT). The method is based on 4 principles namely that i) PEG is essential since it “dries” the cell with YPD which adds an influx ii) LiAc and heat shock enhances the transformation efficiency and makes the cell membrane porous; iii) DTT makes the cell membrane porous and iv) the highest efficiency is obtained when the cells are at the mid-log phase.

*Kluyveromyces marxianus* competent cells were prepared in the laboratory using the one step method. The glycerol stock of the *Kluyveromyces marxianus* strain was streaked onto YPD agar plates, and grown for 16 hours at 37°C. One 5 ml culture of YPD medium was inoculated with *Kluyveromyces marxianus*. The culture was grown until an OD of between 0.6-0.8nm at $A_{600}$ was obtained. The cells were harvested by centrifugation (800 µl) at 16 000 x g for 1 minute and the pellet used for transformation.

### 3.2.4.3 Transformation of yeast strains

#### 3.2.4.3.1 Linearization of DNA for transformation

Restriction enzyme digestion reactions were performed in reaction volumes of 10 µl. Each reaction mixture consisted of the applicable restriction enzyme (NotI) and buffer, approximately 500-1000 ng DNA and nuclease free water to a final volume of 10 µl. Reaction mixtures were incubated at 37°C for approximately 1.5 hours. The digested samples were analysed with agarose gel electrophoresis (section 2.2.3.4) and further used for transformation.

#### 3.2.4.3.2 Bicine method transformation

The transformation reaction consisted of 1-5 ng linearized DNA reaction and 50 µg carrier DNA (10 mg/ml stock) in 200 µl cell suspension. Frozen chemical competent cells were removed from storage at -80°C and thawed on ice. The ligation reaction/carrier DNA was added to the competent cells and gently mixed. The cell mixture was incubated at 37°C for 5 minutes (when fresh cells were used) and at 37°C until melting is completed (when frozen cells were used). The cells were suspended in solution 2 (200 mM bicine pH 8 and 40% PEG 1000) and incubated for 1 hour at 30°C. After incubation cells were pelleted at 5 000 x g for 5 seconds and the supernatant discarded. The pellets were washed with solution 3 (0.15 M NaCl and 10 mM bicine pH 8.35) and pelleted at 5 000 x g for 5 seconds. The cell pellets were re-suspended in 200 µl solution 3 and plated on selective (600 $mg.l^{-1}$ hygromycin B) YPD plates. The plates were left for fifteen minutes at room temperature to absorb the liquid and then incubated upside-down at 30°C until colonies formed.

#### 3.2.4.3.3 One step method transformation

The transformation reaction consisted of 500 ng linearized DNA reaction and 100 µl pre-heated fresh one step buffer (OSB - 60% PEG4000, 2 M lithium acetate, 1 M DTT and 2 µg/µl salmon sperm). The fresh chemical competent cell pellet was disrupted when the ligation reaction DNA was added to the competent cells. The cell pellet/DNA ligation mixture was re-suspended with the one step buffer (60% PEG4000, 2 M lithium acetate, 1 M DTT and 2 µg/µl salmon sperm) and incubated at 42°C for 60 minutes. YPD medium was added to the mixture and the cells were recovered at 37°C for 2 hours. After incubation cells were pelleted at 16 000 x g for 30 seconds and the supernatant discarded. The cell pellets were re-suspended in 100 µl remaining supernatant and plated on selective (600 $mg.l^{-1}$

hygromycin B) YPD plates. The plates were left for fifteen minutes at room temperature to absorb the liquid and then incubated upside down at 37°C until colonies formed.

#### 3.2.4.4 Analysis of protein expression

The expression of the proteins was analysed by using two methods, namely sodium dodecyl sulphate polyacrylamide gel electrophoresis and western blot analyses.

##### 3.2.4.4.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SDS-PAGE was done as discussed in Chapter 2, section 2.2.6. After transformation of the several yeast strains, containing the possible protein of interest, colonies were chosen of each yeast strain and inocculated in 5 ml YPD medium overnight. Protein samples were prepared by harvesting 4 ml of the cells (with the construct of interest) at 16 000 x g for 1 min. The supernatant was discarded and 10 µl 4 x Laemmli sample buffer (Biorad) was added to the pellet. The samples were then mixed and boiled for 5 minutes at 98°C. Unless otherwise stated, 10 µl of this mixture was loaded on to the gel.

The loaded gel was electrophoresed in 1 x Tris Glycine SDS (TGS - 25 mM Tris, 2 M Glycine and 0.1% SDS) buffer at a current of 25mA using a Bio-Rad PowerPac Basic system. Electrophoresis was carried out for about 2 hours or until the pink dye front was at the bottom of the gel. The gels were removed from the glass plates and incubated in 1 x Transfer buffer (0.025 M Tris, 0.2 M Glycine and 20% methanol, pH 8.4), for western blot analysis (section 3.2.4.4.2).

##### 3.2.4.4.2 Western blot analysis

The recombinant protein samples were subjected to SDS-PAGE (section 3.2.4.4.1). The proteins were electro-blotted onto the nitrocellulose membrane for 1 hour at 100V in transfer buffer (25 mM Tris, 192 mM Glycine, 20% Methanol). Protein transfer was verified by staining the membrane with Ponceau S solution. The staining solution was removed after protein verification and the nitrocellulose membrane washed with distilled water. Two types of development methods were used for the western blot analysis namely the Pierce ECL western blotting substrate as well as the BCIP-T/NBT substrate. Both will be discussed below. Two antibodies were used (i) VP6/group specific antibodies and (ii) primary goat

polyclonal anti-rotavirus antibody with the secondary antibody, Donkey anti-Goat IgG (Nebraska calf diarrhoea virus (NCDV).

***Pierce ECL Western blotting substrate***

Western blot analysis was performed with the Pierce ECL western blotting substrate to verify the co-expression of the recombinant proteins. After transfer of the proteins onto a nitrocellulose membrane, the membrane was blocked in 5% non-fat milk in TNT buffer (0.05% Tween, 0.2 M NaCl, 0.05 M Tris-HCl [pH 8.5]) for 3.5 hours at 4°C, to prevent non-specific binding of the antibodies to the proteins. The membrane was washed three times with the TNT buffer and incubated at 4°C for 8 hours with the goat polyclonal anti-rotavirus antibody prepared against rotavirus NCDV, which was diluted to 1:1000 in TNT buffer. The membrane was washed three times with TNT buffer and incubated with the secondary antibody donkey horseradish peroxidase conjugated anti-goat IgG, which was specific to the primary antibody used. Incubation of the secondary antibody was diluted to 1:500 in TNT buffer, for 1 hour at 4°C. Following secondary antibody incubation the membrane was washed three times with TNT buffer to remove any traces of remaining antibody. The membrane was developed with the Pierce ECL western blotting substrate. The substrate contained a 1:1 ratio of detection reagent 1 (Peroxide solution) and detection reagent 2 (Luminol enhancer solution). The substrate binds to the horseradish peroxidase enzyme which in turn leads to the decay of oxidised luminal radicals and results in light emission that can be captured on X-ray film. The film was exposed to AGFA rapid developing solution until the protein bands were visible. The film was then fixed for 5 minutes in AFGA Rapid Fixer solution.

***BCIP-T/NBT Western blotting substrate***

Western blot analysis was performed with the BCIP-T/NBT western blotting substrate to verify the co-expression of the proteins. After transfer of the proteins onto a nitrocellulose membrane, the membrane was blocked in 5% non-fat milk in TBST buffer (20 mM Tris-HCl pH 7.4, 137 mM NaCl and 0.1% Tween 20) for 1 hour at 37°C, to prevent non-specific binding of the antibodies to the proteins. The membrane was washed twice, for 5 min, with the TBST buffer and incubated at room temperature for 2 hours with the rotavirus group specific antigen, which was diluted to 1:1000 in TBST buffer. The membrane was washed twice, for 5 minutes, with TBST buffer and incubated with the secondary antibody goat anti-goat IgG-AP, which was specific to the primary antibody used. Incubation of the secondary

antibody was diluted to 1:5000 in TBST buffer, for 1 hour at 37°C. Following secondary antibody incubation the membrane was washed two times with TBST buffer to remove any traces of remaining antibody. The membrane was developed with BCIP-T/NBT western blotting substrate. Incubation was done until bands were visible in the BCIP-T/NBT alkaline phosphatase solution that produces an insoluble blue end product that can be observed visually. BCIP-T (5-bromo-4-chloro-3-indolyl phosphate, p-toluidine salt) is a chromogenic substrate for alkaline phosphatase that enhances a blue colour development when used in conjunction with the oxidant NBT (nitro blue tetrazolium). Alkaline phosphatase catalyses the removal of a phosphate group from BCIP-T, to generate a product that oxidizes and dimerizes to dibromodichloro indigo. The reducing equivalents produced during the dimerization reaction reduce NBT to an insoluble purple dye, diformazan. The substrate was prepared by adding 33 µl BCIP-T (50 mg/ml) and 44 µl NBT (75 mg/ml) to 10 ml alkaline phosphatase (AP) buffer (100 mM Tris-HCl (pH 9.5), 100 mM NaCl and 10 mM $MgCl_2$). Alternatively all steps can also be carried out at 4°C.

#### 3.2.4.5 Genomic DNA isolation

Propagation of yeast containing the construct of interest was inoculated into 5 ml YPD media and grown for 24 hours at 30°C, unless otherwise stated, while shaking. The next day the cultures, containing the appropriate construct, were centrifuged at 16 000 x g for 1 minute and the supernatants discarded.

A commercial genomic isolation kit, ZR Fungal/Bacterial DNA MiniPrep™ (Zymo research), was used to do the genomic DNA isolation. The isolation was done according to the instructions of the manufacturer. This kit is designed for the simple, rapid isolation of DNA from though-to-lyse fungi as well as from mycelium and bacteria. This system incorporates a ZR BashingBead™ lyses tube that rapidly and efficiently lyses fungal/bacterial samples by beads beating without using organic denaturants or proteinases. DNA is isolated using fast-spin column technology that is ideal for downstream molecular-based applications such as PCR, array etc.

This procedure consists of five steps namely i) preparation and lysis of yeast cell cultures, ii) DNA binding, iii) pre-wash wash step, iv) wash step and v) elution. The bashing beads minimize DNA shearing by chemical and physical methods since the beads are fracture resistant and chemically inert. The yeast cells are lysed and pelleted and the released genomic DNA is bound to the Zymo-Spin™ IIC column in the presence of the DNA Binding buffer**.** Under these conditions only the genomic DNA will bind to Zymo-Spin™ IIC column

while most of the contaminating RNA and cellular proteinaceous components are removed in the flow through. The bound DNA is then washed to remove any remaining impurities. Genomic DNA was eluted in 50 µl DNA elution buffer.

### 3.3 Results and Discussion

As mentioned previously, this chapter focuses on the introduction and expression of the open reading frames of the wild type genome segment 2 (VP2) and genome segment 6 (VP6) of the GR10924 G9P[6] rotavirus strain in several yeast strains. The cloning strategy used to generate the dual expression vector, pKM173_VP2/6, for expression is indicated in Figure 3.2A and B.

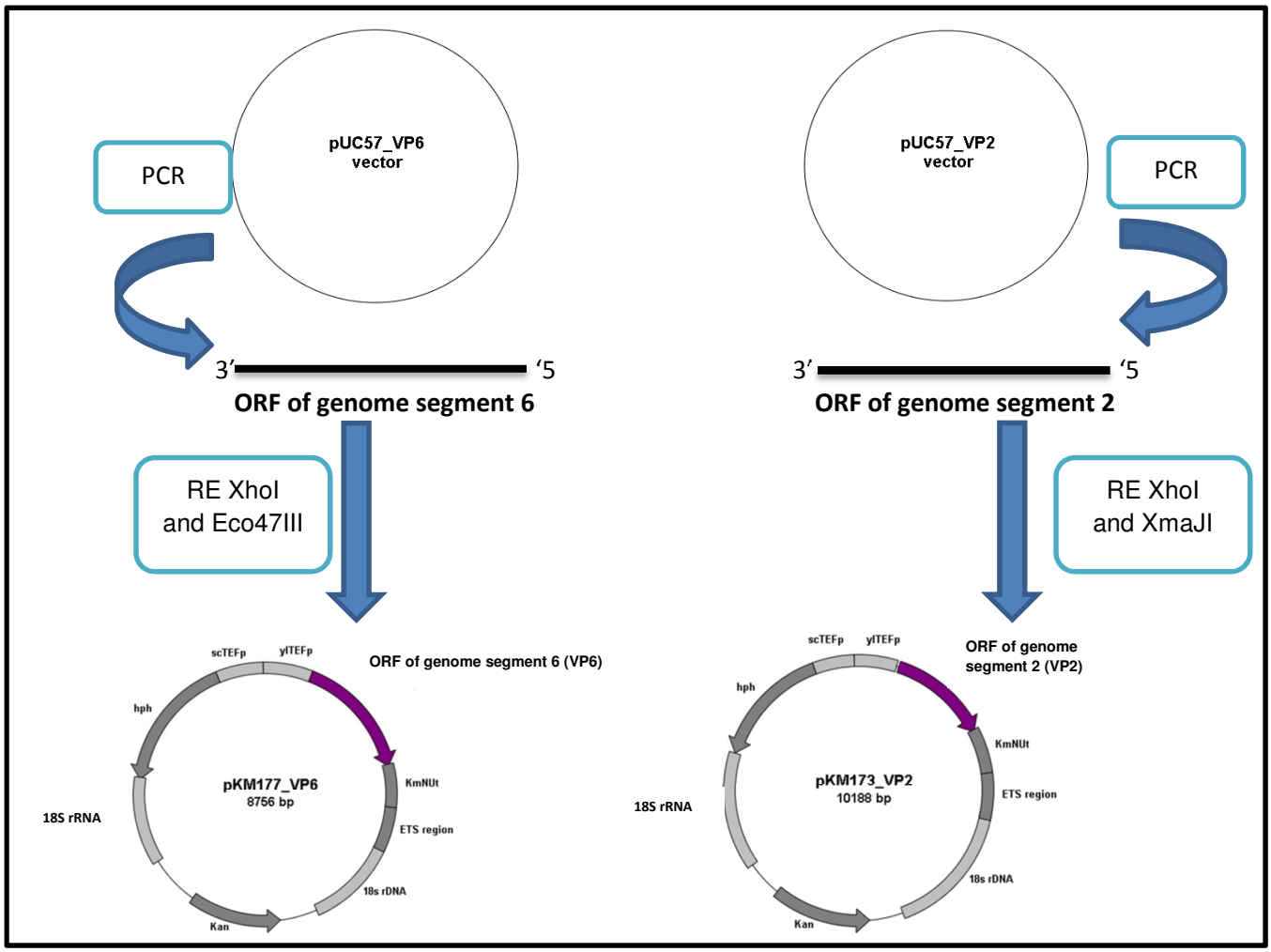


**Figure 3.2A: A schematic illustration of the cloning strategy to generate pKM173_VP2 and pKM177_VP6.** PCR was used to amplify the ORF encoding VP2 (genome segment 2) and VP6 (genome segment 6) from pUC57. The ORF of genome segment 2 and genome segment 6 as well as their respective plasmids, pKM173 and pKM177 were digested with the appropriate enzymes (ORF of genome segment 2 and pKM173: XhoI and XmaJI: ORF of genome segment 6 and pKM177: XhoI and Eco47III). The ligation reactions yielded the vectors, pKM173_VP2 and pKM177_VP6.

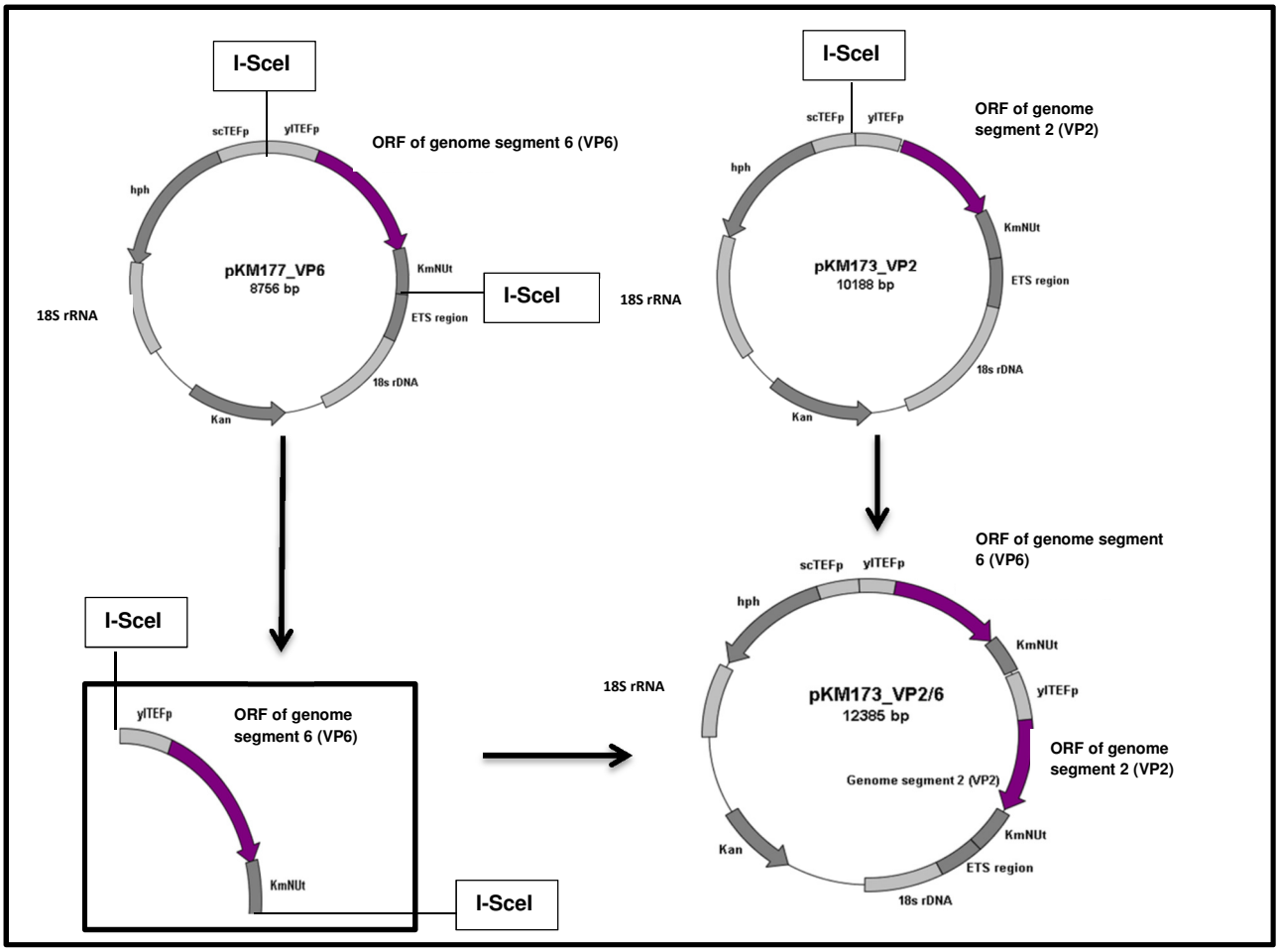


**Figure 3.2B: A schematic illustration of the cloning strategy to generate the dual expression vector for expression in several yeast strains.** Both pKM173_VP2 and pKM177_VP6 were both digested with I-SceI and pKM173_VP2 dephosphorylated before ligation reactions which yielded the expression vector, pKM173_VP2/6.

### 3.3.1 Cloning of the open reading frame of genome segment 2 (VP2) into pKM173

The wild type ORF of genome segment 2 was purchased from GenScript which provided the ORF of genome segment 2 in the plasmid pUC57.

In order to clone the ORF of genome segment 2 (VP2), into the expression vector (pKM173, Figure 3.1), it was PCR amplified from pUC57_VP2 with primers that were designed to contain the restriction enzymes, XhoI and XmaJI, to facilitate directional cloning. The enzyme sites are contained in the multiple cloning site of the cloning vector pKM173 to facilitate the insertion of the amplified ORF of genome segment 2 (VP2). The primers were designed using the wild type sequence of genome segment 2 (VP2) (Figure 3.3).

Restriction enzyme **XhoI**

Forward primer 5′-CTCA**CTCGAG**ATGG CGTACAGGAA AC-3′

```
1       GGCTATTAAA GGCTCAATGG CGTACAGGAA ACGTGGAGCG CGCCGTGAGG CGAACTTAAA
        CCGATAATTT CCGAGTTACC GCATGTCCTT TGCACCTCGC GCGGCACTCC GCTTGAATTT

61      TAATAATGAC CGAATGCAGG AGAAAATTGA TGAAAAACAA GATTCAAATA AAATACAATT
        ATTATTACTG GCTTACGTCC TCTTTTAACT ACTTTTTGTT CTAAGTTTAT TTTATGTTAA

121     ATCCGATAAG GTACTTTCGA AGAAAGAAGA AATTGTAACG GATAGTCATG AGGAAGTTAA
        TAGGCTATTC CATGAAAGCT TCTTTCTTCT TTAACATTGC CTATCAGTAC TCCTTCAATT

181     AGTTACTGAT GAGTTAAAAA AATCAACGAA AGAAGAATCA AAACAATTGC TTGAAGTGTT
        TCAATGACTA CTCAATTTTT TTAGTTGCTT TCTTCTTAGT TTTGTTAACG AACTTCACAA

241     GAAAACAAAG GAAGAACATC AGAAAGAAAT ACAGTATGAA ATATTACAGA AAACTATACC
        CTTTTGTTTC CTTCTTGTAG TCTTTCTTTA TGTCATACTT TATAATGTCT TTTGATATGG

301     AACATTCGAA CCTAAAGAGA CGATATTGAG AAAATTAGAG GATATTCAAC CAGAACTAGC
        TTGTAAGCTT GGATTTCTCT GCTATAACTC TTTTAATCTC CTATAAGTTG GTCTTGATCG

361     GAAAAAACAG ACTAAGTTAT TTAGAATATT TGAACCGAAA CAATTACCGA TTTATAGAGC
        CTTTTTTGTC TGATTCAATA AATCTTATAA ACTTGGCTTT GTTAATGGCT AAATATCTCG

421     AAATGGAGAG AGAGAATTGC GTAATAGATG GTATTGGAAA TTAAAAAAAG ATACACTACC
        TTTACCTCTC TCTCTTAACG CATTATCTAC CATAACCTTT AATTTTTTTC TATGTGATGG

481     AGACGGAGAC TATGATGTGA GAGAGTATTT TCTGAATTTG TATGATCAAG TGCTTACTGA
        TCTGCCTCTG ATACTACACT CTCTCATAAA AGACTTAAAC ATACTAGTTC ACGAATGACT

541     AATGCCAGAC TACTTATTAT TGAAAGATAT GGCAGTAGAA AATAAGAACT CTAGGGATGC
        TTACGGTCTG ATGAATAATA ACTTTCTATA CCGTCATCTT TTATTCTTGA GATCCCTACG

601     AGGTAAAGTT GTTGACTCAG AAACGGCTAG TATATGCGAT GCCATATTTC AAGATGAAGA
        TCCATTTCAA CAACTGAGTC TTTGCCGATC ATATACGCTA CGGTATAAAG TTCTACTTCT

661     AACGGAAGGT GCCGTTAGAA GATTCATTGC AGAAATGAGA CAACGTGTGC AAGCTGATAG
        TTGCCTTCCA CGGCAATCTT CTAAGTAACG TCTTTACTCT GTTGCACACG TTCGACTATC

721     AAATGTTGTC AATTATCCAT CAATATTACA TCCAATAGAT TATGCATTTA ATGAATACTT
        TTTACAACAG TTAATAGGTA GTTATAATGT AGGTTATCTA ATACGTAAAT TACTTATGAA

781     TTTACAACAT CAATTGGTTG AACCATTGAA TAATGATATA ATATTTAATT ATATACCAGA
        AAATGTTGTA GTTAACCAAC TTGGTAACTT ATTACTATAT TATAAATTAA TATATGGTCT

841     AAGGATAAGA AATGATGTTA ATTATATTCT CAATATGGAC AGAAATTTAC CATCAACTGC
        TTCCTATTCT TTACTACAAT TAATATAAGA GTTATACCTG TCTTTAAATG GTAGTTGACG

901     CAGATATATA AGACCTAATT TACTTCAAGA TAGATTAAAT TTGCACGATA ATTTTGAATC
        GTCTATATAT TCTGGATTAA ATGAAGTTCT ATCTAATTTA AACGTGCTAT TAAAACTTAG

961     ACTATGGGAT ACAATAACTA CATCAAATTA TATTTTGGCG AGATCGGTAG TACCAGATTT
        TGATACCCTA TGTTATTGAT GTAGTTTAAT ATAAAACCGC TCTAGCCATC ATGGTCTAAA

1021    AAAGGAATTA GTGTCAACGG AAGCACAAAT TCAGAAAATG TCACAAGATT TGCAATTAGA
        TTTCCTTAAT CACAGTTGCC TTCGTGTTTA AGTCTTTTAC AGTGTTCTAA ACGTTAATCT

1081    AGCATTAACA ATTCAGTCAG AAACACAATT TCTAACAGGT ATAAATTCAC AAGCAGCTAA
        TCGTAATTGT TAAGTCAGTC TTTGTGTTAA AGATTGTCCA TATTTAAGTG TTCGTCGATT

1141    CGATTGTTTT AAAACCTTAA TTGCAGCAAT GTTAAGTCAA CGTACTATGT CATTAGATTT
        GCTAACAAAA TTTTGGAATT AACGTCGTTA CAATTCAGTT GCATGATACA GTAATCTAAA
```

```
1201    TGTAACTACT AATTATATGT CATTGATTTC AGGTATGTGG CTATTGACGG TTGTGCCAAA
        ACATTGATGA TTAATATACA GTAACTAAAG TCCATACACC GATAACTGCC AACACGGTTT

1261    TGATATGTTT ATAAGGGAAT CGTTAGTCGC GTGTCAACTA GCTATAGTAA ATACAATAAT
        ACTATACAAA TATTCCCTTA GCAATCAGCG CACAGTTGAT CGATATCATT TATGTTATTA

1321    CTATCCAGCA TTTGGAATGC AACGAATGCA TTATAGAAAC GGGGATCCAC AAACACCGTT
        GATAGGTCGT AAACCTTACG TTGCTTACGT AATATCTTTG CCCCTAGGTG TTTGTGGCAA

1381    TCAGATAGCA GAACAGCAGA TTCAAAATTT CCAAGTCGCA AATTGGTTAC ATTTTGTTAA
        AGTCTATCGT CTTGTCGTCT AAGTTTTAAA GGTTCAGCGT TTAACCAATG TAAAACAATT

1441    TAATAATCAA TTTAGACAGG CAGTTATTGA TGGTGTATTG AATCAGGTAC TGAATGACAA
        ATTATTAGTT AAATCTGTCC GTCAATAACT ACCACATAAC TTAGTCCATG ACTTACTGTT

1501    TATTAGAAAT GGTCATGTTA TTAACCAACT GATGGAGGCT CTAATGCAGC TGTCGCGACA
        ATAATCTTTA CCAGTACAAT AATTGGTTGA CTACCTCCGA GATTACGTCG ACAGCGCTGT

1561    ACAATTTCCA ACCATGCCAA TTGATTATAA GAGATCAATT CAACGTGGAA TATTACTGTT
        TGTTAAAGGT TGGTACGGTT AACTAATATT CTCTAGTTAA GTTGCACCTT ATAATGACAA

1621    ATCTAACAGA CTTGGTCAGT TAGTTGATTT AACTAGATTA TTAGCTTACA ATTATGAGAC
        TAGATTGTCT GAACCAGTCA ATCAACTAAA TTGATCTAAT AATCGAATGT TAATACTCTG

1681    ATTAATGGCA TGCATTACAA TGAACATGCA ACATGTTCAA ACCTTAACAA CAGAAAAATT
        TAATTACCGT ACGTAATGTT ACTTGTACGT TGTACAAGTT TGGAATTGTT GTCTTTTTAA

1741    ACAATTAACG TCAGTTACAT CATTATGTAT GCTTATTGGA AATGCGACTG TTATACCAAG
        TGTTAATTGC AGTCAATGTA GTAATACATA CGAATAACCT TTACGCTGAC AATATGGTTC

1801    TCCACAAACA TTATTTCATT ATTATAACGT TAACGTTAAT TTTCATTCAA ATTACAATGA
        AGGTGTTTGT AATAAAGTAA TAATATTGCA ATTGCAATTA AAAGTAAGTT TAATGTTACT

1861    GAGAATTAAT GATGCAGTAG CTATAATAAC TGCTGCTAAC AGACTGAATC TATATCAGAA
        CTCTTAATTA CTACGTCATC GATATTATTG ACGACGATTG TCTGACTTAG ATATAGTCTT

1921    AAAAATGAAG GCTATTGTTG AGGATTTCTT AAAAAGATTA TACATTTTTG ATGTATCTAG
        TTTTTACTTC CGATAACAAC TCCTAAAGAA TTTTTCTAAT ATGTAAAAAC TACATAGATC

1981    AGTTCCGGAC GACCAAATGT ATAGATTAAG GGATAGATTA CGCTTATTGC CAGTAGAAAT
        TCAAGGCCTG CTGGTTTACA TATCTAATTC CCTATCTAAT GCGAATAACG GTCATCTTTA

2041    CAGAAGATTA GATATCTTCA ATCTAATACT AATGAACATG GATCAAATTG AACGTGCCTC
        GTCTTCTAAT CTATAGAAGT TAGATTATGA TTACTTGTAC CTAGTTTAAC TTGCACGGAG

2101    AGATAAAATT GCTCAAGGTG TAATCATTGC TTATCGTGAC ATGCATCTTG AAAGAGATGA
        TCTATTTTAA CGAGTTCCAC ATTAGTAACG AATAGCACTG TACGTAGAAC TTTCTCTACT

2161    GATGTACGGA TATGTAAATA TAGCTAGAAA TTTAGAGGGA TTTCAACAGA TAAATTTAGA
        CTACATGCCT ATACATTTAT ATCGATCTTT AAATCTCCCT AAAGTTGTCT ATTTAAATCT

2221    GGAGCTGATG AGATCAGGTG ACTATGCGCA AATAACTAAC ATGCTTTTGA ATAATCAACC
        CCTCGACTAC TCTAGTCCAC TGATACGCGT TTATTGATTG TACGAAAACT TATTAGTTGG

2281    AGTAGCATTG GTTGGAGCAC TTCCATTTAT TACTGATTCA TCAGTTATAT CGCTAATAGC
        TCATCGTAAC CAACCTCGTG AAGGTAAATA ATGACTAAGT AGTCAATATA GCGATTATCG

2341    AAAACTTGAC GCTACAGTGT TCGCTCAAAT AGTTAAATTA CGAAAAGTTG ATACTTTAAA
        TTTTGAACTG CGATGTCACA AGCGAGTTTA TCAATTTAAT GCTTTTCAAC TATGAAATTT
```

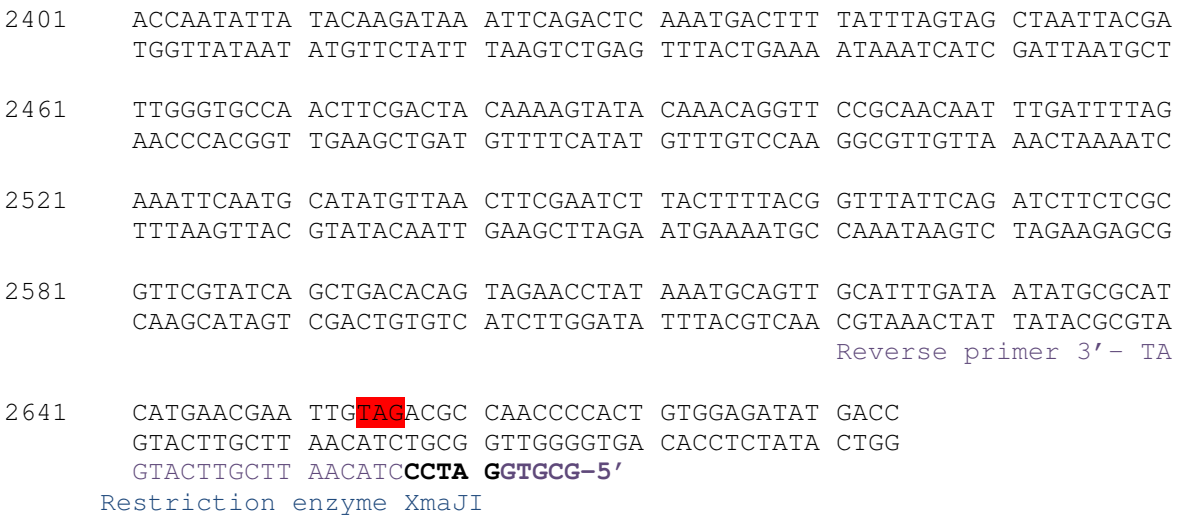


**Figure 3.3: Sequence of the wild type ORF of genome segment 2 (GR10924 G9P[6]) indicating the primers used for amplification.** The forward and reverse primers are indicated in purple and the restriction enzymes sites in black. The start codon is indicated in green and the stop codon in red.

In order to PCR amplify the open reading frame of genome segment 2 from pUC57 to facilitate cloning into the expression vector (pKM173); optimisation of the annealing temperatures had to be done. A temperature gradient ranging from 55°C - 65°C, was used. A 5 µl sample of the PCR annealing temperature optimisation was analysed on a 1% agarose gel, as shown in Figure 3.4.

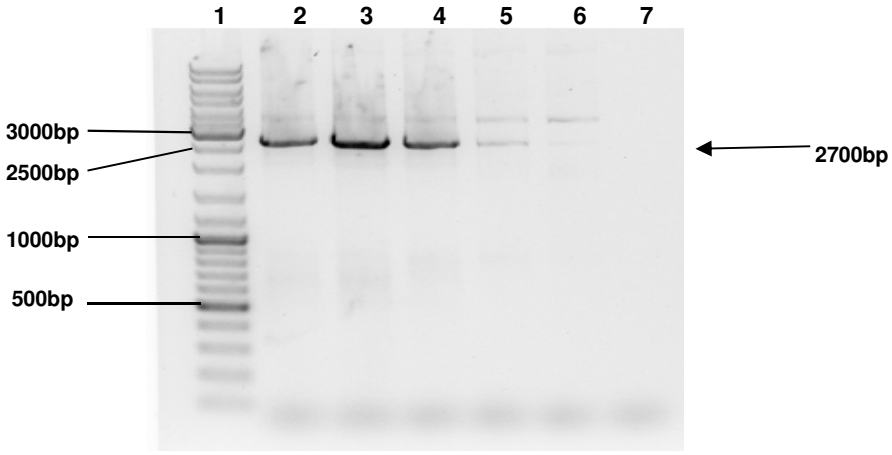


**Figure 3.4: Analysis by agarose gel electrophoresis of the PCR amplification temperature gradient of the ORF of genome segment 2 (VP2) from pUC57.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl volume of each 50 µl PCR reaction of the ORF of genome segment 2 was loaded as follows; lanes: 2) amplification reaction at 55°C; 3) amplification reaction at 57°C; 4) amplification reaction at 60°C; 5) amplification reaction at 63°C; 6) reaction at 65°C; 7) no template control.

The temperature gradient amplification of the ORF of genome segment 2 from pUC57 was successful as seen in Figure 3.4. Lanes 2 to 6 show a 2700bp amplicon as expected, however, the yield of amplicon at 63°C and 65°C (lane 5 and lane 6) was very low. The lack of a band in lane 7, which serves as the negative control (the negative control has no template DNA and is usually water instead of DNA), indicates that no contamination took place. The ORF of genome segment 2 was subsequently amplified in duplicate as described in section 3.2.3.2, with the annealing temperature at 57°C since the band was the clearest at this temperature during PCR annealing temperature optimisation. A 5 µl sample of the duplicate amplification reaction, of the ORF of genome segment 2, was analysed on a 1% agarose gel, as shown in Figure 3.5.

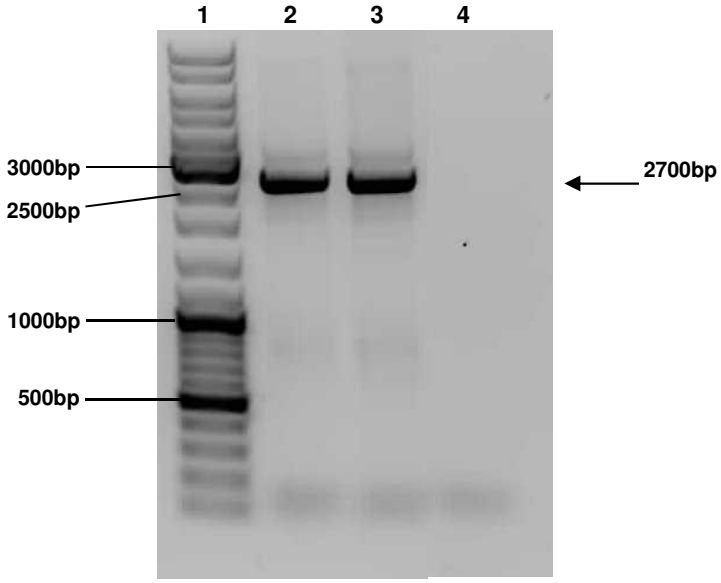


**Figure 3.5: Analysis by agarose gel electrophoresis of the PCR amplification of the ORF of genome segment 2 (VP2) from pUC57 at 57°C.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl of each 50 µl PCR reaction was loaded as follows; lanes: 2-3) ORF of genome segment 2 amplification reactions (done in duplicate); 4) negative control.

The PCR amplification was successful as seen in Figure 3.5. Lane 2 to 3 show a 2700bp amplicon as expected (the same reaction done in duplicate). The lack of a band in lane 4, which serves as the negative control, indicates that no contamination took place. The two PCR amplified ORF of genome segment 2 reactions were pooled and purified by removing oligonucleotides and other contaminations by means of PCR clean-up to ensure good quality DNA for cloning, as described in section 2.2.3.3 (Chapter 2). This yielded about 64.3 ng/µl product of the ORF of genome segment 2 (VP2), which was stored at 4°C for cloning. A 5 µl sample of the purified ORF of genome segment 2 was analysed using a 1% agarose gel, as shown in Figure 3.6.

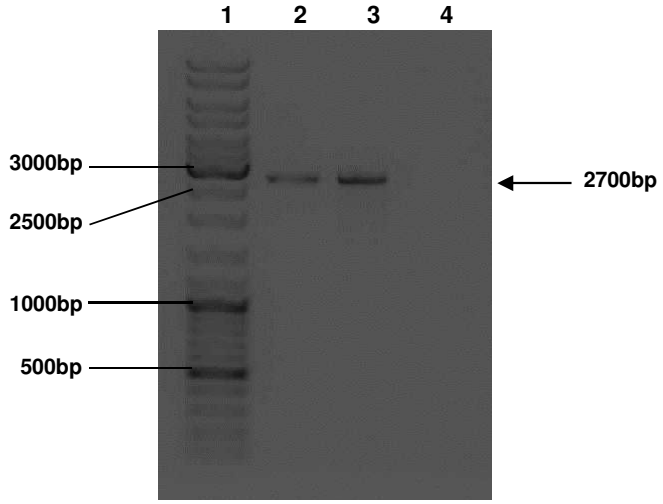


**Figure 3.6: Analysis by agarose gel electrophoresis of the PCR clean-up of the amplified ORF of genome segment 2.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl reaction of the 50 µl gel extracted reaction was loaded as follows; lanes: 2-3) ORF of genome segment 2.

Double restriction endonuclease digestion was done with XhoI and XmaJI, as described in section 3.2.3.3., for the amplified ORF of genome segment 2 and pKM173. These digestions were analysed by a 1% agarose gel, as shown in Figure 3.7. The expected results for the restriction digestion were for pKM173 to give two bands and the ORF of genome segment 2 to give one band, since the enzymes only cuts both (pKM173 and ORF of genome segment 2) at one site. Fragments of 7461bp and 57bp were expected for pKM173 and a fragment of 2684bp was expected for the ORF of genome segment 2 (VP2).

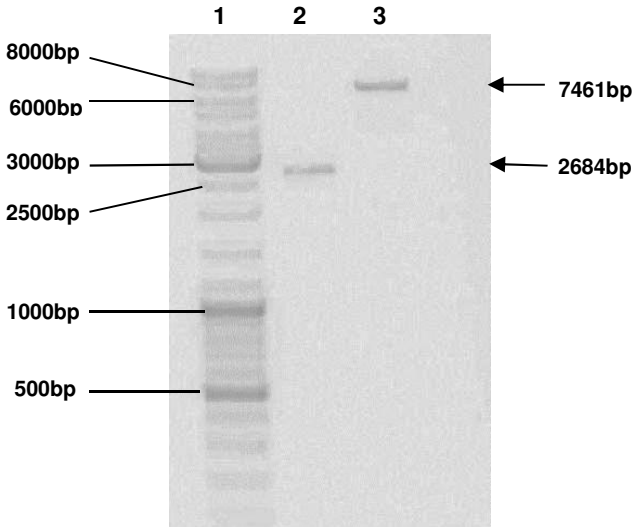


**Figure 3.7: Analysis by agarose gel electrophoresis XhoI and XmaJI restriction enzyme digestions of pKM173 and the ORF of genome segment 2.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl volume of 30 µl XhoI *and* XmaJI digestion reaction was loaded as follows; lanes: 2) ORF of genome segment 2; 3) pKM173 vector.

A band at 7461bp for the restriction digest of pKM173 was obtained (Figure 3.7) as well as a band at 2684bp for the restriction digest of the ORF of genome segment 2 (Figure 3.7), indicating that the expected results were obtained. However, the second band for pKM173 at 57bp could not be observed on the agarose gel due to the small size of the band. The remaining sample of the double restriction digestions were analysed on a 0.8% gel and purified by means of gel extraction. This resulted in about 35 µl gel purified product of both the ORF of genome segment 2 and pKM173 (26.2 ng/µl ORF of genome segment 2 and 59.2 ng/µl pKM173). The ligation reaction of the ORF of genome segment 2 into pKM173 was done as described in section 2.2.3.8 (Chapter 2). The ligation background control contained no DNA insert. Transformation of competent JM109 cells using above mentioned ligation reaction mixture was done. More than 100 colonies were obtained. Twenty colonies of possible pKM173_VP2 plasmids were chosen to perform colony screening on by means of PCR, as described in section 3.2.3.5. A colony was considered positive if the PCR reaction gave an amplicon of 2684bp when analysed on a 1% agarose gel. The results of the PCR colony screening of possible pKM173_VP2 plasmids are shown in Figure 3.8.

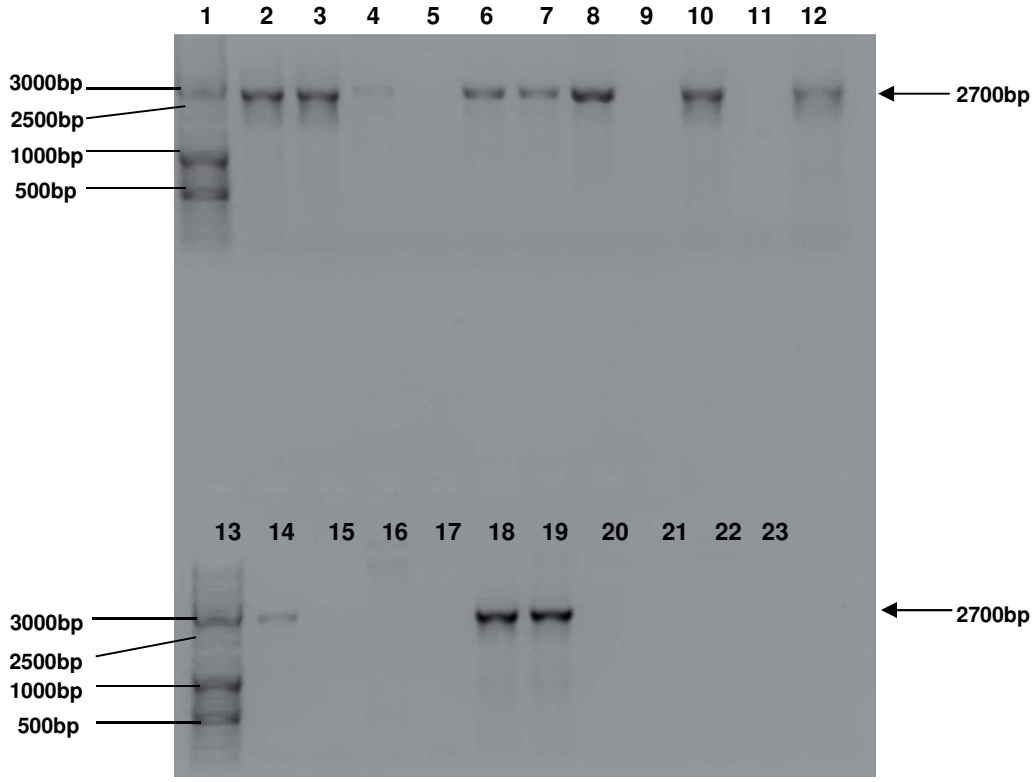


**Figure 3.8: Analysis by agarose gel electrophoresis of the PCR colony screening of the ORF of genome segment 2 (VP2) for possible pKM173_VP2 plasmids.** Lanes 1 and 13) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl volume of each 25 µl pKM173_VP2 colony screening reaction was loaded as follows; lanes: 2) Colony 1; 3) Colony 2; 4) Colony 3; 5) Colony 4; 6) Colony 5; 7) Colony 6; 8) Colony 7; 9) Colony 8; 10) Colony 9; 11) Colony 10; 12) Colony 11; 14) Colony 12; 15) Colony 13; 16) Colony 14; 17) Colony 15; 18) Colony 16; 19) Colony 17; 20) Colony 18; 21) Colony 19; 22) Colony 20; 23) No template control.

Colonies 1-3, 5-7, 9, 11-12, and 16-17 (Lanes 2-4, 6-8, 10, 12, 14, and 18-19 ) gave the expected results of a 2684bp amplicon. However, colonies 4, 8, 10, 13-15 and 19-20 (Lanes 5-7, 9, 11, 15-17 and 21-22) showed no visible band at 2684bp. Lane 23 was the negative control containing no DNA template. Mini-preparations of plasmid DNA and restriction enzyme digestions were done on overnight LB medium cultures of the colonies that gave the expected result with the PCR colony screening to see if the ORF of genome segment 2 was cloned into the pKM173 vector. The plasmid map that would result when the ORF of genome segment 2 was cloned into pKM173, is shown in Figure 3.9.

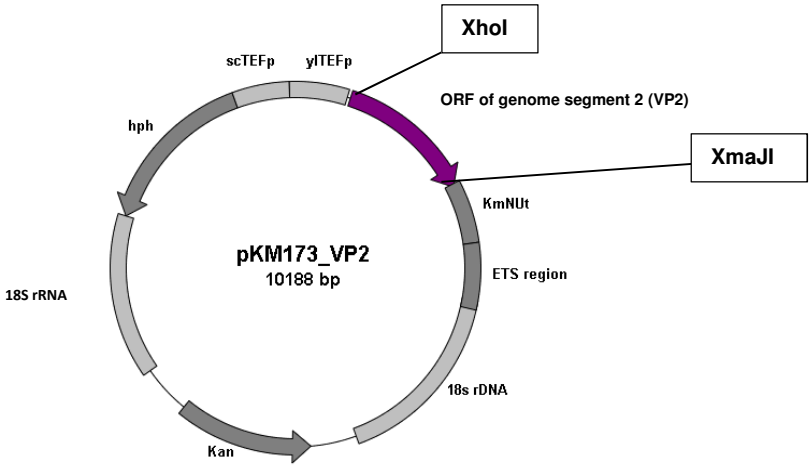


**Figure 3.9: Plasmid map that would result when the ORF of genome segment 2 is cloned into the pKM173 vector.** The XhoI and XmaJI recognition sites indicate the enzymes used to clone the ORF of genome segment 2 (indicated in purple) into the pKM173 vector. The *Yarrowia lipolytica* translation elongation factor (ylTEF) promoter is indicated in light grey, and the *kanamycin* resistance gene in dark grey. The Hygromycin B resistance gene (hph) is the antibiotic used during yeast transformation and the KmlNut is the *Kluyvermomyces marxianus* inulinase target.

The samples that gave the expected result with the PCR colony screening were subjected to a double restriction endonuclease digestion with the restriction enzymes XhoI and XmaJI, since these were the enzymes used for cloning. A 10 µl sample of each 25 µl double restriction digestion was analysed by a 1% agarose gel, as shown in Figure 3.10. A colony was considered positive if two fragments (7461bp and 2684bp) were visible after the double restriction digestion reaction.

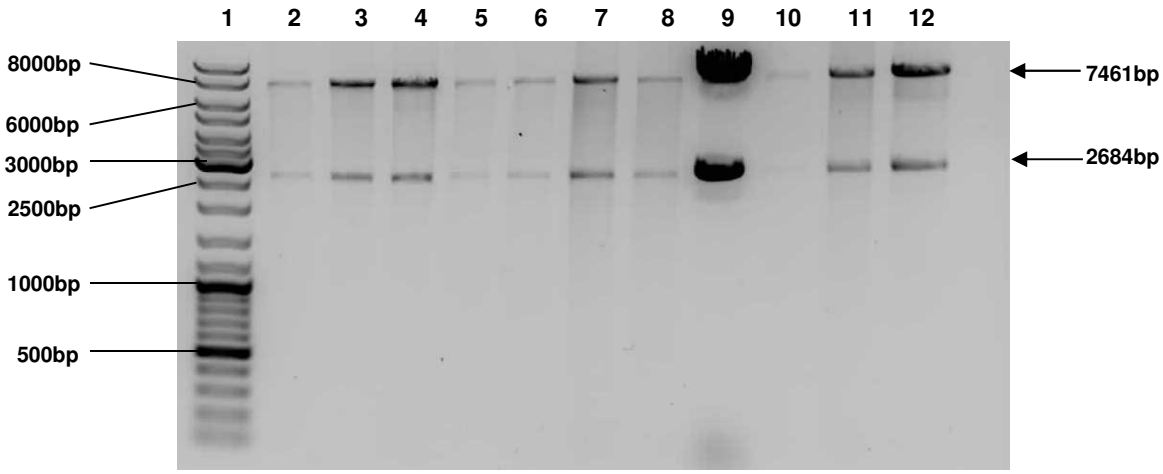


**Figure 3.10: Analysis by agarose gel electrophoresis of XhoI and XmaJI restriction enzyme digestion of possible pKM173_VP2 plasmids.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 25 µl XhoI and XmaJI digestion reaction of possible pKM173_VP2 plasmids was loaded as follows; lanes: 2) Colony 2; 3) Colony 3; 4) Colony 4; 5) Colony 6; 6) Colony 7; 7) Colony 8; 8) Colony 10; 9) Colony 12; 10) Colony 14; 11) Colony 18; 12) Colony 19.

All eleven colonies (colonies 1-3, 5-7, 9, 11-12, and 16-17) gave the expected result namely two bands, one at 7461bp and the other one at 2684bp. A midi-preparation of the plasmid DNA was done for two colonies (colony 2 - lane 3 and colony 3 - lane 4), as described in section 3.2.3.1. A 10 µl sample of each 400 µl plasmid preparation was analysed on a 1% agarose gel, as shown in Figure 3.11.

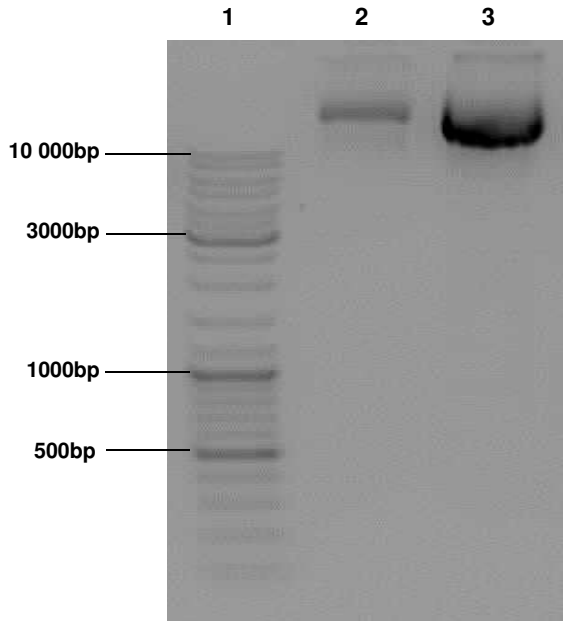


**Figure 3.11: Analysis by agarose gel electrophoresis of plasmid extraction of pKM173_VP2.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 400 µl plasmid extraction was loaded as follows; lanes: 2) pKM173_VP2 colony 2; 3) pKM173_VP2 colony 3.

The results in Figure 3.11 showed that plasmid DNA was obtained for both pKM173_VP2 colonies. However, to make sure that the ORF of genome segment 2 was cloned into pKM173 successfully, both colonies were subjected to a restriction enzyme digestion with XbaI. This enzyme was chosen to confirm the correct insertion of the ORF of genome segment 2 into pKM173, since the restriction enzyme (XbaI) has a recognition site within the ORF of genome segment 2. A 10 µl sample of each 25 µl restriction digestion was analysed on a 1% agarose gel (Figure 3.12). Two fragments (1823bp and 8280bp) were expected for pKM173_VP2, after restriction digestion.

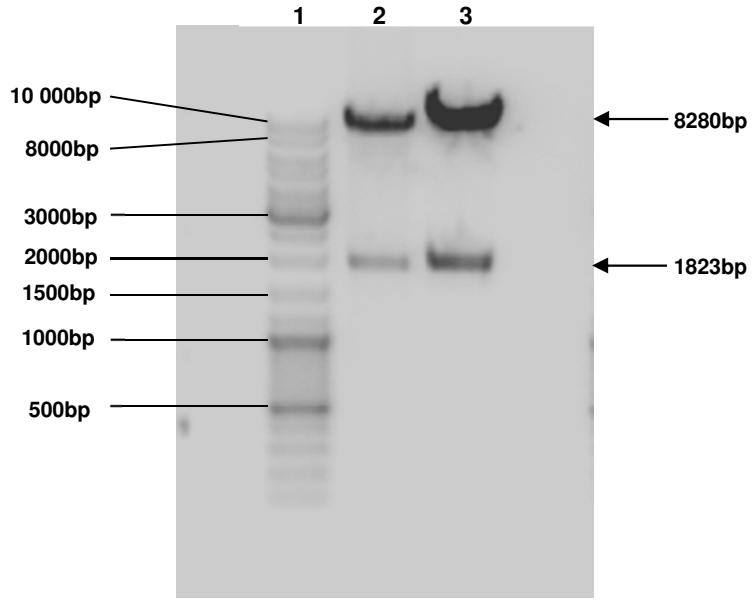


**Figure 3.12: Analysis by agarose gel electrophoresis of XbaI restriction enzyme digestion of pKM173_VP2 plasmids following midi-plasmid preparation.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 25 µl XbaI digestion reaction was loaded as follows; lanes: 2) pKM173_VP2 colony 2; 3) pKM173_VP2 colony 3.

Figure 3.12 shows that the expected results were obtained for both colonies namely two fragments (1823bp and 8280bp). This provides further conformation that the ORF of genome segment 2 was cloned into pKM173. Although these results were encouraging sequencing was still necessary to ensure successful cloning. The midi-plasmid preparation of both pKM173_VP2 colonies was sent for sequencing, to ensure that the ORF of genome segment 2 was cloned into pKM173 and that no mutations had occurred. Primers designed for pKM173 sequencing (Table 3.4) were used. The sequence alignment, of only pKM173_VP2 colony 2, is shown in Figure 3.13.

```
pKM173_VP2      TATAAAAGACCACCGTCCCCGAATTACCTTTCCTCTTCTT    360
pKM173_VP2_F    ..........................CTTTCCTCTTCTT     13
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0


pKM173_VP2      TTCTCTCTCTCCTTGTCAACTCACACCCGAAATGCTCGAG    400
pKM173_VP2_F    TTCTCTCTCTCCTTGTCAACTCACACCCGAAATGCTCGAG     52
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0


pKM173_VP2      ATGGCGTACAGGAAACGTGGAGCGCGCCGTGAGGCGAACT    440
pKM173_VP2_F    ATGGCGTACAGGAAACGTGGAGCGCGCCGTGAGGCGAACT     92
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0


pKM173_VP2      TAAATAATAATGACCGAATGCAGGAGAAAATTGATGAAAA    480
pKM173_VP2_F    TAAATAATAATGACCGAATGCAGGAGAAAATTGATGAAAA    132
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0


pKM173_VP2      ACAAGATTCAAATAAAATACAATTATCCGATAAGGTACTT    520
pKM173_VP2_F    ACAAGATTCAAATAAAATACAATTATCCGATAAGGTACTT    172
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0


pKM173_VP2      TCGAAGAAAGAAGAAATTGTAACGGATAGTCATGAGGAAG    560
pKM173_VP2_F    TCGAAGAAAGAAGAAATTGTAACGGATAGTCATGAGGAAG    212
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0


pKM173_VP2      TTAAAGTTACTGATGAGTTAAAAAAATCAACGAAAGAAGA    600
pKM173_VP2_F    TTAAAGTTACTGATGAGTTAAAAAAATCAACGAAAGAAGA    252
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0


pKM173_VP2      ATCAAAACAATTGCTTGAAGTGTTGAAAACAAAGGAAGAA    640
pKM173_VP2_F    ATCAAAACAATTGCTTGAAGTGTTGAAAACAAAGGAAGAA    292
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0


pKM173_VP2      CATCAGAAAGAAATACAGTATGAAATATTACAGAAAACTA    680
pKM173_VP2_F    CATCAGAAAGAAATACAGTATGAAATATTACAGAAAACTA    332
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0


pKM173_VP2      TACCAACATTCGAACCTAAAGAGACGATATTGAGAAAATT    720
pKM173_VP2_F    TACCAACATTCGAACCTAAAGAGACGATATTGAGAAAATT    372
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0
```

```
pKM173_VP2     AGAGGATATTCAACCAGAACTAGCGAAAAAACAGACTAAG    760
pKM173_VP2_F   AGAGGATATTCAACCAGAACTAGCGAAAAAACAGACTAAG    412
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0

pKM173_VP2     TTATTTAGAATATTTGAACCGAAACAATTACCGATTTATA    800
pKM173_VP2_F   TTATTTAGAATATTTGAACCGAAACAATTACCGATTTATA    452
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0

pKM173_VP2     GAGCAAATGGAGAGAGAGAATTGCGTAATAGATGGTATTG    840
pKM173_VP2_F   GAGCAAATGGAGAGAGAGAATTGCGTAATAGATGGTATTG    492
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0

pKM173_VP2     GAAATTAAAAAAAGATACACTACCAGACGGAGACTATGAT    880
pKM173_VP2_F   GAAATTAAAAAAAGATACACTACCAGACGGAGACTATGAT    532
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0

pKM173_VP2     GTGAGAGAGTATTTTCTGAATTTGTATGATCAAGTGCTTA    920
pKM173_VP2_F   GTGAGAGAGTATTTTCTGAATTTGTATGATCAAGTGCTTA    572
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0

pKM173_VP2     CTGAAATGCCAGACTACTTATTATTGAAAGATATGGCAGT    960
pKM173_VP2_F   CTGAAATGCCAGACTACTTATTATTGAAAGATATGGCAGT    612
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0

pKM173_VP2     AGAAAATAAGAACTCTAGGGATGCAGGTAAAGTTGTTGAC   1000
pKM173_VP2_F   AGAAAATAAGAACTCTAGGGATGCAGGTAAAGTTGTTGAC    652
pKM173_VP2_Int ........................................      0
pKM173_VP2_R   ........................................      0

pKM173_VP2     TCAGAAACGGCTAGTATATGCGATGCCATATTTCAAGATG   1040
pKM173_VP2_F   TCAGAAACGGCTAGTATATGCGATGCCATATTTCAAGATG    692
pKM173_VP2_Int .....................................AGATG      5
pKM173_VP2_R   ........................................      0

pKM173_VP2     AAGAAACGGAAGGTGCCGTTAGAAGATTCATTGCAGAAAT   1080
pKM173_VP2_F   AAGAAACGGAAGGTGCCGTTAGAAGATTCATTGCAGAAAT    732
pKM173_VP2_Int AAGAAACGGAAGGTGCCGTTAGAAGATTCATTGCAGAAAT     45
pKM173_VP2_R   ........................................      0

pKM173_VP2     GAGACAACGTGTGCAAGCTGATAGAAATGTTGTCAATTAT   1120
pKM173_VP2_F   GAGACAACGTGTGCAAGCTGATAGAAATGTTGTCAATTAT    772
pKM173_VP2_Int GAGACAACGTGTGCAAGCTGATAGAAATGTTGTCAATTAT     85
pKM173_VP2_R   ........................................      0
```

```
pKM173_VP2     CCATCAATATTACATCCAATAGATTATGCATTTAATGAAT   1160
pKM173_VP2_F   CCATCAATATTACATCCAATAGATTATGCATTTAATGAAT    812
pKM173_VP2_Int CCATCAATATTACATCCAATAGATTATGCATTTAATGAAT    125
pKM173_VP2_R   ....................................        0

pKM173_VP2     ACTTTTTACAACATCAATTGGTTGAACCATTGAATAATGA   1200
pKM173_VP2_F   ACTTTTTACAACATCAATTGGTTGAACCATTGAATAATGA    852
pKM173_VP2_Int ACTTTTTACAACATCAATTGGTTGAACCATTGAATAATGA    165
pKM173_VP2_R   ....................................        0

pKM173_VP2     TATAATATTTAATTATATACCAGAAAGGATAAGAAATGAT   1240
pKM173_VP2_F   TATAATATTTAATTATATACCAGAAAGGATAAGAAATGAT    892
pKM173_VP2_Int TATAATATTTAATTATATACCAGAAAGGATAAGAAATGAT    205
pKM173_VP2_R   ....................................        0

pKM173_VP2     GTTAATTATATTCTCAATATGGACAGAAATTTACCATCAA   1280
pKM173_VP2_F   GTTAATTATATTCTCAATATGGACAGAAATTTACCATCAA    932
pKM173_VP2_Int GTTAATTATATTCTCAATATGGACAGAAATTTACCATCAA    245
pKM173_VP2_R   ....................................        0

pKM173_VP2     CTGCCAGATATATAAGACCTAATTTACTTCAAGATAGATT   1320
pKM173_VP2_F   CTGCCAGATATATAAGACCTAATTTACTTCAAGATAGATT    972
pKM173_VP2_Int CTGCCAGATATATAAGACCTAATTTACTTCAAGATAGATT    285
pKM173_VP2_R   ....................................        0

pKM173_VP2     AAATTTGCACGATAATTTTGAATCACTATGGGATACAATA   1360
pKM173_VP2_F   AAATTTGCACGATAATTTTGAATCACTATGGGATACAATA   1012
pKM173_VP2_Int AAATTTGCACGATAATTTTGAATCACTATGGGATACAATA    325
pKM173_VP2_R   ....................................        0

pKM173_VP2     ACTACATCAAATTATATTTTGGCGAGATCGGTAGTACCAG   1400
pKM173_VP2_F   ACTACATCAAATTATATTTTGGCGAGATCGGTAGTACCAG   1052
pKM173_VP2_Int ACTACATCAAATTATATTTTGGCGAGATCGGTAGTACCAG    365
pKM173_VP2_R   ....................................        0

pKM173_VP2     ATTTAAAGGAATTAGTGTCAACGGAAGCACAAATTCAGAA   1440
pKM173_VP2_F   ATTTAAAGGAATTAGTGTCAACGGAAGCACAAATTCAGAA   1092
pKM173_VP2_Int ATTTAAAGGAATTAGTGTCAACGGAAGCACAAATTCAGAA    405
pKM173_VP2_R   ....................................        0

pKM173_VP2     AATGTCACAAGATTTGCAATTAGAAGCATTAACAATTCAG   1480
pKM173_VP2_F   AATGTCACAAGATTTGCAATTAGAAGCATTAACAATTCAG   1132
pKM173_VP2_Int AATGTCACAAGATTTGCAATTAGAAGCATTAACAATTCAG    445
pKM173_VP2_R   ....................................        0

pKM173_VP2     TCAGAAACACAATTTCTAACAGGTATAAATTCACAAGCAG   1520
pKM173_VP2_F   TCAGAAACACAATTTCTAACAGGTA...............   1157
pKM173_VP2_Int TCAGAAACACAATTTCTAACAGGTATAAATTCACAAGCAG    485
pKM173_VP2_R   ....................................        0
```

```
pKM173_VP2     CTAACGATTGTTTTAAAACCTTAATTGCAGCAATGTTAAG   1560
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int CTAACGATTGTTTTAAAACCTTAATTGCAGCAATGTTAAG    525
pKM173_VP2_R   ....................................      0


pKM173_VP2     TCAACGTACTATGTCATTAGATTTTGTAACTACTAATTAT   1600
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int TCAACGTACTATGTCATTAGATTTTGTAACTACTAATTAT    565
pKM173_VP2_R   ....................................      0


pKM173_VP2     ATGTCATTGATTTCAGGTATGTGGCTATTGACGGTTGTGC   1640
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ATGTCATTGATTTCAGGTATGTGGCTATTGACGGTTGTGC    605
pKM173_VP2_R   ....................................      0


pKM173_VP2     CAAATGATATGTTTATAAGGGAATCGTTAGTCGCGTGTCA   1680
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int CAAATGATATGTTTATAAGGGAATCGTTAGTCGCGTGTCA    645
pKM173_VP2_R   ....................................      0


pKM173_VP2     ACTAGCTATAGTAAATACAATAATCTATCCAGCATTTGGA   1720
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ACTAGCTATAGTAAATACAATAATCTATCCAGCATTTGGA    685
pKM173_VP2_R   ....................................      0


pKM173_VP2     ATGCAACGAATGCATTATAGAAACGGGGATCCACAAACAC   1760
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ATGCAACGAATGCATTATAGAAACGGGGATCCACAAACAC    725
pKM173_VP2_R   ....................................      0


pKM173_VP2     CGTTTCAGATAGCAGAACAGCAGATTCAAAATTTCCAAGT   1800
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int CGTTTCAGATAGCAGAACAGCAGATTCAAAATTTCCAAGT    765
pKM173_VP2_R   ....................................      0


pKM173_VP2     CGCAAATTGGTTACATTTTGTTAATAATAATCAATTTAGA   1840
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int CGCAAATTGGTTACATTTTGTTAATAATAATCAATTTAGA    805
pKM173_VP2_R   ....................................      0


pKM173_VP2     CAGGCAGTTATTGATGGTGTATTGAATCAGGTACTGAATG   1880
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int CAGGCAGTTATTGATGGTGTATTGAATCAGGTACTGAATG    845
pKM173_VP2_R   ....................................      0


pKM173_VP2     ACAATATTAGAAATGGTCATGTTATTAACCAACTGATGGA   1920
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ACAATATTAGAAATGGTCATGTTATTAACCAACTGATGGA    885
pKM173_VP2_R   ....................................      0
```

```
pKM173_VP2     GGCTCTAATGCAGCTGTCGCGACAACAATTTCCAACCATG   1960
pKM173_VP2_F   ........................................   1157
pKM173_VP2_Int GGCTCTAATGCAGCTGTCGCGACAACAATTTCCAACCATG    925
pKM173_VP2_R   ........................................      0


pKM173_VP2     CCAATTGATTATAAGAGATCAATTCAACGTGGAATATTAC   2000
pKM173_VP2_F   ........................................   1157
pKM173_VP2_Int CCAATTGATTATAAGAGATCAATTCAACGTGGAATATTAC    965
pKM173_VP2_R   ........................................      0


pKM173_VP2     TGTTATCTAACAGACTTGGTCAGTTAGTTGATTTAACTAG   2040
pKM173_VP2_F   ........................................   1157
pKM173_VP2_Int TGTTATCTAACAGACTTGGTCAGTTAGTTGATTTAACTAG   1005
pKM173_VP2_R   ........................................      0


pKM173_VP2     ATTATTAGCTTACAATTATGAGACATTAATGGCATGCATT   2080
pKM173_VP2_F   ........................................   1157
pKM173_VP2_Int ATTATTAGCTTACAATTATGAGACATTAATGGCATGCATT   1045
pKM173_VP2_R   ........................TAATGGCATGCATT     14


pKM173_VP2     ACAATGAACATGCAACATGTTCAAACCTTAACAACAGAAA   2120
pKM173_VP2_F   ........................................   1157
pKM173_VP2_Int ACAATGAACATGCAACATGTTCAAACCTTAACAACAGAAA   1085
pKM173_VP2_R   ACAATGAACATGCAACATGTTCAAACCTTAACAACAGAAA     54


pKM173_VP2     AATTACAATTAACGTCAGTTACATCATTATGTATGCTTAT   2160
pKM173_VP2_F   ........................................   1157
pKM173_VP2_Int AATTACAATTAACGTCAGTTACATCATTATGTATGCTTAT   1125
pKM173_VP2_R   AATTACAATTAACGTCAGTTACATCATTATGTATGCTTAT     94


pKM173_VP2     TGGAAATGCGACTGTATACCAAGTCCACAAACATTATTTC   2200
pKM173_VP2_F   ........................................   1157
pKM173_VP2_Int TGGAAATGCGACTGTATACCAAGTCCACAAACATTATTTC   1165
pKM173_VP2_R   TGGAAATGCGACTGTATACCAAGTCCACAAACATTATTTC    133


pKM173_VP2     ATTATTATAACGTTAACGTTAATTTTCATTCAAATTACAA   2240
pKM173_VP2_F   ........................................   1157
pKM173_VP2_Int ATTATTATAACGTTAACGTTAATTTTCATTCA........   1197
pKM173_VP2_R   ATTATTATAACGTTAACGTTAATTTTCATTCAAATTACAA    173


pKM173_VP2     TGAGAGAATTAATGATGCAGTAGCTATAATAACTGCTGCT   2280
pKM173_VP2_F   ........................................   1157
pKM173_VP2_Int ........................................   1197
pKM173_VP2_R   TGAGAGAATTAATGATGCAGTAGCTATAATAACTGCTGCT    213


pKM173_VP2     AACAGACTGAATCTATATCAGAAAAAAATGAAGGCTATTG   2320
pKM173_VP2_F   ........................................   1157
pKM173_VP2_Int ........................................   1197
pKM173_VP2_R   AACAGACTGAATCTATATCAGAAAAAAATGAAGGCTATTG    253
```

```
pKM173_VP2     TTGAGGATTTCTTAAAAAGATTATACATTTTTGATGTATC   2360
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ....................................   1197
pKM173_VP2_R   TTGAGGATTTCTTAAAAAGATTATACATTTTTGATGTATC    293


pKM173_VP2     TAGAGTTCCGGACGACCAAATGTATAGATTAAGGGATAGA   2400
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ....................................   1197
pKM173_VP2_R   TAGAGTTCCGGACGACCAAATGTATAGATTAAGGGATAGA    333


pKM173_VP2     TTACGCTTATTGCCAGTAGAAATCAGAAGATTAGATATCT   2440
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ....................................   1197
pKM173_VP2_R   TTACGCTTATTGCCAGTAGAAATCAGAAGATTAGATATCT    373


pKM173_VP2     TCAATCTAATACTAATGAACATGGATCAAATTGAACGTGC   2480
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ....................................   1197
pKM173_VP2_R   TCAATCTAATACTAATGAACATGGATCAAATTGAACGTGC    413


pKM173_VP2     CTCAGATAAAATTGCTCAAGGTGTAATCATTGCTTATCGT   2520
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ....................................   1197
pKM173_VP2_R   CTCAGATAAAATTGCTCAAGGTGTAATCATTGCTTATCGT    453


pKM173_VP2     GACATGCATCTGAAAGAGATGAGATGTACGGATATGTAAA   2560
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ....................................   1197
pKM173_VP2_R   GACATGCATCTGAAAGAGATGAGATGTACGGATATGTAAA    493


pKM173_VP2     TATAGCTAGAAATTTAGAGGGATTTCAACAGATAAATTTA   2600
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ....................................   1197
pKM173_VP2_R   TATAGCTAGAAATTTAGAGGGATTTCAACAGATAAATTTA    533


pKM173_VP2     GAGGAGCTGATGAGATCAGGTGACTATGCGCAAATAACTA   2640
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ....................................   1197
pKM173_VP2_R   GAGGAGCTGATGAGATCAGGTGACTATGCGCAAATAACTA    573


pKM173_VP2     ACATGCTTTTGAATAATCAACCAGTAGCATTGGTTGGAGC   2679
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ....................................   1197
pKM173_VP2_R   ACATGCTTTTGAATAATCAACCAGTAGCATTGGTTGGAGC    613


pKM173_VP2     ACTTCCATTTATTACTGATTCATCAGTTATATCGCTAATA   2719
pKM173_VP2_F   ....................................   1157
pKM173_VP2_Int ....................................   1197
pKM173_VP2_R   ACTTCCATTTATTACTGATTCATCAGTTATATCGCTAATA    653
```

```
pKM173_VP2     GCAAAACTTGACGCTACAGTGTTCGCTCAAATAGTTAAAT   2759
pKM173_VP2_F   .........................................   1157
pKM173_VP2_Int .........................................   1197
pKM173_VP2_R   GCAAAACTTGACGCTACAGTGTTCGCTCAAATAGTTAAAT    693

pKM173_VP2     TACGAAAAGTTGATACTTTAAAACCAATATTATACAAGAT   2799
pKM173_VP2_F   .........................................   1157
pKM173_VP2_Int .........................................   1197
pKM173_VP2_R   TACGAAAAGTTGATACTTTAAAACCAATATTATACAAGAT    733

pKM173_VP2     AAATTCAGACTCAAATGACTTTTATTTAGTAGCTAATTAC   2839
pKM173_VP2_F   .........................................   1157
pKM173_VP2_Int .........................................   1197
pKM173_VP2_R   AAATTCAGACTCAAATGACTTTTATTTAGTAGCTAATTAC    773

pKM173_VP2     GATTGGGTGCCAACTTCGACTACAAAAGTATACAAACAGG   2879
pKM173_VP2_F   .........................................   1157
pKM173_VP2_Int .........................................   1197
pKM173_VP2_R   GATTGGGTGCCAACTTCGACTACAAAAGTATACAAACAGG    813

pKM173_VP2     TTCCGCAACAATTTGATTTTAGAAATTCAATGCATATGTT   2919
pKM173_VP2_F   .........................................   1157
pKM173_VP2_Int .........................................   1197
pKM173_VP2_R   TTCCGCAACAATTTGATTTTAGAAATTCAATGCATATGTT    853

pKM173_VP2     AACTTCGAATCTTACTTTTACGGTTTATTCAGATCTTCTC   2959
pKM173_VP2_F   .........................................   1157
pKM173_VP2_Int .........................................   1197
pKM173_VP2_R   AACTTCGAATCTTACTTTTACGGTTTATTCAGATCTTCTC    893

pKM173_VP2     GCGTTCGTATCAGCTGACACAGTAGAACCTATAAATGCAG   2999
pKM173_VP2_F   .........................................   1157
pKM173_VP2_Int .........................................   1197
pKM173_VP2_R   GCGTTCGTATCAGCTGACACAGTAGAACCTATAAATGCAG    933

pKM173_VP2     TTGCATTTGATAATATGCGCATCATGAACGAATTGTAGCC   3039
pKM173_VP2_F   .........................................   1157
pKM173_VP2_Int .........................................   1197
pKM173_VP2_R   TTGCATTTGATAATATGCGCATCATGAACGAATTGTAGCC    973

pKM173_VP2     TAGGTGATCTGATCTGCTTACTTTACTTAACGACCAAAGA   3079
pKM173_VP2_F   .........................................   1157
pKM173_VP2_Int .........................................   1197
pKM173_VP2_R   TAGGTGATCTGATCTGCTTACTTTACTTAACGACCAAAGA   1013

pKM173_VP2     AAAACGACAAAAAAAAAATATTACTACTATTAAAATAAAT   3119
pKM173_VP2_F   .........................................   1157
pKM173_VP2_Int .........................................   1197
pKM173_VP2_R   AAAACGACAAAAAAAAAATATTACTACTATTAAAATAAAT   1053
```

```
pKM173_VP2      TAGTATTTTTCTCTTCTTACGATATGATATGATGCTATGA   3159
pKM173_VP2_F    ..........................................   1157
pKM173_VP2_Int  ..........................................   1197
pKM173_VP2_R    TAGTATTTTTCTCTTCTTACGATATGATATGATGCTATGA   1093


pKM173_VP2      AATCATCATCTTCTTAACTTTCTTGTCTCTTACACGTCAC   3199
pKM173_VP2_F    ..........................................   1157
pKM173_VP2_Int  ..........................................   1197
pKM173_VP2_R    AATCATCATCTTCTTAACTTTCTTGTCTCTTACACGTCAC   1133
```

**Figure 3.13: Nucleotide sequence alignments of the recombinant pKM173 vector containing the ORF of genome segment 2 (VP2) with the *in silico* clone of pKM173_VP2.** The *in silico* clone of pKM173_VP2 is indicated as pKM173_VP2, the forward sequence is indicated as pKM173_VP2_F nucleotide sequence, the reverse sequence is indicated as pKM173_VP2_R nucleotide sequence and the internal sequence is indicated as pKM173_VP2_Int nucleotide sequence. The restriction enzymes used for the cloning of the ORF of genome segment 2 are indicated in grey. The start codon of the ORF of genome segment 2 is indicated in green and the stop codon is in red.

Both pKM173_VP2 colonies gave the expected result, however, only the sequence results of pKM173_VP2 colony 2 is shown in Figure 3.13. Figure 3.13 provides information that the open reading frame of genome segment 2 was correctly cloned into pKM173 and that no mutations had occurred. The primers designed for sequencing made it possible to see the restriction enzymes used for cloning as well as the start and stop codon of the ORF of genome segment 2. However, a part of the ORF of genome segment 2 (from 1505bp to 206bp) could not be sequenced with the reverse and forward primer designed for sequencing. Therefore, an internal primer set was designed to sequence this part of the ORF of genome segment 2. Figure 3.13 also shows the sequencing results of the internal primer, indicating that the entire ORF of genome segment 2 was correctly cloned into pKM173. All further experiments were performed using pKM173_VP2 colony 2.

### 3.3.2 Cloning of the ORF of genome segment 6 (VP6) into pKM177

The wild type ORF of genome segment 6 was purchased from GenScript which provided the gene in the plasmid pUC57.

In order to clone the ORF of genome segment 6 (VP6), into the pKM177 vector (Figure 3.1), it was PCR amplified from pUC57 with primers that were designed to contain the restriction enzyme sites, XhoI and Eco47III to facilitate directional cloning. The ORF of genome segment 6 did not contain these two restriction enzyme sites and these enzyme sites are contained in the multiple cloning site of the cloning vector pKM177. The primers were designed using the wild type sequence of the ORF of genome segment 6 (VP6) (Figure 3.14).

Restriction enzyme: XhoI

Forward primer 5′-CAAC**CTC** **GAG**ATGGATG TCCTGTAC-3′

```
1       GGCTTTAAAA CGAAGTCTTC AACATGGATG TCCTGTACTC CTTATCAAAA ACTCTTAAAG
        CCGAAATTTT GCTTCAGAAG TTGTACCTAC AGGACATGAG GAATAGTTTT TGAGAATTTC

61      ATGCTAGAGA CAAAATTGTC GAAGGCACAT TATACTCTAA TGTGAGTGAT CTAATTCAAC
        TACGATCTCT GTTTTAACAG CTTCCGTGTA ATATGAGATT ACACTCACTA GATTAAGTTG

121     AATTTAACCA AATGATAATT ACTATGAATG GAAATGAGTT CCAAACTGGA GGAATTGGTA
        TTAAATTGGT TTACTATTAA TGATACTTAC CTTTACTCAA GGTTTGACCT CCTTAACCAT

181     ATCTACCAAT TAGAAATTGG AATTTTGATT TTGGATTACT TGGAACAACT CTACTAAATT
        TAGATGGTTA ATCTTTAACC TTAAAACTAA AACCTAATGA ACCTTGTTGA GATGATTTAA

241     TAGACGCTAA CTACGTCGAA ACAGCCCGTA ACACAATTGA TTATTTTGTA GATTTTGTAG
        ATCTGCGATT GATGCAGCTT TGTCGGGCAT TGTGTTAACT AATAAAACAT CTAAAACATC

301     ATAACGTATG TATGGATGAA ATGGTTAGAG AATCACAAAG AAATGGAATT GCACCACAAT
        TATTGCATAC ATACCTACTT TACCAATCTC TTAGTGTTTC TTTACCTTAA CGTGGTGTTA

361     CAGACTCACT TAGAAAATTG TCAGGCATTA AGTTCAAAAG GATAAATTTT GATAATTCAT
        GTCTGAGTGA ATCTTTTAAC AGTCCGTAAT TCAAGTTTTC CTATTTAAAA CTATTAAGTA

421     CGGAATATAT AGAGAACTGG AATCTACAAA ACAGAAGACA ACGAACAGGT TTTACATTTC
        GCCTTATATA TCTCTTGACC TTAGATGTTT TGTCTTCTGT TGCTTGTCCA AAATGTAAAG

481     ATAAACCAAA TATTTTCCCT TATTCAGCGT CATTCACACT GAATAGATCA CAACCAGCTC
        TATTTGGTTT ATAAAAGGGA ATAAGTCGCA GTAAGTGTGA CTTATCTAGT GTTGGTCGAG

541     ATGATAACTT GATGGGTACA ATGTGGCTGA ACGCAGGATC AGAAATTCAG GTCGCTGGAT
        TACTATTGAA CTACCCATGT TACACCGACT TGCGTCCTAG TCTTTAAGTC CAGCGACCTA

601     TCGACTATTC GTGTGCAATT AATGCGCCAG CTAATACACA ACAATTTGAA CATATTGTAC
        AGCTGATAAG CACACGTTAA TTACGCGGTC GATTATGTGT TGTTAAACTT GTATAACATG

661     AGCTCCGAAG AGTTTTAACT ACAGCTACAA TAACACTTTT ACCGGATGCA GAAAGATTCA
        TCGAGGCTTC TCAAAATTGA TGTCGATGTT ATTGTGAAAA TGGCCTACGT CTTTCTAAGT

721     GTTTTCCAAG AGTGATTAAT TCAGCTGACG GAGCAACTAC ATGGTATTTT AATCCAGTAA
        CAAAAGGTTC TCACTAATTA AGTCGACTGC CTCGTTGATG TACCATAAAA TTAGGTCATT

781     TTCTTAGACC AAATAACGTT GAAGTGGAGT TTCTACTAAA CGGGCAGATA ATAAACACTT
        AAGAATCTGG TTTATTGCAA CTTCACCTCA AAGATGATTT GCCCGTCTAT TATTTGTGAA

841     ACCAGGCTAG ATTTGGAACG ATCGTAGCTA GAAATTTTGA TACAATCAGA TTGTCGTTTC
        TGGTCCGATC TAAACCTTGC TAGCATCGAT CTTTAAAACT ATGTTAGTCT AACAGCAAAG

901     AGTTGATGAG ACCACCAAAT ATGACACCAT CGGTAGCAGC ATTATTTCCA AATGCGCAAC
        TCAACTACTC TGGTGGTTTA TACTGTGGTA GCCATCGTCG TAATAAAGGT TTACGCGTTG

961     CATTTGAACA TCATGCTACA GTAGGACTTA CATTGAAAAT TGAATCTGCA GTTTGTGAAT
        GTAAACTTGT AGTACGATGT CATCCTGAAT GTAACTTTTA ACTTAGACGT CAAACACTTA

1021    CTGTACTTGC TGACGCAAGC GAGACAATGC TAGCAAATGT GACATCTGTT AGACAAGAAT
        GACATGAACG ACTGCGTTCG CTCTGTTACG ATCGTTTACA CTGTAGACAA TCTGTTCTTA

1081    ACGCGATACC AGTTGGACCA GTCTTTCCAC CAGGTATGAA TTGGACTGAT TTGATCACTA
        TGCGCTATGG TCAACCTGGT CAGAAAGGTG GTCCATACTT AACCTGACTA AACTAGTGAT
```

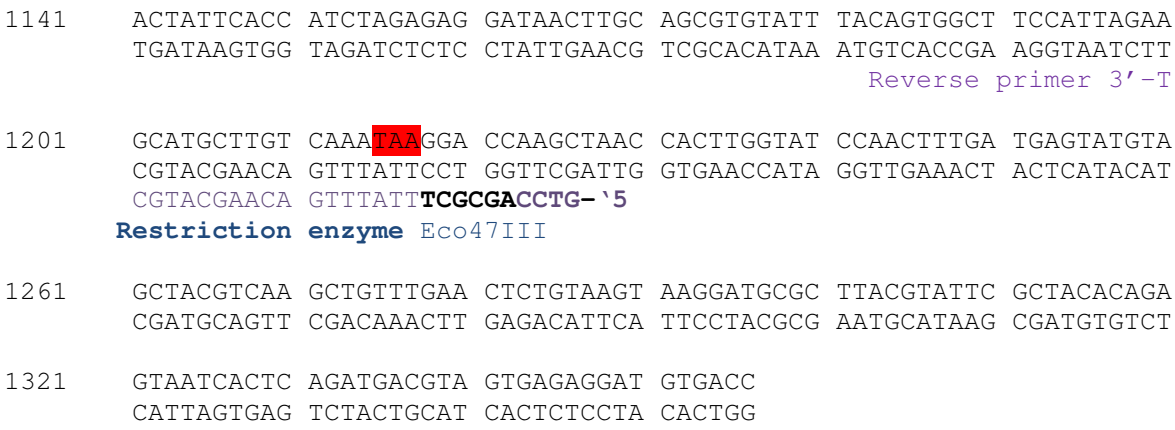


**Figure 3.14: Sequence of wild type ORF of genome segment 6 (GR10924 G9P[6]) indicating the primers used for amplification.** The forward and reverse primers are indicated in purple and the restriction enzymes sites in black. The start codon is indicated in green and the stop codon in red.

In order to obtain the ORF of genome segment 6 amplicon, PCR amplification was done as described in section 2.2.3.2 (Chapter 2). The annealing temperature was 55°C. A 5 µl sample of the PCR amplification reaction was analysed on a 1% agarose gel, as shown in Figure 3.15.

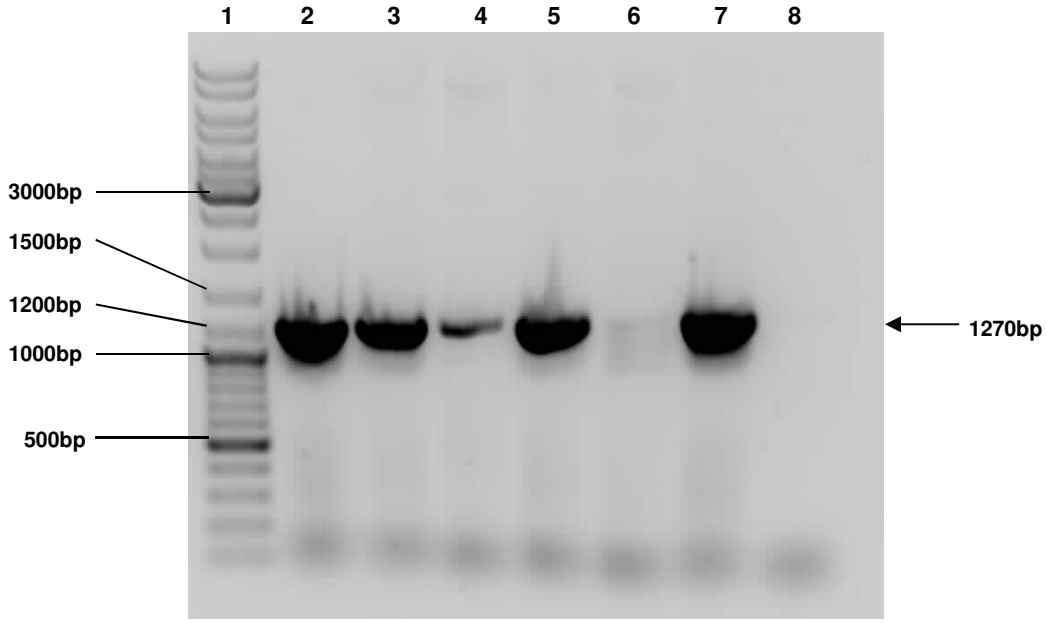


**Figure 3.15: Analysis by agarose gel electrophoresis of the PCR amplification of the ORF of genome segment 6 (VP6) from pUC57.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl of each 50 µl PCR reaction was loaded as follows; lanes: 2-7) the ORF of genome segment 6 amplification reactions (done in duplicate); 8) negative control.

The PCR amplification was successful as seen in Figure 3.15. Lanes 2 to 7 show a 1270bp amplicon as expected (the same reaction done in duplicate). The lack of a band in lane 8, which serves as the negative control, indicates that no contamination took place. The six PCR amplified ORF of genome segment 6 reactions were pooled (three reactions were pooled together) and purified by removing oligonucleotides and other contaminations by means of PCR clean-up to ensure good quality DNA for cloning. This yielded about 120 ng/µl product of the ORF of genome segment 6 (VP6), which was stored at 4°C, for cloning. A 5 µl sample of the purified ORF of genome segment 6 was analysed on a 1% agarose gel, as shown in Figure 3.16.

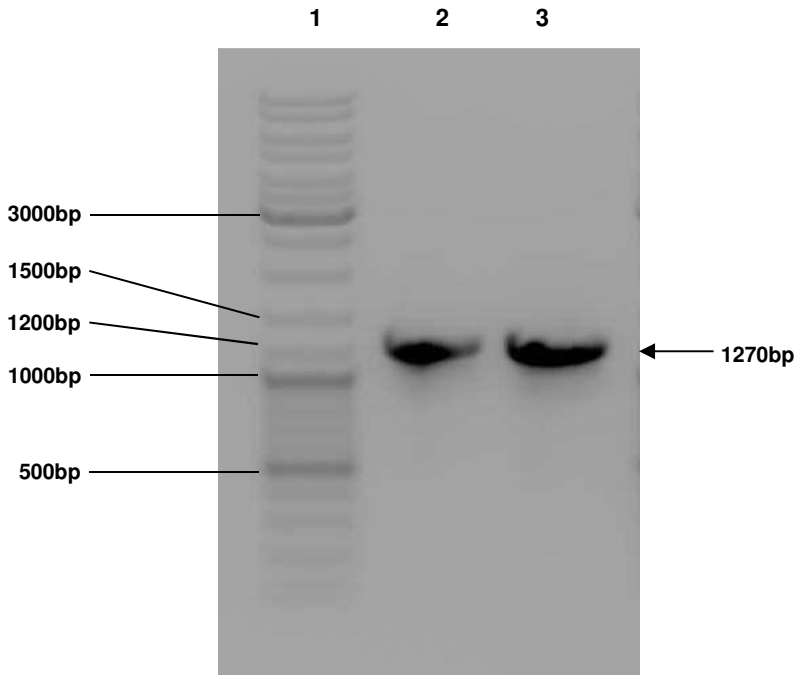


**Figure 3.16: Analysis by agarose gel electrophoresis of the PCR clean-up of the amplified ORF of genome segment 6.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl reaction of each 40 µl gel extracted reaction was loaded as follows; lanes: 2-3) pool of amplification reactions of the ORF of genome segment 6.

Double restriction endonuclease digestion was done with XhoI and Eco47III, as described in section 3.2.3.3, for the amplified ORF of genome segment 6 and pKM177. These digestions were analysed by loading a 5 µl sample of each reaction on a 1% agarose gel, as shown in Figure 3.17. The expected results for the restriction digestion were two bands (7489bp and 77bp) for pKM177 and one band (1190bp) for the ORF of genome segment 6, since the enzymes cuts both (pKM177 and ORF of genome segment 6) at only one site.

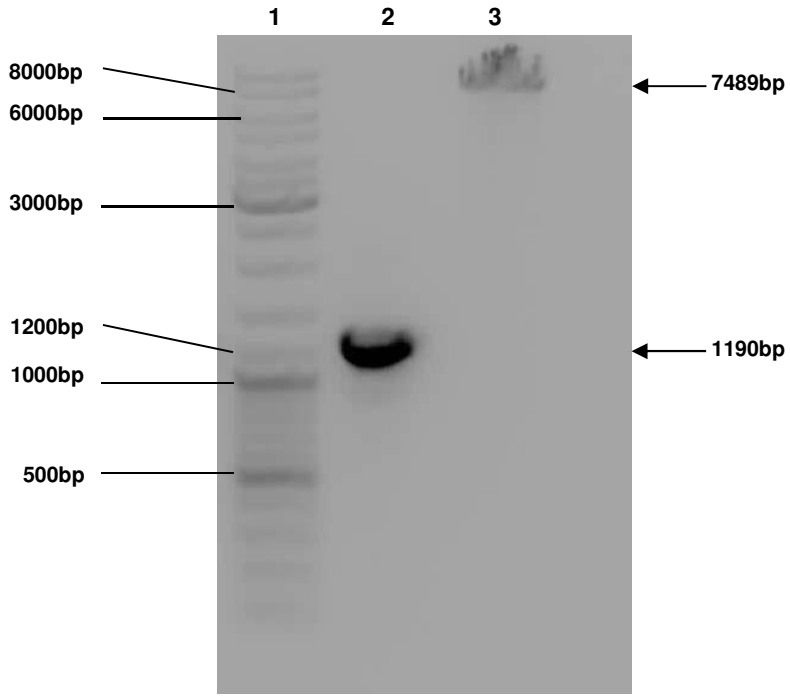


**Figure 3.17: Analysis by agarose gel electrophoresis of XhoI and Eco47III restriction enzyme analysis of pKM177 and the ORF of genome segment 6.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl volume of 30 µl XhoI and Eco47III digestion reactions was loaded as follows; lanes: 2) ORF of genome segment 6; 3) pKM177 vector.

Figure 3.17 provides information that the expected result of a band at 7489bp for the restriction digestion of pKM177 was obtained (lane 3) but the 77bp band could not be observed on the agarose gel due to the small size of the band. The expected result of a band at 1190bp was also obtained for the ORF of genome segment 6, after restriction digestion (Figure 3.17, lane 2). The remaining samples of the double restriction digestions were analysed with a 0.8% agarose gel and purified by means of a gel extraction as described in section 2.2.3.7 (Chapter 2). The double restriction enzyme reaction yielded about 35 µl purified product of both the ORF of genome segment 6 and pKM177 (118.2 ng/µl ORF of genome segment 6 and 42.2 ng/µl pKM177). Ligation of the ORF of genome segment 6 and pKM177 was done as described in section 2.2.3.8 (Chapter 2). The ligation background control contained no DNA insert. Transformation of competent JM109 cells, using the above mentioned ligation reaction mixture, was done as described in section 2.2.3.10 (Chapter 2). Twenty colonies of possible pKM177_VP6 plasmids were chosen to perform colony screening on, by means of PCR, as described in section 3.2.3.5. A colony

was considered positive if the PCR colony screening gave an amplicon of 1190bp when analysed on a 1% agarose gel (Figure 3.18).

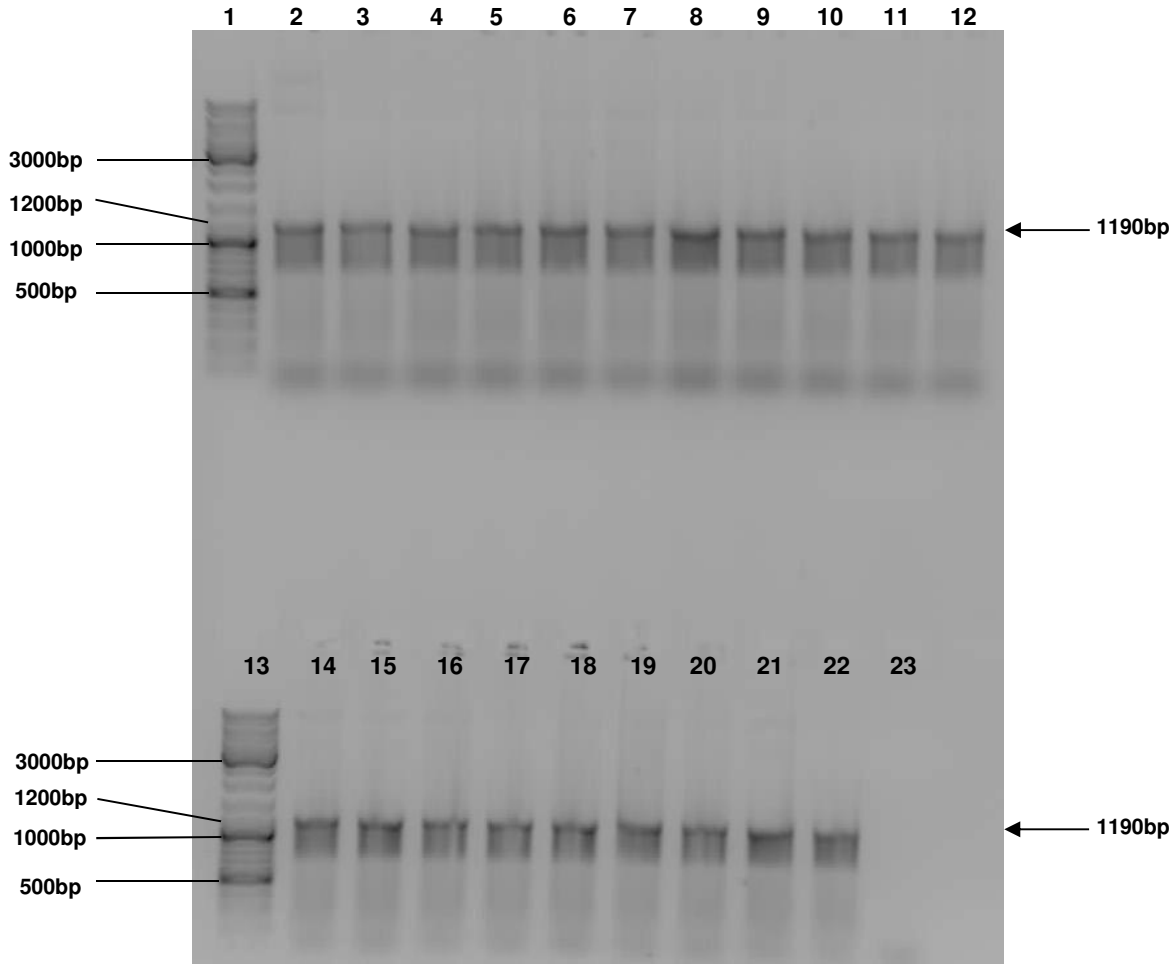


**Figure 3.18: Analysis by agarose gel electrophoresis of PCR colony screening of the ORF of genome segment 6 (VP6) for possible pKM177_VP6.** Lanes 1 and 13) 10 000bp O'Generuler DNA marker (Fermentas). A 5 µl volume of each 25 µl pKM177_VP6 colony screening reaction was loaded as follows; lanes: 2) Colony 1; 3) Colony 2; 4) Colony 3; 5) Colony 4; 6) Colony 5; 7) Colony 6; 8) Colony 7; 9) Colony 8; 10) Colony 9; 11) Colony 10; 12) Colony 11; 14) Colony 12; 15) Colony 13; 16) Colony 14; 17) Colony 15; 18) Colony 16; 19) Colony 17; 20) Colony 18; 21) Colony 19; 22) Colony 20; 23) No template control.

Colonies 1-20 (Lanes 2-12 and 14-22) gave the expected result of a band at 1190bp. Lane 23 was the negative control containing no DNA template. Mini-preparations of plasmid DNA and restriction enzyme digestions were done on overnight LB medium of all 20 colonies, to

see if the ORF of genome segment 6 was cloned into the pKM177 vector. The plasmid map that would result, when the ORF of genome segment 6 is cloned into pKM177, is illustrated in Figure 3.19.

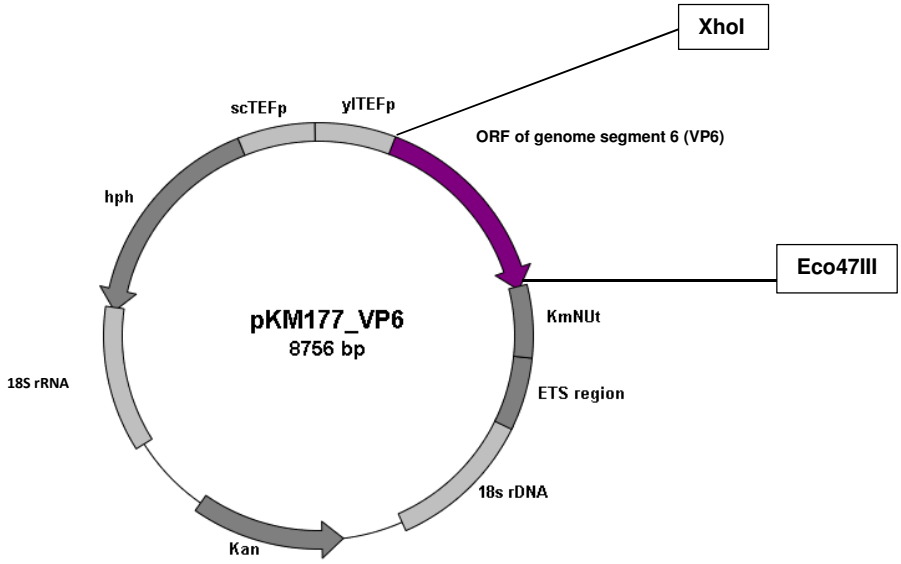


**Figure 3.19: Plasmid map that would result when the ORF of genome segment 6 is cloned into the pKM177 vector.** The XhoI and Eco47III recognition sites indicate the enzymes used to clone the ORF of genome segment 6 (indicated in purple) into the pKM177 vector. The *Yarrowia lipolytica* translation elongation factor (ylTEF) promoter is indicated in light grey, and the *kanamycin* resistance gene in dark grey. The Hygromycin B resistance gene (hph) is the antibiotic used during yeast transformation and the KmlNut is the *Kluyvermomyces marxianus* inulinase target.

Each sample was subjected to a double restriction endonuclease digestion with the restriction enzymes, XhoI and EcoR47III, since these were the enzymes used for cloning. A 10 µl sample of each 25 µl double restriction digestion was analysed on a 1% agarose gel as shown in Figure 3.20. A colony was considered positive if two fragments (7489bp and 1190bp) were visible after the double restriction digestion reaction.

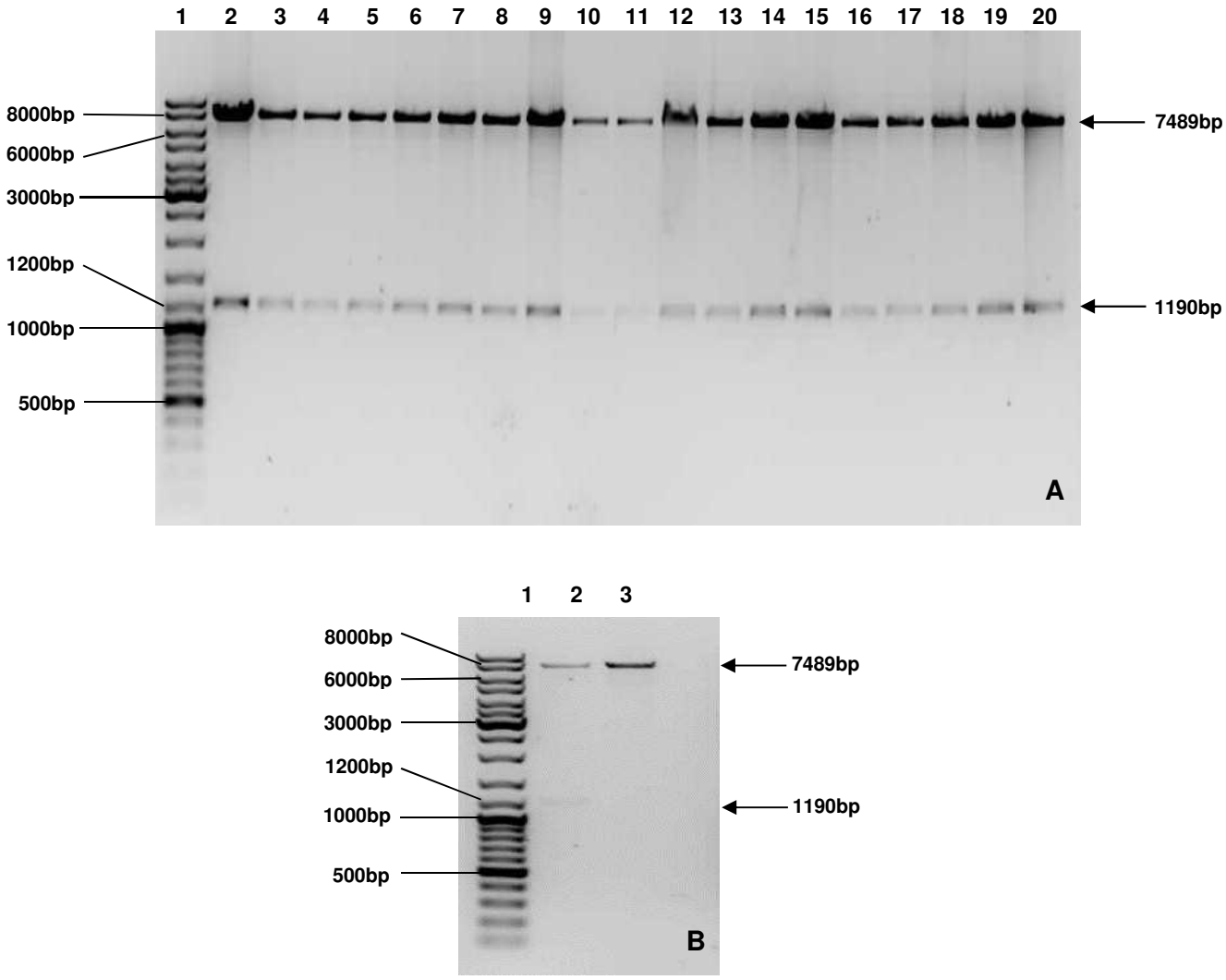


**Figure 3.20: Analysis by agarose gel electrophoresis of XhoI and Eco47III restriction enzyme digestion of possible pKM177_VP6 plasmids. (A)** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 25 µl XhoI and Eco47III pKM177_VP6 digestion reactions were loaded as follows; lanes: 2) Colony 1; 3) Colony 2; 4) Colony 3; 5) Colony 4; 6) Colony 5; 7) Colony 6; 8) Colony 7; 9) Colony 8; 10) Colony 9; 11) Colony 10; 12) Colony 11; 13) Colony 12; 14) Colony 13; 15) Colony 14; 16) Colony 15; 17) Colony 16; 18) Colony 17; 19) Colony 18; 20) Colony 19. **(B)** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 25 µl XhoI and Eco47III pKM177_VP6 digestion reactions were loaded as follows; lanes: 2) Colony 20; 3) Negative control, digested pKM177.

Colonies 1-20 in Figure 3.20A and lane 2 in Figure 3.20B gave the expected result namely two bands one at 7489bp and the other at 1190bp. Lane 3 (Figure 3.20B) served as the negative control namely digested pKM177 vector containing no insert. A midi-preparation of plasmid DNA was done for two colonies (colony 1 - lane 2 and colony 2 – lane 3) as described in section 3.2.3.1. A 10 µl sample of each 400 µl plasmid preparation was analysed on a 1% agarose gel, as shown in Figure 3.21.

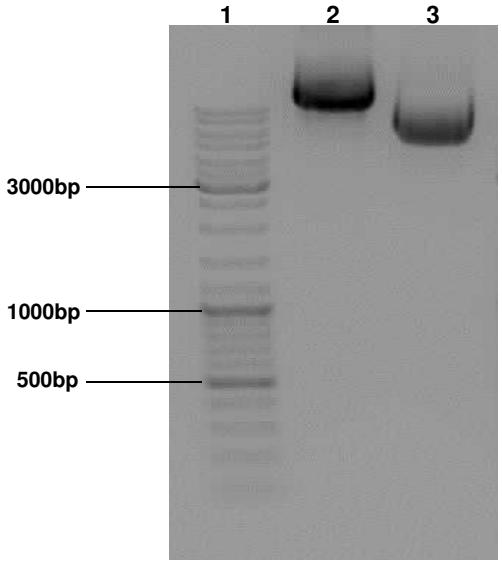


**Figure 3.21: Analysis by agarose gel electrophoresis of plasmid preparation of pKM177_VP6 colonies.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 400 µl plasmid extraction was loaded as follows; lanes: 2) pKM177_VP6 colony 1; 3) pKM177_VP6 colony 2.

The results in Figure 3.21 showed that plasmid DNA was obtained for both pKM177_VP6 colonies. However, to make sure that the ORF of genome segment 6 was cloned into pKM177 successfully, both colonies (colonies 1 and 2) were subjected to a restriction digestion, with XbaI. A 10 µl sample of each 25 µl restriction digestion was analysed on a 1% agarose gel (Figure 3.22). Two fragments (7450bp and 1241bp) were expected for pKM177_VP6 after restriction digestion reaction.

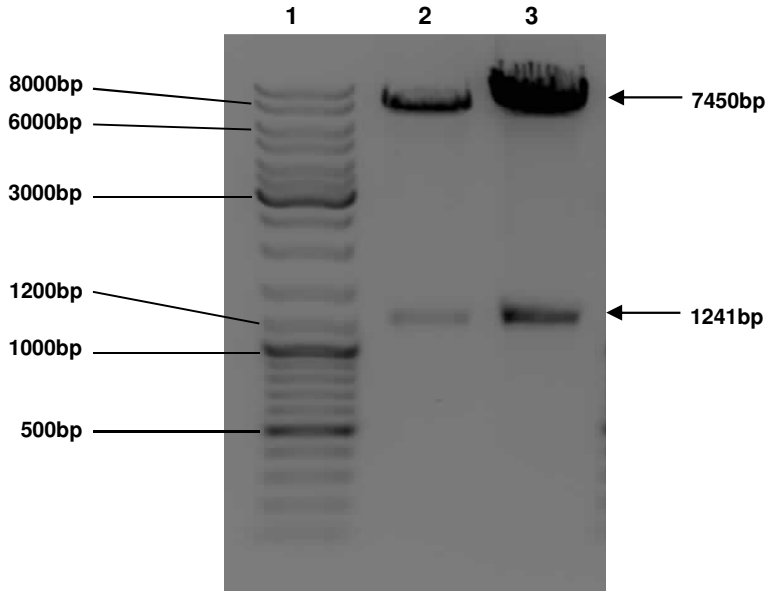


**Figure 3.22: Analysis by agarose gel electrophoresis of XbaI restriction enzyme analysis of pKM177_VP6 plasmids, from the midi-plasmid preparation.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 20 µl XbaI digestion reaction was loaded as follows; lanes: 2) pKM173_VP6 colony 1; 3) pKM173_VP6 colony 2

Figure 3.22 shows that the expected results were obtained for both pKM177_VP6 colonies namely two fragments, one at 7450bp and the other at 1241bp. This indicates that the ORF of genome segment 6 was cloned into pKM177. These results were encouraging but sequencing was still necessary to ensure that the plasmid contained the coding region of interest. The midi-plasmid preparation of both pKM177_VP6 colonies were sent for sequencing, to ensure that the ORF of genome segment 6 was cloned into pKM177 and that no mutations occurred.

Primers designed for the sequencing of pKM177 (Table 3.4) were used, as described in section 2.2.3.12 (Chapter 2). Both pKM177_VP6 colonies were sent for sequencing, however, only the sequence alignment of pKM177_VP6 colony 1 is shown in Figure 3.23.

```
pKM177_VP6    AAAATTTTTTTGCTTTGTGGTTGGGACTTTAGCCAAGGGT    320
pKM177_VP6_F  ........................................      0
pKM177_VP6_R  ........................................      0

pKM177_VP6    ATAAAAGACCACCGTCCCCGAATTACCTTTCCTCTTCTTT    360
pKM177_VP6_F  ..................GAATTACCTTTCCTCTTCTTT     21
pKM177_VP6_R  ........................................      0


pKM177_VP6    TCTCTCTCTCCTTGTCAACTCACACCCGAAATGCTCGAGA    400
pKM177_VP6_F  TCTCTCTCTCCTTGTCAACTCACACCCGAAATGCTCGAGA     61
pKM177_VP6_R  ........................................      0


pKM177_VP6    TGGATGTCCTGTACTCCTTATCAAAAACTCTTAAAGATGC    440
pKM177_VP6_F  TGGATGTCCTGTACTCCTTATCAAAAACTCTTAAAGATGC    101
pKM177_VP6_R  ........................................      0


pKM177_VP6    TAGAGACAAAATTGTCGAAGGCACATTATACTCTAATGTG    480
pKM177_VP6_F  TAGAGACAAAATTGTCGAAGGCACATTATACTCTAATGTG    141
pKM177_VP6_R  ........................................      0


pKM177_VP6    AGTGATCTAATTCAACAATTTAACCAAATGATAATTACTA    520
pKM177_VP6_F  AGTGATCTAATTCAACAATTTAACCAAATGATAATTACTA    181
pKM177_VP6_R  ........................................      0


pKM177_VP6    TGAATGGAAATGAGTTCCAAACTGGAGGAATTGGTAATCT    560
pKM177_VP6_F  TGAATGGAAATGAGTTCCAAACTGGAGGAATTGGTAATCT    221
pKM177_VP6_R  ........................................      0


pKM177_VP6    ACCAATTAGAAATTGGAATTTTGATTTTGGATTACTTGGA    600
pKM177_VP6_F  ACCAATTAGAAATTGGAATTTTGATTTTGGATTACTTGGA    261
pKM177_VP6_R  ........................................      0


pKM177_VP6    ACAACTCTACTAAATTTAGACGCTAACTACGTCGAAACAG    640
pKM177_VP6_F  ACAACTCTACTAAATTTAGACGCTAACTACGTCGAAACAG    301
pKM177_VP6_R  ................................AACAG      5
                                              *****

pKM177_VP6    CCCGTAACACAATTGATTATTTTGTAGATTTTGTAGATAA    680
pKM177_VP6_F  CCCGTAACACAATTGATTATTTTGTAGATTTTGTAGATAA    341
pKM177_VP6_R  CCCGTAACACAATTGATTATTTTGTAGATTTTGTAGATAA     45
              ****************************************

pKM177_VP6    CGTATGTATGGATGAAATGGTTAGAGAATCACAAAGAAAT    720
pKM177_VP6_F  CGTATGTATGGATGAAATGGTTAGAGAATCACAAAGAAAT    381
pKM177_VP6_R  CGTATGTATGGATGAAATGGTTAGAGAATCACAAAGAAAT     85
              ****************************************

pKM177_VP6    GGAATTGCACCACAATCAGACTCACTTAGAAAATTGTCAG    760
pKM177_VP6_F  GGAATTGCACCACAATCAGACTCACTTAGAAAATTGTCAG    421
pKM177_VP6_R  GGAATTGCACCACAATCAGACTCACTTAGAAAATTGTCAG    125
              ****************************************
```

```
pKM177_VP6    GCATTAAGTTCAAAAGGATAAATTTTGATAATTCATCGGA    800
pKM177_VP6_F  GCATTAAGTTCAAAAGGATAAATTTTGATAATTCATCGGA    461
pKM177_VP6_R  GCATTAAGTTCAAAAGGATAAATTTTGATAATTCATCGGA    165
              ****************************************

pKM177_VP6    ATATATAGAGAACTGGAATCTACAAAACAGAAGACAACGA    840
pKM177_VP6_F  ATATATAGAGAACTGGAATCTACAAAACAGAAGACAACGA    501
pKM177_VP6_R  ATATATAGAGAACTGGAATCTACAAAACAGAAGACAACGA    205
              ****************************************

pKM177_VP6    ACAGGTTTTACATTTCATAAACCAAATATTTTCCCTTATT    880
pKM177_VP6_F  ACAGGTTTTACATTTCATAAACCAAATATTTTCCCTTATT    541
pKM177_VP6_R  ACAGGTTTTACATTTCATAAACCAAATATTTTCCCTTATT    245
              ****************************************

pKM177_VP6    CAGCGTCATTCACACTGAATAGATCACAACCAGCTCATGA    920
pKM177_VP6_F  CAGCGTCATTCACACTGAATAGATCACAACCAGCTCATGA    581
pKM177_VP6_R  CAGCGTCATTCACACTGAATAGATCACAACCAGCTCATGA    285
              ****************************************

pKM177_VP6    TAACTTGATGGGTACAATGTGGCTGAACGCAGGATCAGAA    960
pKM177_VP6_F  TAACTTGATGGGTACAATGTGGCTGAACGCAGGATCAGAA    621
pKM177_VP6_R  TAACTTGATGGGTACAATGTGGCTGAACGCAGGATCAGAA    325
              ****************************************

pKM177_VP6    ATTCAGGTCGCTGGATTCGACTATTCGTGTGCAATTAATG   1000
pKM177_VP6_F  ATTCAGGTCGCTGGATTCGACTATTCGTGTGCAATTAATG    661
pKM177_VP6_R  ATTCAGGTCGCTGGATTCGACTATTCGTGTGCAATTAATG    365
              ****************************************

pKM177_VP6    CGCCAGCTAATACACAACAATTTGAACATATTGTACAGCT   1040
pKM177_VP6_F  CGCCAGCTAATACACAACAATTTGAACATATTGTACAGCT    701
pKM177_VP6_R  CGCCAGCTAATACACAACAATTTGAACATATTGTACAGCT    405
              ****************************************

pKM177_VP6    CCGAAGAGTTTTAACTACAGCTACAATAACACTTTTACCG   1080
pKM177_VP6_F  CCGAAGAGTTTTAACTACAGCTACAATAACACTTTTACCG    741
pKM177_VP6_R  CCGAAGAGTTTTAACTACAGCTACAATAACACTTTTACCG    445
              ****************************************

pKM177_VP6    GATGCAGAAAGATTCAGTTTTCCAAGAGTGATTAATTCAG   1120
pKM177_VP6_F  GATGCAGAAAGATTCAGTTTTCCAAGAGTGATTAATTCAG    781
pKM177_VP6_R  GATGCAGAAAGATTCAGTTTTCCAAGAGTGATTAATTCAG    485
              ****************************************

pKM177_VP6    CTGACGGAGCAACTACATGGTATTTTAATCCAGTAATTCT   1160
pKM177_VP6_F  CTGACGGAGCAACTACATGGTATTTTAATCCAGTAATTCT    821
pKM177_VP6_R  CTGACGGAGCAACTACATGGTATTTTAATCCAGTAATTCT    525
              ****************************************

pKM177_VP6    TAGACCAAATAACGTTGAAGTGGAGTTTCTACTAAACGGG   1200
pKM177_VP6_F  TAGACCAAATAACGTTGAAGTGGAGTTTCTACTAAACGGG    861
pKM177_VP6_R  TAGACCAAATAACGTTGAAGTGGAGTTTCTACTAAACGGG    565
              ****************************************

pKM177_VP6    CAGATAATAAACACTTACCAGGCTAGATTTGGAACGATCG   1240
pKM177_VP6_F  CAGATAATAAACACTTACCAGGCTAGATTTGGAACGATCG    901
pKM177_VP6_R  CAGATAATAAACACTTACCAGGCTAGATTTGGAACGATCG    605
              ****************************************
```

```
pKM177_VP6    TAGCTAGAAATTTTGATACAATCAGATTGTCGTTTCAGTT   1280
pKM177_VP6_F  TAGCTAGAAATTTTGATACAATCAGATTGTCGTTTCAGTT    941
pKM177_VP6_R  TAGCTAGAAATTTTGATACAATCAGATTGTCGTTTCAGTT    645
              ***************************************

pKM177_VP6    GATGAGACCACCAAATATGACACCATCGGTAGCAGCATTA   1320
pKM177_VP6_F  GATGAGACCACCAAATATGACACCATCGGTAGCAGCATTA    981
pKM177_VP6_R  GATGAGACCACCAAATATGACACCATCGGTAGCAGCATTA    685
              ***************************************

pKM177_VP6    TTTCCAAATGCGCAACCATTTGAACATCATGCTACAGTAG   1360
pKM177_VP6_F  TTTCCAAATGCGCAACCATTTGAACATCATGCTACAGTAG   1021
pKM177_VP6_R  TTTCCAAATGCGCAACCATTTGAACATCATGCTACAGTAG    725
              ***************************************

pKM177_VP6    GACTTACATTGAAAATTGAATCTGCAGTTTGTGAATCTGT   1400
pKM177_VP6_F  GACTTACATTGAAAATTGAATCTGCAGTTTGTGAATCTGT   1061
pKM177_VP6_R  GACTTACATTGAAAATTGAATCTGCAGTTTGTGAATCTGT    765
              ***************************************

pKM177_VP6    ACTTGCTGACGCAAGCGAGACAATGCTAGCAAATGTGACA   1440
pKM177_VP6_F  ACTTGCTGACGCAAGCGAGACAATGCTAGCAAATGTGACA   1101
pKM177_VP6_R  ACTTGCTGACGCAAGCGAGACAATGCTAGCAAATGTGACA    805
              ***************************************

pKM177_VP6    TCTGTTAGACAAGAATACGCGATACCAGTTGGACCAGTCT   1480
pKM177_VP6_F  TCTGTTAGACAAGAATACGCGATACCAGTTGGACCAGTCT   1141
pKM177_VP6_R  TCTGTTAGACAAGAATACGCGATACCAGTTGGACCAGTCT    845
              ***************************************

pKM177_VP6    TTCCACCAGGTATGAATTGGACTGATTTGATCACTAACTA   1520
pKM177_VP6_F  TTCCACCAGGTATGAATTGGACTGATTTGATCACTAACTA   1181
pKM177_VP6_R  TTCCACCAGGTATGAATTGGACTGATTTGATCACTAACTA    885
              ***************************************

pKM177_VP6    TTCACCATCTAGAGAGGATAACTTGCAGCGTGTATTTACA   1560
pKM177_VP6_F  TTCACCATCTAGAGAGGATAACTTGCAGCGT.........   1212
pKM177_VP6_R  TTCACCATCTAGAGAGGATAACTTGCAGCGTGTATTTACA    925
              ******************************

pKM177_VP6    GTGGCTTCCATTAGAAGCATGCTTGTCAAATAAAGCGCTA   1600
pKM177_VP6_F  ........................................   1212
pKM177_VP6_R  GTGGCTTCCATTAGAAGCATGCTTGTCAAATAAAGCGCTA    965


pKM177_VP6    TTAATCCTAGGTGATCTGATCTGCTTACTTTACTTAACGA   1640
pKM177_VP6_F  ........................................   1212
pKM177_VP6_R  TTAATCCTAGGTGATCTGATCTGCTTACTTTACTTAACGA   1005


pKM177_VP6    CCAAAGAAAAACGACAAAAAAAAAAATATTACTACTATTAA   1680
pKM177_VP6_F  ........................................   1212
pKM177_VP6_R  CCAAAGAAAAACGACAAAAAAAAAAATATTACTACTATTAA   1045


pKM177_VP6    AATAAATTAGTATTTTTCTCTTCTTACGATATGATATGAT   1720
pKM177_VP6_F  ........................................   1212
pKM177_VP6_R  AATAAATTAGTATTTTTCTCTTCTTACGATATGATATGAT   1085
```

```
pKM177_VP6    GCTATGAAATCATCATCTTCTTAACTTTCTTGTCTCTTAC   1760
pKM177_VP6_F  ....................................     1212
pKM177_VP6_R  GCTATGAAATCATCATCTTCTTAACTTTCTTGTCTCTTAC   1125


pKM177_VP6    ACGTCACTTACTCTATATACCCGTTTATATAAGTGTACGT   1800
pKM177_VP6_F  ....................................     1212
pKM177_VP6_R  ACGTCACTTACTCTATATACCCGTTTATATAAGTGTACGT   1165


pKM177_VP6    ATTTTCTTTTTTTTAAAAAATTTCTATTCTATCCTTAGAA   1840
pKM177_VP6_F  ....................................     1212
pKM177_VP6_R  ATTTTCTTTTTTTTAAAAAATTTCTATTCTATCCTTAGAA   1205
```

**Figure 3.23: Nucleotide sequence alignments of the recombinant pKM177 vector containing the ORF of genome segment 6 (VP6) using the *in silico* clone of pKM177_VP6.** The *in silico* clone of pKM177_VP6 is indicated as pKM177_VP6, the forward sequence is indicated as pKM177_VP6_F nucleotide sequence and the reverse sequence is indicated as pKM177_VP6_R nucleotide sequence. The restriction enzymes used for the amplification of the ORF of genome segment 6 are indicated in grey. The start codon of the ORF of genome segment 6 is indicated in green and the stop codon is in red.

Both pKM177_VP6 colonies (colony 1 and colony 2) gave the expected results, however, only pKM177_VP6 colony 1 sequence results are shown. It is clear from Figure 3.23 that the open reading frame of genome segment 6 was cloned into pKM177. The primers designed for pKM177 made it possible to see the restriction enzymes used for cloning as well as the start and stop codon of the ORF of genome segment 6. All further experiments were performed using pKM177_VP6 colony 1.

### 3.3.3 Cloning of cassette, containing the ORF of genome segment 6 (VP6), into pKM173_VP2

In order to obtain a plasmid containing both the ORF of genome segment 2 and genome segment 6 the pKM177_VP6 and pKM173_VP2 plasmids had to be digested with the restriction enzyme I-SceI (see Figure 3.2B – illustration of cloning strategy). Restriction enzyme digestion was done, as described in section 3.2.3.3, for both pKM177_VP6 and pKM173_VP2. The digestions were done to clone the cassette containing: the ORF of genome segment 6, the *Yarrow lipolytica* translation elongation factor promoter (ylTEF) and the *Kluveromyces marxianus* inulinase target (KmlNut) into the expression vector pKM173_VP2. These digestions were analysed by loading a 5 µl sample of each reaction on a 1% agarose gel to evaluate if the digestions had worked (results not shown). The expected result for pKM173_VP2 was a linearized vector (10 103bp) and two bands (6493bp and 2263bp) for pKM177_VP6 since the plasmid contains two I-SceI restriction sites.

The expected results for the restriction digestion were obtained for both pKM177_VP6 and pKM173_VP2 (results not shown). The remaining sample of the restriction digestions were run on a 0.8% gel and purified by means of a gel extraction as described in section 2.2.3.7 (Chapter 2). This yielded about 35 µl gel purified product of both the cassette containing the ORF of genome segment 6 and pKM173_VP2 (96 ng/µl of the cassette containing the ORF of genome segment 6 and 67 ng/µl pKM173_VP2). Before the ligation reaction could be performed, pKM173_VP2 had to be dephosphorylated as described in section 3.2.3.4. This was done in order for pKM173_VP2 not to ligate to itself when performing the ligation reaction. The ligation background control contained no DNA insert. Transformation of competent JM109 cells, using the above mentioned ligation reaction mixture, was done as described in section 2.2.3.10 (Chapter 2). More than 100 colonies were obtained. One hundred and fifty colonies of possible pKM173_VP2/6 plasmids were chosen for screening by means of PCR colony screening, as described in section 3.2.3.5. Two colonies were pooled into one PCR reaction. A colony was considered positive if the PCR reaction gave one band at 1190bp (since the ORF of genome segment 6 PCR amplification primers were used) when analysed on a 1% agarose gel. The results of the PCR colony screening are not shown due to the bad quality of the agarose gels.

Colonies 1-2, 7-8, 15-16, 19-20, 35-36 and 40-41 gave the expected results of a band at 1190bp. Mini-preparations of plasmid DNA and restriction enzyme digestions were done on overnight LB medium cultures, of the colonies that gave the expected result with the PCR colony screening, to see if the cassette containing the ORF of genome segment 6 was cloned into pKM173_VP2. The plasmid map that would result, when the cassette containing the ORF of genome segment 6 was cloned into pKM173_VP2, is shown in Figure 3.24.

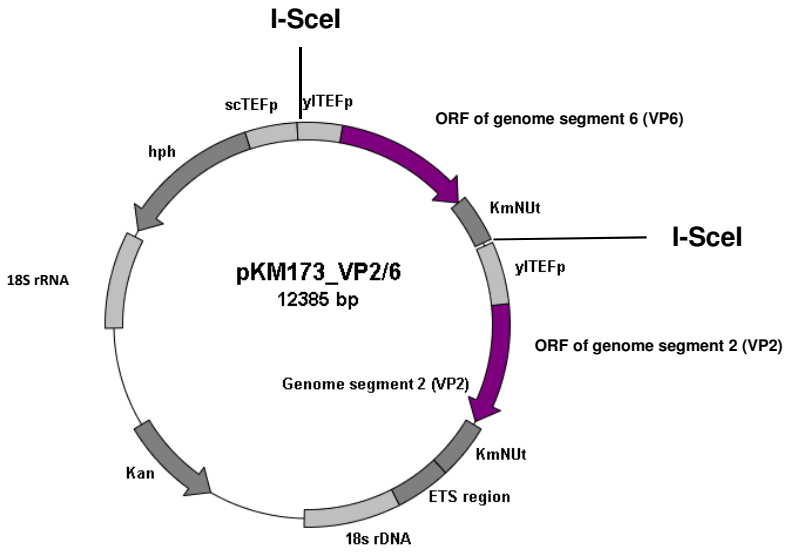


**Figure 3.24: Plasmid map that would result when the cassette containing the ORF of genome segment 6 is cloned into pKM173_VP2.** The I-SceI recognition sites, used to clone the cassette containing the ORF of genome segment 6 into pKM173_VP2, are indicated. The *Yarrowia lipolytica* translation elongation factor (ylTEF) promoters are indicated in light grey, and the *kanamycin* resistance gene in dark grey. The hygromycin B resistance gene (hph) is the antibiotic used during yeast transformation and the KmlNut is the *Kluyvermomyces marxianus* inulinase target sites.

Each sample was subjected to a restriction endonuclease digestion with the restriction enzyme XhoI. A 10 µl sample of each 25 µl double restriction digestion was analysed on a 1% agarose gel as shown in Figure 3.25. A colony was considered positive if two fragments were visible, namely 2101bp and 10 103bp, after the digestion reaction.

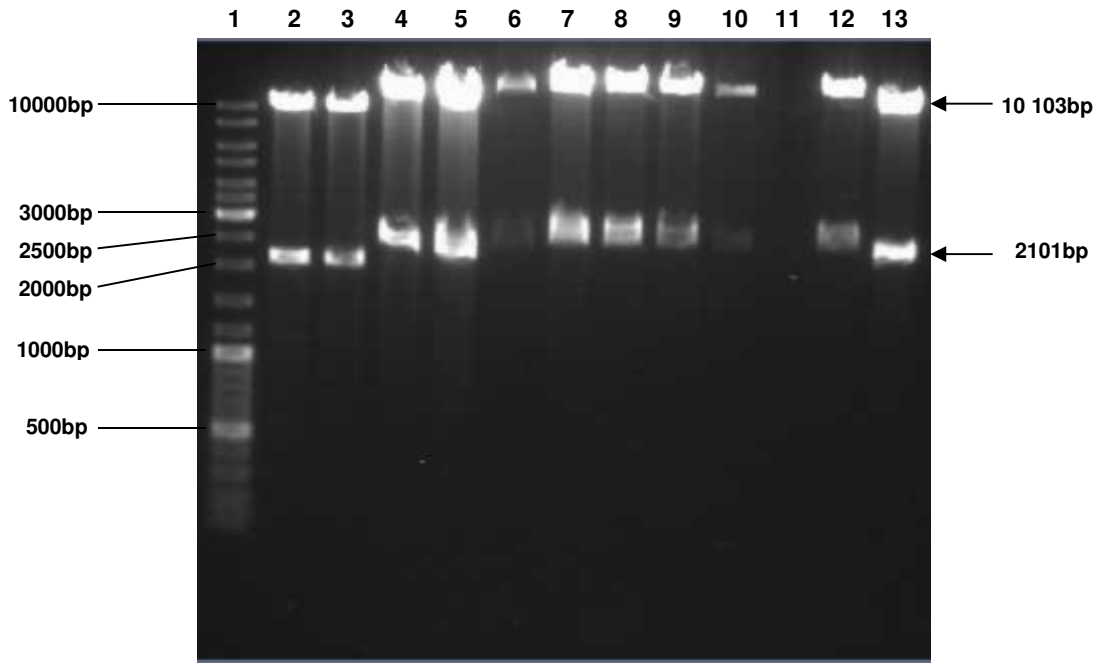


**Figure 3.25: Analysis by agarose gel electrophoresis of XhoI restriction enzyme digestion of possible pKM177_VP2/6 plasmids.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 25 µl pKM177_VP2/6 XhoI digestion reaction was loaded as follows; lanes: 2) Colony 1; 3) Colony 2; 4) Colony 7; 5) Colony 8; 6) Colony 15; 7) Colony 16; 8) Colony 19; 9) Colony 20; 10) Colony 35; 11) Colony 36; 12) Colony 40; 13) Colony 41.

Colonies 1, 2 and 41 (lanes 2, 3 and 13) gave the expected result namely two band one at 2101bp and the other at 10 103bp. Only two colonies (colony 1 and colony 41) were used for further analysis. A midi-preparation of plasmid DNA was done for both as described in section 3.2.3.1 (results not shown).

The midi-preparation showed that plasmid DNA was obtained for both pKM173_VP2/6 colonies (results not shown). However, to confirm that the cassette containing the ORF of genome segment 6 was cloned into pKM173_VP2, the two pKM173_VP2/6 colonies (colonies 1 and 41) were subjected to a restriction digestion with NotI. A 10 µl sample of each 25 µl double restriction digestion was analysed on a 1% agarose gel, as shown in Figure 3.26. Two fragments (2210bp and 9994bp) were expected after the double restriction digestion.

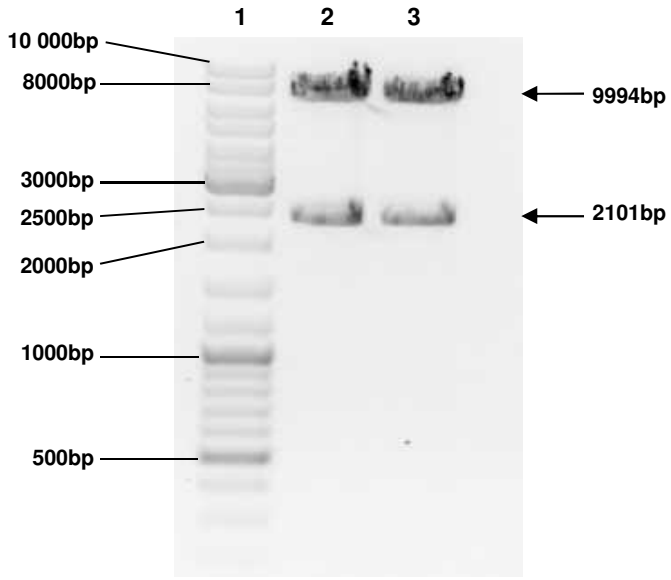


**Figure 3.26: Analysis by agarose gel electrophoresis of NotI restriction enzyme digest of pKM173_VP2/6 plasmids.** Lane: 1) 10 000bp O'Generuler DNA marker (Fermentas). A 10 µl volume of 25 µl NotI digestion reaction was loaded as follows; lanes: 2) pKM173_VP2/6 colony 1; 3) pKM173_VP2/6 colony 41.

The expected results were obtained for both colonies, as shown in Figure 3.26, namely two fragments, one at 2210bp and the other at 9994bp. Although these results were encouraging for both pKM173_VP2/6 constructs, sequencing was still necessary to ensure that the pKM173_VP2 plasmid contained the coding region of both genome segments. The midi-plasmid preparation of only pKM173_VP2/6 colony 1 was sent for sequencing, to ensure that the cassette containing the ORF of genome segment 6 was cloned into pKM173_VP2 and that no mutations occurred. Primers were designed, for pKM173_VP2/6 (Table 3.4 and indicated in Figure 3.27) sequencing in order to determine if the ORF of genome segment 6 was cloned into pKM173_VP2. Sequencing was done as described in section 3.2.3.7.

```
1       CTAGGGATAA CAGGGTAATG GTACCAGAGA CCGGGTTGGC GGCGTATTTG TGTCCCAAAA
        GATCCCTATT GTCCCATTAC CATGGTCTCT GGCCCAACCG CCGCATAAAC ACAGGGTTTT

61      AACAGCCCCA ATTGCCCCAA TTGACCCCAA ATTGACCCAG TAGCGGGCCC AACCCCGGCG
        TTGTCGGGGT TAACGGGGTT AACTGGGGTT TAACTGGGTC ATCGCCCGGG TTGGGGCCGC

121     AGAGCCCCCT TCACCCCACA TATCAAACCT CCCCCGGTTC CCACACTTGC CGTTAAGGGC
        TCTCGGGGGA AGTGGGGTGT ATAGTTTGGA GGGGGCCAAG GGTGTGAACG GCAATTCCCG

181     GTAGGGTACT GCAGTCTGGA ATCTACGCTT GTTCAGACTT TGTACTAGTT TCTTTGTCTG
        CATCCCATGA CGTCAGACCT TAGATGCGAA CAAGTCTGAA ACATGATCAA AGAAACAGAC

241     GCCATCCGGG TAACCCATGC CGGACGCAAA ATAGACTACT GAAAATTTTT TTGCTTTGTG
        CGGTAGGCCC ATTGGGTACG GCCTGCGTTT TATCTGATGA CTTTTAAAAA AACGAAACAC

301     GTTGGGACTT TAGCCAAGGG TATAAAAGAC CACCGTCCCC GAATTACCTT TCCTCTTCTT
        CAACCCTGAA ATCGGTTCCC ATATTTTCTG GTGGCAGGGG CTTAATGGAA AGGAGAAGAA

361     TTCTCTCTCT CCTTGTCAAC TCACACCCGA AATGCTCGAG ATGGATGTCC TGTACTCCTT
        AAGAGAGAGA GGAACAGTTG AGTGTGGGCT TTACGAGCTC TACCTACAGG ACATGAGGAA

421     ATCAAAAACT CTTAAAGATG CTAGAGACAA AATTGTCGAA GGCACATTAT ACTCTAATGT
        TAGTTTTTGA GAATTTCTAC GATCTCTGTT TTAACAGCTT CCGTGTAATA TGAGATTACA

481     GAGTGATCTA ATTCAACAAT TTAACCAAAT GATAATTACT ATGAATGGAA ATGAGTTCCA
        CTCACTAGAT TAAGTTGTTA AATTGGTTTA CTATTAATGA TACTTACCTT TACTCAAGGT

541     AACTGGAGGA ATTGGTAATC TACCAATTAG AAATTGGAAT TTTGATTTTG GATTACTTGG
        TTGACCTCCT TAACCATTAG ATGGTTAATC TTTAACCTTA AAACTAAAAC CTAATGAACC

601     AACAACTCTA CTAAATTTAG ACGCTAACTA CGTCGAAACA GCCCGTAACA CAATTGATTA
        TTGTTGAGAT GATTTAAATC TGCGATTGAT GCAGCTTTGT CGGGCATTGT GTTAACTAAT

661     TTTTGTAGAT TTTGTAGATA ACGTATGTAT GGATGAAATG GTTAGAGAAT CACAAAGAAA
        AAAACATCTA AAACATCTAT TGCATACATA CCTACTTTAC CAATCTCTTA GTGTTTCTTT

                    Reverse primer 1
721     TGGAATTGCA CCACAATCAG ACTCACTTAG AAAATTGTCA GGCATTAAGT TCAAAAGGAT
        ACCTTAACGT GGTGTTAGTC TGAGTGAATC TTTTAACAGT CCGTAATTCA AGTTTTCCTA

781     AAATTTTGAT AATTCATCGG AATATATAGA GAACTGGAAT CTACAAAACA GAAGACAACG
        TTTAAAACTA TTAAGTAGCC TTATATATCT CTTGACCTTA GATGTTTTGT CTTCTGTTGC

841     AACAGGTTTT ACATTTCATA AACCAAATAT TTTCCCTTAT TCAGCGTCAT TCACACTGAA
        TTGTCCAAAA TGTAAAGTAT TTGGTTTATA AAAGGGAATA AGTCGCAGTA AGTGTGACTT

901     TAGATCACAA CCAGCTCATG ATAACTTGAT GGGTACAATG TGGCTGAACG CAGGATCAGA
        ATCTAGTGTT GGTCGAGTAC TATTGAACTA CCCATGTTAC ACCGACTTGC GTCCTAGTCT

961     AATTCAGGTC GCTGGATTCG ACTATTCGTG TGCAATTAAT GCGCCAGCTA ATACACAACA
        TTAAGTCCAG CGACCTAAGC TGATAAGCAC ACGTTAATTA CGCGGTCGAT TATGTGTTGT

1021    ATTTGAACAT ATTGTACAGC TCCGAAGAGT TTTAACTACA GCTACAATAA CACTTTTACC
        TAAACTTGTA TAACATGTCG AGGCTTCTCA AAATTGATGT CGATGTTATT GTGAAAATGG

1081    GGATGCAGAA AGATTCAGTT TTCCAAGAGT GATTAATTCA GCTGACGGAG CAACTACATG
        CCTACGTCTT TCTAAGTCAA AAGGTTCTCA CTAATTAAGT CGACTGCCTC GTTGATGTAC
```

***Forward primer 1***

```
1141    GTATTTTAAT CCAGTAATTC TTAGACCAAA TAACGTTGAA GTGGAGTTTC TACTAAACGG
        CATAAAATTA GGTCATTAAG AATCTGGTTT ATTGCAACTT CACCTCAAAG ATGATTTGCC

1201    GCAGATAATA AACACTTACC AGGCTAGATT TGGAACGATC GTAGCTAGAA ATTTTGATAC
        CGTCTATTAT TTGTGAATGG TCCGATCTAA ACCTTGCTAG CATCGATCTT TAAAACTATG
```

***Reverse primer 2***

```
1261    AATCAGATTG TCGTTTCAGT TGATGAGACC ACCAAATATG ACACCATCGG TAGCAGCATT
        TTAGTCTAAC AGCAAAGTCA ACTACTCTGG TGGTTTATAC TGTGGTAGCC ATCGTCGTAA

1321    ATTTCCAAAT GCGCAACCAT TTGAACATCA TGCTACAGTA GGACTTACAT TGAAAATTGA
        TAAAGGTTTA CGCGTTGGTA AACTTGTAGT ACGATGTCAT CCTGAATGTA ACTTTTAACT

1381    ATCTGCAGTT TGTGAATCTG TACTTGCTGA CGCAAGCGAG ACAATGCTAG CAAATGTGAC
        TAGACGTCAA ACACTTAGAC ATGAACGACT GCGTTCGCTC TGTTACGATC GTTTACACTG

1441    ATCTGTTAGA CAAGAATACG CGATACCAGT TGGACCAGTC TTTCCACCAG GTATGAATTG
        TAGACAATCT GTTCTTATGC GCTATGGTCA ACCTGGTCAG AAAGGTGGTC CATACTTAAC
```

***Forward primer 2***

```
1501    GACTGATTTG ATCACTAACT ATTCACCATC TAGAGAGGAT AACTTGCAGC GTGTATTTAC
        CTGACTAAAC TAGTGATTGA TAAGTGGTAG ATCTCTCCTA TTGAACGTCG CACATAAATG

1561    AGTGGCTTCC ATTAGAAGCA TGCTTGTCAA ATAAAGCGCT ATTAATCCTA GGTGATCTGA
        TCACCGAAGG TAATCTTCGT ACGAACAGTT TATTTCGCGA TAATTAGGAT CCACTAGACT

1621    TCTGCTTACT TTACTTAACG ACCAAAGAAA AACGACAAAA AAAAAATATT ACTACTATTA
        AGACGAATGA AATGAATTGC TGGTTTCTTT TTGCTGTTTT TTTTTTATAA TGATGATAAT

1681    AAATAAATTA GTATTTTTCT CTTCTTACGA TATGATATGA TGCTATGAAA TCATCATCTT
        TTTATTTAAT CATAAAAAGA GAAGAATGCT ATACTATACT ACGATACTTT AGTAGTAGAA

1741    CTTAACTTTC TTGTCTCTTA CACGTCACTT ACTCTATATA CCCGTTTATA TAAGTGTACG
        GAATTGAAAG AACAGAGAAT GTGCAGTGAA TGAGATATAT GGGCAAATAT ATTCACATGC

1801    TATTTTCTTT TTTTTAAAAA ATTTCTATTC TATCCTTAGA AAAGTGCCCT TACATCAGTT
        ATAAAAGAAA AAAAATTTTT TAAAGATAAG ATAGGAATCT TTTCACGGGA ATGTAGTCAA

1861    CCAACGCACT CTAGCTGTTC TGGCACTGTA TCTTCATCAT GTGCCGGTCG TTTTCCACCC
        GGTTGCGTGA GATCGACAAG ACCGTGACAT AGAAGTAGTA CACGGCCAGC AAAAGGTGGG

1921    CAAAAATAAC TTTCTTCCCT TTTCCTTTCA ATTAATGGCC TGGAATTCCG AACCCATTTT
        GTTTTTATTG AAAGAAGGGA AAAGGAAAGT TAATTACCGG ACCTTAAGGC TTGGGTAAAA

1981    CGCATCTGAA ACTAATTCTC GAAACCTTTA ATATCAAACA ATTGAAAAGA TCATCATCAC
        GCGTAGACTT TGATTAAGAG CTTTGGAAAT TATAGTTTGT TAACTTTTCT AGTAGTAGTG

2041    TAGAAATGAG AAAAAGATCA ACAGCACTTA ATAACAGTAC GAAAGAAAGA TCGCTCGGAT
        ATCTTTACTC TTTTTCTAGT TGTCGTGAAT TATTGTCATG CTTTCTTTCT AGCGAGCCTA

2101    CCTAGGGATA ACAGGGTA GGTACCAGAG ACCGGGTTGG CGGCGTATTT GTGTCCCAAA
        GGATCCCTAT TGTCCCATTA CCATGGTCTC TGGCCCAACC GCCGCATAAA CACAGGGTTT

2161    AAACAGCCCC AATTGCCCCA ATTGACCCCA AATTGACCCA GTAGCGGGCC CAACCCCGGC
        TTTGTCGGGG TTAACGGGGT TAACTGGGGT TTAACTGGGT CATCGCCCGG GTTGGGGCCG

2221    GAGAGCCCCC TTCACCCCAC ATATCAAACC TCCCCCGGTT CCCACACTTG CCGTTAAGGG
        CTCTCGGGGG AAGTGGGGTG TATAGTTTGG AGGGGGCCAA GGGTGTGAAC GGCAATTCCC
```

```
2281    CGTAGGGTAC TGCAGTCTGG AATCTACGCT TGTTCAGACT TTGTACTAGT TTCTTTGTCT
        GCATCCCATG ACGTCAGACC TTAGATGCGA ACAAGTCTGA AACATGATCA AAGAAACAGA

2341    GGCCATCCGG GTAACCCATG CCGGACGCAA AATAGACTAC TGAAAATTTT TTTGCTTTGT
        CCGGTAGGCC CATTGGGTAC GGCCTGCGTT TTATCTGATG ACTTTTAAAA AAACGAAACA


2401    GGTTGGGACT TTAGCCAAGG GTATAAAAGA CCACCGTCCC CGAATTACCT TTCCTCTTCT
        CCAACCCTGA AATCGGTTCC CATATTTTCT GGTGGCAGGG GCTTAATGGA AAGGAGAAGA

2461    TTTCTCTCTC TCCTTGTCAA CTCACACCCG AAATGCTCGA GATCGCGTAC AGGAAACGTG
        AAAGAGAGAG AGGAACAGTT GAGTGTGGGC TTTACGAGCT CTACCGCATG TCCTTTGCAC

2521    GAGCGCGCCG TGAGGCGAAC TTAAATAATA ATGACCGAAT GCAGGAGAAA ATTGATGAAA
        CTCGCGCGGC ACTCCGCTTG AATTTATTAT TACTGGCTTA CGTCCTCTTT TAACTACTTT

2581    AACAAGATTC AAATAAAATA CAATTATCCG ATAAGGTACT TTCGAAGAAA GAAGAAATTG
        TTGTTCTAAG TTTATTTTAT GTTAATAGGC TATTCCATGA AAGCTTCTTT CTTCTTTAAC

2641    TAACGGATAG TCATGAGGAA GTTAAAGTTA CTGATGAGTT AAAAAAATCA ACGAAAGAAG
        ATTGCCTATC AGTACTCCTT CAATTTCAAT GACTACTCAA TTTTTTTAGT TGCTTTCTTC

2701    AATCAAAACA ATTGCTTGAA GTGTTGAAAA CAAAGGAAGA ACATCAGAAA GAAATACAGT
        TTAGTTTTGT TAACGAACTT CACAACTTTT GTTTCCTTCT TGTAGTCTTT CTTTATGTCA

2761    ATGAAATATT ACAGAAAACT ATACCAACAT TCGAACCTAA AGAGACGATA TTGAGAAAAT
        TACTTTATAA TGTCTTTTGA TATGGTTGTA AGCTTGGATT TCTCTGCTAT AACTCTTTTA

2821    TAGAGGATAT TCAACCAGAA CTAGCGAAAA AACAGACTAA GTTATTTAGA ATATTTGAAC
        ATCTCCTATA AGTTGGTCTT GATCGCTTTT TTGTCTGATT CAATAAATCT TATAAACTTG

2881    CGAAACAATT ACCGATTTAT AGAGCAAATG GAGAGAGAGA ATTGCGTAAT AGATGGTATT
        GCTTTGTTAA TGGCTAAATA TCTCGTTTAC CTCTCTCTCT TAACGCATTA TCTACCATAA

2941    GGAAATTAAA AAAAGATACA CTACCAGACG GAGACTATGA TGTGAGAGAG TATTTTCTGA
        CCTTTAATTT TTTTCTATGT GATGGTCTGC CTCTGATACT ACACTCTCTC ATAAAAGACT

3001    ATTTGTATGA TCAAGTGCTT ACTGAAATGC CAGACTACTT ATTATTGAAA GATATGGCAG
        TAAACATACT AGTTCACGAA TGACTTTACG GTCTGATGAA TAATAACTTT CTATACCGTC

3061    TAGAAAATAA GAACTCTAGG GATGCAGGTA AAGTTGTTGA CTCAGAAACG GCTAGTATAT
        ATCTTTTATT CTTGAGATCC CTACGTCCAT TTCAACAACT GAGTCTTTGC CGATCATATA

3121    GCGATGCCAT ATTTCAAGAT GAAGAAACGG AAGGTGCCGT TAGAAGATTC ATTGCAGAAA
        CGCTACGGTA TAAAGTTCTA CTTCTTTGCC TTCCACGGCA ATCTTCTAAG TAACGTCTTT

3181    TGAGACAACG TGTGCAAGCT GATAGAAATG TTGTCAATTA TCCATCAATA TTACATCCAA
        ACTCTGTTGC ACACGTTCGA CTATCTTTAC AACAGTTAAT AGGTAGTTAT AATGTAGGTT

3241    TAGATTATGC ATTTAATGAA TACTTTTTAC AACATCAATT GGTTGAACCA TTGAATAATG
        ATCTAATACG TAAATTACTT ATGAAAAATG TTGTAGTTAA CCAACTTGGT AACTTATTAC

3301    ATATAATATT TAATTATATA CCAGAAAGGA TAAGAAATGA TGTTAATTAT ATTCTCAATA
        TATATTATAA ATTAATATAT GGTCTTTCCT ATTCTTTACT ACAATTAATA TAAGAGTTAT

3361    TGGACAGAAA TTTACCATCA ACTGCCAGAT ATATAAGACC TAATTTACTT CAAGATAGAT
        ACCTGTCTTT AAATGGTAGT TGACGGTCTA TATATTCTGG ATTAAATGAA GTTCTATCTA

3421    TAAATTTGCA CGATAATTTT GAATCACTAT GGGATACAAT AACTACATCA AATTATATTT
        ATTTAAACGT GCTATTAAAA CTTAGTGATA CCCTATGTTA TTGATGTAGT TTAATATAAA
```

```
3481    TGGCGAGATC GGTAGTACCA GATTTAAAGG AATTAGTGTC AACGGAAGCA CAAATTCAGA
        ACCGCTCTAG CCATCATGGT CTAAATTTCC TTAATCACAG TTGCCTTCGT GTTTAAGTCT

3541    AAATGTCACA AGATTTGCAA TTAGAAGCAT TAACAATTCA GTCAGAAACA CAATTTCTAA
        TTTACAGTGT TCTAAACGTT AATCTTCGTA ATTGTTAAGT CAGTCTTTGT GTTAAAGATT

3601    CAGGTATAAA TTCACAAGCA GCTAACGATT GTTTTAAAAC CTTAATTGCA GCAATGTTAA
        GTCCATATTT AAGTGTTCGT CGATTGCTAA CAAAATTTTG GAATTAACGT CGTTACAATT

3661    GTCAACGTAC TATGTCATTA GATTTTGTAA CTACTAATTA TATGTCATTG ATTTCAGGTA
        CAGTTGCATG ATACAGTAAT CTAAAACATT GATGATTAAT ATACAGTAAC TAAAGTCCAT

3721    TGTGGCTATT GACGGTTGTG CCAAATGATA TGTTTATAAG GGAATCGTTA GTCGCGTGTC
        ACACCGATAA CTGCCAACAC GGTTTACTAT ACAAATATTC CCTTAGCAAT CAGCGCACAG

3781    AACTAGCTAT AGTAAATACA ATAATCTATC CAGCATTTGG AATGCAACGA ATGCATTATA
        TTGATCGATA TCATTTATGT TATTAGATAG GTCGTAAACC TTACGTTGCT TACGTAATAT

3841    GAAACGGGGA TCCACAAACA CCGTTTCAGA TAGCAGAACA GCAGATTCAA AATTTCCAAG
        CTTTGCCCCT AGGTGTTTGT GGCAAAGTCT ATCGTCTTGT CGTCTAAGTT TTAAAGGTTC

3901    TCGCAAATTG GTTACATTTT GTTAATAATA ATCAATTTAG ACAGGCAGTT ATTGATGGTG
        AGCGTTTAAC CAATGTAAAA CAATTATTAT TAGTTAAATC TGTCCGTCAA TAACTACCAC

3961    TATTGAATCA GGTACTGAAT GACAATATTA GAAATGGTCA TGTTATTAAC CAACTGATGG
        ATAACTTAGT CCATGACTTA CTGTTATAAT CTTTACCAGT ACAATAATTG GTTGACTACC

4021    AGGCTCTAAT GCAGCTGTCG CGACAACAAT TTCCAACCAT GCCAATTGAT TATAAGAGAT
        TCCGAGATTA CGTCGACAGC GCTGTTGTTA AAGGTTGGTA CGGTTAACTA ATATTCTCTA

4081    CAATTCAACG TGGAATATTA CTGTTATCTA ACAGACTTGG TCAGTTAGTT GATTTAACTA
        GTTAAGTTGC ACCTTATAAT GACAATAGAT TGTCTGAACC AGTCAATCAA CTAAATTGAT

4141    GATTATTAGC TTACAATTAT GAGACATTAA TGGCATGCAT TACAATGAAC ATGCAACATG
        CTAATAATCG AATGTTAATA CTCTGTAATT ACCGTACGTA ATGTTACTTG TACGTTGTAC

4201    TTCAAACCTT AACAACAGAA AAATTACAAT TAACGTCAGT TACATCATTA TGTATGCTTA
        AAGTTTGGAA TTGTTGTCTT TTTAATGTTA ATTGCAGTCA ATGTAGTAAT ACATACGAAT

4261    TTGGAAATGC GACTGTATAC CAAGTCCACA AACATTATTT CATTATTATA ACGTTAACGT
        AACCTTTACG CTGACATATG GTTCAGGTGT TTGTAATAAA GTAATAATAT TGCAATTGCA

4321    TAATTTTCAT TCAAATTACA ATGAGAGAAT TAATGATGCA GTAGCTATAA TAACTGCTGC
        ATTAAAAGTA AGTTTAATGT TACTCTCTTA ATTACTACGT CATCGATATT ATTGACGACG

4381    TAACAGACTG AATCTATATC AGAAAAAAAT GAAGGCTATT GTTGAGGATT TCTTAAAAAG
        ATTGTCTGAC TTAGATATAG TCTTTTTTTA CTTCCGATAA CAACTCCTAA AGAATTTTTC

4441    ATTATACATT TTTGATGTAT CTAGAGTTCC GGACGACCAA ATGTATAGAT TAAGGGATAG
        TAATATGTAA AAACTACATA GATCTCAAGG CCTGCTGGTT TACATATCTA ATTCCCTATC

4501    ATTACGCTTA TTGCCAGTAG AAATCAGAAG ATTAGATATC TTCAATCTAA TACTAATGAA
        TAATGCGAAT AACGGTCATC TTTAGTCTTC TAATCTATAG AAGTTAGATT ATGATTACTT

4561    CATGGATCAA ATTGAACGTG CCTCAGATAA AATTGCTCAA GGTGTAATCA TTGCTTATCG
        GTACCTAGTT TAACTTGCAC GGAGTCTATT TTAACGAGTT CCACATTAGT AACGAATAGC

4621    TGACATGCAT CTGAAAGAGA TGAGATGTAC GGATATGTAA ATATAGCTAG AAATTTAGAG
        ACTGTACGTA GACTTTCTCT ACTCTACATG CCTATACATT TATATCGATC TTTAAATCTC
```


**Figure 3.27: Primer design for the sequencing of pKM173_VP2/6 construct to determine if the cassette containing the ORF of genome segment 6 was cloned into pKM173_VP2/6.** The forward and reverse primers designed for sequencing are highlighted in yellow. All the restriction enzymes used for cloning of the ORF of genome segment 2 and 6 into their respective plasmid are highlighted in grey. The start codons of the ORF of the genome segments are highlighted with green and the stop codons in red. The I-SceI restriction enzyme sequences are highlighted in blue. The sequence of the ORF of genome segment 6 is indicated in purple and the sequence of the ORF of genome segment 2 in blue. The sequence of the pKM173 plasmid is indicated in black.

Primers designed for the sequenicng of pKM173_VP2/6 (Table 3.4, Figure 3.27) were used, as described in section 3.2.3.7. The sequence alignment of pKM173_VP2/6 colony 1 is indicated in Figure 2.28A and B.

```
pKM173_VP2/6     CTAGGGATAACAGGGTAATGGTACCAGAGACCGGGTTGGC     40
pKM173_VP2/6_F1  .....................................      0
pKM173_VP2/6_R1  .....................................      0


pKM173_VP2/6     GGCGTATTTGTGTCCCAAAAAACAGCCCCAATTGCCCCAA     80
pKM173_VP2/6_F1  .....................................      0
pKM173_VP2/6_R1  .....................................      0


pKM173_VP2/6     TTGACCCCAAATTGACCCAGTAGCGGGCCCAACCCCGGCG    120
pKM173_VP2/6_F1  .....................................      0
pKM173_VP2/6_R1  .....................................      0


pKM173_VP2/6     AGAGCCCCCTTCACCCCACATATCAAACCTCCCCCGGTTC    160
pKM173_VP2/6_F1  .....................................      0
pKM173_VP2/6_R1  ..........TCACCCCACATATCAAACCTCCCCCGGTTC     30


pKM173_VP2/6     CCACACTTGCCGTTAAGGGCGTAGGGTACTGCAGTCTGGA    200
pKM173_VP2/6_F1  .....................................      0
pKM173_VP2/6_R1  CCACACTTGCCGTTAAGGGCGTAGGGTACTGCAGTCTGGA     70


pKM173_VP2/6     ATCTACGCTTGTTCAGACTTTGTACTAGTTTCTTTGTCTG    240
pKM173_VP2/6_F1  .....................................      0
pKM173_VP2/6_R1  ATCTACGCTTGTTCAGACTTTGTACTAGTTTCTTTGTCTG    110


pKM173_VP2/6     GCCATCCGGGTAACCCATGCCGGACGCAAAATAGACTACT    280
pKM173_VP2/6_F1  .....................................      0
pKM173_VP2/6_R1  GCCATCCGGGTAACCCATGCCGGACGCAAAATAGACTACT    150


pKM173_VP2/6     GAAAATTTTTTTGCTTTGTGGTTGGGACTTTAGCCAAGGG    320
pKM173_VP2/6_F1  .....................................      0
pKM173_VP2/6_R1  GAAAATTTTTTTGCTTTGTGGTTGGGACTTTAGCCAAGGG    190


pKM173_VP2/6     TATAAAAGACCACCGTCCCCGAATTACCTTTCCTCTTCTT    360
pKM173_VP2/6_F1  .....................................      0
pKM173_VP2/6_R1  TATAAAAGACCACCGTCCCCGAATTACCTTTCCTCTTCTT    230


pKM173_VP2/6     TTCTCTCTCTCCTTGTCAACTCACACCCGAAATGCTCGAG    400
pKM173_VP2/6_F1  .....................................      0
pKM173_VP2/6_R1  TTCTCTCTCTCCTTGTCAACTCACACCCGAAATGCTCGAG    270


pKM173_VP2/6     ATGGATGTCCTGTACTCCTTATCAAAAACTCTTAAAGATG    440
pKM173_VP2/6_F1  .....................................      0
pKM173_VP2/6_R1  ATGGATGTCCTGTACTCCTTATCAAAAACTCTTAAAGATG    310


pKM173_VP2/6     CTAGAGACAAAATTGTCGAAGGCACATTATACTCTAATGT    480
pKM173_VP2/6_F1  .....................................      0
pKM173_VP2/6_R1  CTAGAGACAAAATTGTCGAAGGCACATTATACTCTAATGT    350
```

```
pKM173_VP2/6     GAGTGATCTAATTCAACAATTTAACCAAATGATAATTACT    520
pKM173_VP2/6_F1  ....................................      0
pKM173_VP2/6_R1  GAGTGATCTAATTCAACAATTTAACCAAATGATAATTACT    390


pKM173_VP2/6     ATGAATGGAAATGAGTTCCAAACTGGAGGAATTGGTAATC    560
pKM173_VP2/6_F1  ....................................      0
pKM173_VP2/6_R1  ATGAATGGAAATGAGTTCCAAACTGGAGGAATTGGTAATC    430


pKM173_VP2/6     TACCAATTAGAAATTGGAATTTTGATTTTGGATTACTTGG    600
pKM173_VP2/6_F1  ....................................      0
pKM173_VP2/6_R1  TACCAATTAGAAATTGGAATTTTGATTTTGGATTACTTGG    470


pKM173_VP2/6     AACAACTCTACTAAATTTAGACGCTAACTACGTCGAAACA    640
pKM173_VP2/6_F1  ....................................      0
pKM173_VP2/6_R1  AACAACTCTACTAAATTTAGACGCTAACTACGTCGAAACA    510


pKM173_VP2/6     GCCCGTAACACAATTGATTATTTTGTAGATTTTGTAGATA    680
pKM173_VP2/6_F1  ....................................      0
pKM173_VP2/6_R1  GCCCGTAACACAATTGATTATTTTGTAGATTTTGTAGATA    550


pKM173_VP2/6     ACGTATGTATGGATGAAATGGTTAGAGAATCACAAAGAAA    720
pKM173_VP2/6_F1  ....................................      0
pKM173_VP2/6_R1  ACGTATGTATGGATGAAATGGTTAGAGAATCACAAAGAA.    589


pKM173_VP2/6     TGGAATTGCACCACAATCAGACTCACTTAGAAAATTGTCA    760
pKM173_VP2/6_F1  ....................................      0
pKM173_VP2/6_R1  ....................................    589


pKM173_VP2/6     GGCATTAAGTTCAAAAGGATAAATTTTGATAATTCATCGG    800
pKM173_VP2/6_F1  ....................................      0
pKM173_VP2/6_R1  ....................................    589


pKM173_VP2/6     AATATATAGAGAACTGGAATCTACAAAACAGAAGACAACG    840
pKM173_VP2/6_F1  ....................................      0
pKM173_VP2/6_R1  ....................................    589


pKM173_VP2/6     AACAGGTTTTACATTTCATAAACCAAATATTTTCCCTTAT    880
pKM173_VP2/6_F1  ....................................      0
pKM173_VP2/6_R1  ....................................    589


pKM173_VP2/6     TCAGCGTCATTCACACTGAATAGATCACAACCAGCTCATG    920
pKM173_VP2/6_F1  ....................................      0
pKM173_VP2/6_R1  ....................................    589


pKM173_VP2/6     ATAACTTGATGGGTACAATGTGGCTGAACGCAGGATCAGA    960
pKM173_VP2/6_F1  ....................................      0
pKM173_VP2/6_R1  ....................................    589
```

```
pKM173_VP2/6     AATTCAGGTCGCTGGATTCGACTATTCGTGTGCAATTAAT   1000
pKM173_VP2/6_F1  ........................................     0
pKM173_VP2/6_R1  ........................................   589

pKM173_VP2/6     GCGCCAGCTAATACACAACAATTTGAACATATTGTACAGC   1040
pKM173_VP2/6_F1  ........................................     0
pKM173_VP2/6_R1  ........................................   589

pKM173_VP2/6     TCCGAAGAGTTTTAACTACAGCTACAATAACACTTTTACC   1080
pKM173_VP2/6_F1  ........................................     0
pKM173_VP2/6_R1  ........................................   589

pKM173_VP2/6     GGATGCAGAAAGATTCAGTTTTCCAAGAGTGATTAATTCA   1120
pKM173_VP2/6_F1  ........................................     0
pKM173_VP2/6_R1  ........................................   589

pKM173_VP2/6     GCTGACGGAGCAACTACATGGTATTTTAATCCAGTAATTC   1160
pKM173_VP2/6_F1  ........................................     0
pKM173_VP2/6_R1  ........................................   589

pKM173_VP2/6     TTAGACCAAATAACGTTGAAGTGGAGTTTCTACTAAACGG   1200
pKM173_VP2/6_F1  ........................................     0
pKM173_VP2/6_R1  ........................................   589

pKM173_VP2/6     GCAGATAATAAACACTTACCAGGCTAGATTTGGAACGATC   1240
pKM173_VP2/6_F1  ......AATAAACACTTACCAGGCTAGATTTGGAACGATC    34
pKM173_VP2/6_R1  ........................................   589

pKM173_VP2/6     GTAGCTAGAAATTTTGATACAATCAGATTGTCGTTTCAGT   1280
pKM173_VP2/6_F1  GTAGCTAGAAATTTTGATACAATCAGATTGTCGTTTCAGT    74
pKM173_VP2/6_R1  ........................................   589

pKM173_VP2/6     TGATGAGACCACCAAATATGACACCATCGGTAGCAGCATT   1320
pKM173_VP2/6_F1  TGATGAGACCACCAAATATGACACCATCGGTAGCAGCATT   114
pKM173_VP2/6_R1  ........................................   589

pKM173_VP2/6     ATTTCCAAATGCGCAACCATTTGAACATCATGCTACAGTA   1360
pKM173_VP2/6_F1  ATTTCCAAATGCGCAACCATTTGAACATCATGCTACAGTA   154
pKM173_VP2/6_R1  ........................................   589

pKM173_VP2/6     GGACTTACATTGAAAATTGAATCTGCAGTTTGTGAATCTG   1400
pKM173_VP2/6_F1  GGACTTACATTGAAAATTGAATCTGCAGTTTGTGAATCTG   194
pKM173_VP2/6_R1  ........................................   589

pKM173_VP2/6     TACTTGCTGACGCAAGCGAGACAATGCTAGCAAATGTGAC   1440
pKM173_VP2/6_F1  TACTTGCTGACGCAAGCGAGACAATGCTAGCAAATGTGAC   234
pKM173_VP2/6_R1  ........................................   589
```

```
pKM173_VP2/6     ATCTGTTAGACAAGAATACGCGATACCAGTTGGACCAGTC   1480
pKM173_VP2/6_F1  ATCTGTTAGACAAGAATACGCGATACCAGTTGGACCAGTC    274
pKM173_VP2/6_R1  .....................................     589

pKM173_VP2/6     TTTCCACCAGGTATGAATTGGACTGATTTGATCACTAACT   1520
pKM173_VP2/6_F1  TTTCCACCAGGTATGAATTGGACTGATTTGATCACTAACT    314
pKM173_VP2/6_R1  .....................................     589

pKM173_VP2/6     ATTCACCATCTAGAGAGGATAACTTGCAGCGTGTATTTAC   1560
pKM173_VP2/6_F1  ATTCACCATCTAGAGAGGATAACTTGCAGCGTGTATTTAC    354
pKM173_VP2/6_R1  .....................................     589

pKM173_VP2/6     AGTGGCTTCCATTAGAAGCATGCTTGTCAAATAAAGCGCT   1600
pKM173_VP2/6_F1  AGTGGCTTCCATTAGAAGCATGCTTGTCAAATAAAGCGCT    394
pKM173_VP2/6_R1  .....................................     589

pKM173_VP2/6     ATTAATCCTAGGTGATCTGATCTGCTTACTTTACTTAACG   1640
pKM173_VP2/6_F1  ATTAATCCTAGGTGATCTGATCTGCTTACTTTACTTAACG    434
pKM173_VP2/6_R1  .....................................     589

pKM173_VP2/6     ACCAAAGAAAAACGACAAAAAAAAAAATATTACTACTATTA   1680
pKM173_VP2/6_F1  ACCAAAGAAAAACGACAAAAAAAAAAATATTACTACTATTA    474
pKM173_VP2/6_R1  .....................................     589

pKM173_VP2/6     AAATAAATTAGTATTTTTCTCTTCTTACGATATGATATGA   1720
pKM173_VP2/6_F1  AAATAAATTAGTATTTTTCTCTTCTTACGATATGATATGA    514
pKM173_VP2/6_R1  .....................................     589

pKM173_VP2/6     TGCTATGAAATCATCATCTTCTTAACTTTCTTGTCTCTTA   1760
pKM173_VP2/6_F1  TGCTATGAAATCATCATCTTCTTAACTTTCTTGTCTCTTA    554
pKM173_VP2/6_R1  .....................................     589

pKM173_VP2/6     CACGTCACTTACTCTATATACCCGTTTATATAAGTGTACG   1800
pKM173_VP2/6_F1  CACGTCACTTACTCTATATACCCGTTTATATAAGTGTACG    594
pKM173_VP2/6_R1  .....................................     589

pKM173_VP2/6     TATTTTCTTTTTTTTAAAAAATTTCTATTCTATCCTTAGA   1840
pKM173_VP2/6_F1  TATTTTCTTTTTTTTAAAAAATTTCTATTCTATCCTTAGA    634
pKM173_VP2/6_R1  .....................................     589

pKM173_VP2/6     AAAGTGCCCTTACATCAGTTCCAACGCACTCTAGCTGTTC   1880
pKM173_VP2/6_F1  AAAGTGCCCTTACATCAGTTCCAACGCACTCTAGCTGTTC    674
pKM173_VP2/6_R1  .....................................     589

pKM173_VP2/6     TGGCACTGTATCTTCATCATGTGCCGGTCGTTTTCCACCC   1920
pKM173_VP2/6_F1  TGGCACTGTATCTTCATCATGTGCCGGTCGTTTTCCACCC    714
pKM173_VP2/6_R1  .....................................     589
```

```
pKM173_VP2/6     CAAAAATAACTTTCTTCCCTTTTCCTTTCAATTAATGGCC   1960
pKM173_VP2/6_F1  CAAAAATAACTTTCTTCCCTTTTCCTTTCAATTAATGGCC    754
pKM173_VP2/6_R1  ......................................    589

pKM173_VP2/6     TGGAATTCCGAACCCATTTTCGCATCTGAAACTAATTCTC   2000
pKM173_VP2/6_F1  TGGAATTCCGAACCCATTTTCGCATCTGAAACTAATTCTC    794
pKM173_VP2/6_R1  ......................................    589

pKM173_VP2/6     GAAACCTTTAATATCAAACAATTGAAAAGATCATCATCAC   2040
pKM173_VP2/6_F1  GAAACCTTTAATATCAAACAATTGAAAAGATCATCATCAC    834
pKM173_VP2/6_R1  ......................................    589

pKM173_VP2/6     TAGAAATGAGAAAAAGATCAACAGCACTTAATAACAGTAC   2080
pKM173_VP2/6_F1  TAGAAATGAGAAAAAGATCAACAGCACTTAATAACAGTAC    874
pKM173_VP2/6_R1  ......................................    589

pKM173_VP2/6     GAAAGAAAGATCGCTCGGATCCTAGGGATAACAGGGTAAT   2120
pKM173_VP2/6_F1  GAAAGAAAGATCGCTCGGATCCTAGGGATAACAGGGTAAT    914
pKM173_VP2/6_R1  ......................................    589

pKM173_VP2/6     GGTACCAGAGACCGGGTTGGCGGCGTATTTGTGTCCCAAA   2160
pKM173_VP2/6_F1  ......................................    914
pKM173_VP2/6_R1  ......................................    589
```

**Figure 3.28A: Nucleotide sequence alignments of the recombinant pKM173_VP2 vector containing the ORF of genome segment 6 (VP6) using the *in silico* clone of pKM173_VP2/6.** The *in silico* clone of pKM173_VP2/6 is indicated as pKM173_VP2/6, the forward (forward primer 1) sequence is indicated as pKM173_VP2/6_F1 nucleotide sequence and the reverse (reverse primer 1) sequence is indicated as pKM173_VP2/6_R1 nucleotide sequence. The restriction enzymes used for the amplification of the ORF of genome segment 6 are indicated in grey. The start codon of the ORF of genome segment 6 is indicated in green and the stop codon is in red. The part of the ORF of genome segment 6 that could not be sequenced with the forward 1 primer and reverse 1 primer is indicated in dark grey.

```
pKM173_VP2/6     CTAGAGACAAAATTGTCGAAGGCACATTATACTCTAATGT    480
pKM173_VP2/6_F2  ....................................        0
pKM173_VP2/6_R2  ............................CTCTAATGT        9

pKM173_VP2/6     GAGTGATCTAATTCAACAATTTAACCAAATGATAATTACT    520
pKM173_VP2/6_F2  ....................................        0
pKM173_VP2/6_R2  GAGTGATCTAATTCAACAATTTAACCAAATGATAATTACT     49

pKM173_VP2/6     ATGAATGGAAATGAGTTCCAAACTGGAGGAATTGGTAATC    560
pKM173_VP2/6_F2  ....................................        0
pKM173_VP2/6_R2  ATGAATGGAAATGAGTTCCAAACTGGAGGAATTGGTAATC     89

pKM173_VP2/6     TACCAATTAGAAATTGGAATTTTGATTTTGGATTACTTGG    600
pKM173_VP2/6_F2  ....................................        0
pKM173_VP2/6_R2  TACCAATTAGAAATTGGAATTTTGATTTTGGATTACTTGG    129

pKM173_VP2/6     AACAACTCTACTAAATTTAGACGCTAACTACGTCGAAACA    640
pKM173_VP2/6_F2  ....................................        0
pKM173_VP2/6_R2  AACAACTCTACTAAATTTAGACGCTAACTACGTCGAAACA    169

pKM173_VP2/6     GCCCGTAACACAATTGATTATTTTGTAGATTTTGTAGATA    680
pKM173_VP2/6_F2  ....................................        0
pKM173_VP2/6_R2  GCCCGTAACACAATTGATTATTTTGTAGATTTTGTAGATA    209

pKM173_VP2/6     ACGTATGTATGGATGAAATGGTTAGAGAATCACAAAGAAA    720
pKM173_VP2/6_F2  ....................................        0
pKM173_VP2/6_R2  ACGTATGTATGGATGAAATGGTTAGAGAATCACAAAGAAA    249

pKM173_VP2/6     TGGAATTGCACCACAATCAGACTCACTTAGAAAATTGTCA    760
pKM173_VP2/6_F2  ....................................        0
pKM173_VP2/6_R2  TGGAATTGCACCACAATCAGACTCACTTAGAAAATTGTCA    289

pKM173_VP2/6     GGCATTAAGTTCAAAAGGATAAATTTTGATAATTCATCGG    800
pKM173_VP2/6_F2  ....................................        0
pKM173_VP2/6_R2  GGCATTAAGTTCAAAAGGATAAATTTTGATAATTCATCGG    329

pKM173_VP2/6     AATATATAGAGAACTGGAATCTACAAAACAGAAGACAACG    840
pKM173_VP2/6_F2  ....................................        0
pKM173_VP2/6_R2  AATATATAGAGAACTGGAATCTACAAAACAGAAGACAACG    369

pKM173_VP2/6     AACAGGTTTTACATTTCATAAACCAAATATTTTCCCTTAT    880
pKM173_VP2/6_F2  ....................................        0
pKM173_VP2/6_R2  AACAGGTTTTACATTTCATAAACCAAATATTTTCCCTTAT    409

pKM173_VP2/6     TCAGCGTCATTCACACTGAATAGATCACAACCAGCTCATG    920
pKM173_VP2/6_F2  ....................................        0
pKM173_VP2/6_R2  TCAGCGTCATTCACACTGAATAGATCACAACCAGCTCATG    449
```

```
pKM173_VP2/6     ATAACTTGATGGGTACAATGTGGCTGAACGCAGGATCAGA    960
pKM173_VP2/6_F2  ........................................      0
pKM173_VP2/6_R2  ATAACTTGATGGGTACAATGTGGCTGAACGCAGGATCAGA    489

pKM173_VP2/6     AATTCAGGTCGCTGGATTCGACTATTCGTGTGCAATTAAT   1000
pKM173_VP2/6_F2  ........................................      0
pKM173_VP2/6_R2  AATTCAGGTCGCTGGATTCGACTATTCGTGTGCAATTAAT    529

pKM173_VP2/6     GCGCCAGCTAATACACAACAATTTGAACATATTGTACAGC   1040
pKM173_VP2/6_F2  ........................................      0
pKM173_VP2/6_R2  GCGCCAGCTAATACACAACAATTTGAACATATTGTACAGC    569

pKM173_VP2/6     TCCGAAGAGTTTTAACTACAGCTACAATAACACTTTTACC   1080
pKM173_VP2/6_F2  ........................................      0
pKM173_VP2/6_R2  TCCGAAGAGTTTTAACTACAGCTACAATAACACTTTTACC    609

pKM173_VP2/6     GGATGCAGAAAGATTCAGTTTTCCAAGAGTGATTAATTCA   1120
pKM173_VP2/6_F2  ........................................      0
pKM173_VP2/6_R2  GGATGCAGAAAGATTCAGTTTTCCAAGAGTGATTAATTCA    649

pKM173_VP2/6     GCTGACGGAGCAACTACATGGTATTTTAATCCAGTAATTC   1160
pKM173_VP2/6_F2  ........................................      0
pKM173_VP2/6_R2  GCTGACGGAGCAACTACATGGTATTTTAATCCAGTAATTC    689

pKM173_VP2/6     TTAGACCAAATAACGTTGAAGTGGAGTTTCTACTAAACGG   1200
pKM173_VP2/6_F2  ........................................      0
pKM173_VP2/6_R2  TTAGACCAAATAACGTTGAAGTGGAGTTTCTACTAAACGG    729

pKM173_VP2/6     GCAGATAATAAACACTTACCAGGCTAGATTTGGAACGATC   1240
pKM173_VP2/6_F2  ........................................      0
pKM173_VP2/6_R2  GCAGATAATAAACACTTACCAGGCTAGATTTGGAACGATC    769

pKM173_VP2/6     GTAGCTAGAAATTTTGATACAATCAGATTGTCGTTTCAGT   1280
pKM173_VP2/6_F2  ........................................
pKM173_VP2/6_R2  GTAGCTAGAAATTTTGATACAATCAGA.............    796

pKM173_VP2/6     TGATGAGACCACCAAATATGACACCATCGGTAGCAGCATT   1320
pKM173_VP2/6_F2  ........................................      0
pKM173_VP2/6_R2  ........................................    796

pKM173_VP2/6     ATTTCCAAATGCGCAACCATTTGAACATCATGCTACAGTA   1360
pKM173_VP2/6_F2  ........................................      0
pKM173_VP2/6_R2  ........................................    796

pKM173_VP2/6     GGACTTACATTGAAAATTGAATCTGCAGTTTGTGAATCTG   1400
pKM173_VP2/6_F2  ........................................      0
pKM173_VP2/6_R2  ........................................    796
```

```
pKM173_VP2/6     TACTTGCTGACGCAAGCGAGACAATGCTAGCAAATGTGAC   1440
pKM173_VP2/6_F2  ........................................     0
pKM173_VP2/6_R2  ........................................   796

pKM173_VP2/6     ATCTGTTAGACAAGAATACGCGATACCAGTTGGACCAGTC   1480
pKM173_VP2/6_F2  ........................................     0
pKM173_VP2/6_R2  ........................................   796

pKM173_VP2/6     TTTCCACCAGGTATGAATTGGACTGATTTGATCACTAACT   1520
pKM173_VP2/6_F2  ........................................     0
pKM173_VP2/6_R2  ........................................   796

pKM173_VP2/6     ATTCACCATCTAGAGAGGATAACTTGCAGCGTGTATTTAC   1560
pKM173_VP2/6_F2  ........................................     0
pKM173_VP2/6_R2  ........................................   796

pKM173_VP2/6     AGTGGCTTCCATTAGAAGCATGCTTGTCAAATAAAGCGCT   1600
pKM173_VP2/6_F2  ..........................AATAAAGCGCT      11
pKM173_VP2/6_R2  ........................................   796

pKM173_VP2/6     ATTAATCCTAGGTGATCTGATCTGCTTACTTTACTTAACG   1640
pKM173_VP2/6_F2  ATTAATCCTAGGTGATCTGATCTGCTTACTTTACTTAACG    51
pKM173_VP2/6_R2  ........................................   796

pKM173_VP2/6     ACCAAAGAAAAACGACAAAAAAAAAAATATTACTACTATTA   1680
pKM173_VP2/6_F2  ACCAAAGAAAAACGACAAAAAAAAAAATATTACTACTATTA    91
pKM173_VP2/6_R2  ........................................   796

pKM173_VP2/6     AAATAAATTAGTATTTTTCTCTTCTTACGATATGATATGA   1720
pKM173_VP2/6_F2  AAATAAATTAGTATTTTTCTCTTCTTACGATATGATATGA    131
pKM173_VP2/6_R2  ........................................   796

pKM173_VP2/6     TGCTATGAAATCATCATCTTCTTAACTTTCTTGTCTCTTA   1760
pKM173_VP2/6_F2  TGCTATGAAATCATCATCTTCTTAACTTTCTTGTCTCTTA    171
pKM173_VP2/6_R2  ........................................   796

pKM173_VP2/6     CACGTCACTTACTCTATATACCCGTTTATATAAGTGTACG   1800
pKM173_VP2/6_F2  CACGTCACTTACTCTATATACCCGTTTATATAAGTGTACG    211
pKM173_VP2/6_R2  ........................................   796

pKM173_VP2/6     TATTTTCTTTTTTTTAAAAAATTTCTATTCTATCCTTAGA   1840
pKM173_VP2/6_F2  TATTTTCTTTTTTTTAAAAAATTTCTATTCTATCCTTAGA    251
pKM173_VP2/6_R2  ........................................   796

pKM173_VP2/6     AAAGTGCCCTTACATCAGTTCCAACGCACTCTAGCTGTTC   1880
pKM173_VP2/6_F2  AAAGTGCCCTTACATCAGTTCCAACGCACTCTAGCTGTTC    291
pKM173_VP2/6_R2  ........................................   796
```

```
pKM173_VP2/6     TGGCACTGTATCTTCATCATGTGCCGGTCGTTTTCCACCC   1920
pKM173_VP2/6_F2  TGGCACTGTATCTTCATCATGTGCCGGTCGTTTTCCACCC    331
pKM173_VP2/6_R2  .....................................     796


pKM173_VP2/6     CAAAAATAACTTTCTTCCCTTTTCCTTTCAATTAATGGCC   1960
pKM173_VP2/6_F2  CAAAAATAACTTTCTTCCCTTTTCCTTTCAATTAATGGCC    371
pKM173_VP2/6_R2  .....................................     796


pKM173_VP2/6     TGGAATTCCGAACCCATTTTCGCATCTGAAACTAATTCTC   2000
pKM173_VP2/6_F2  TGGAATTCCGAACCCATTTTCGCATCTGAAACTAATTCTC    411
pKM173_VP2/6_R2  .....................................     796


pKM173_VP2/6     GAAACCTTTAATATCAAACAATTGAAAAGATCATCATCAC   2040
pKM173_VP2/6_F2  GAAACCTTTAATATCAAACAATTGAAAAGATCATCATCAC    451
pKM173_VP2/6_R2  .....................................     796


pKM173_VP2/6     TAGAAATGAGAAAAAGATCAACAGCACTTAATAACAGTAC   2080
pKM173_VP2/6_F2  TAGAAATGAGAAAAAGATCAACAGCACTTAATAACAGTAC    491
pKM173_VP2/6_R2  .....................................     796


pKM173_VP2/6     GAAAGAAAGATCGCTCGGATCCTAGGGATAACAGGGTAAT   2120
pKM173_VP2/6_F2  GAAAGAAAGATCGCTCGGATCCTAGGGATAACAGGGTAAT    531
pKM173_VP2/6_R2  .....................................     796


pKM173_VP2/6     GGTACCAGAGACCGGGTTGGCGGCGTATTTGTGTCCCAAA   2160
pKM173_VP2/6_F2  .....................................     531
pKM173_VP2/6_R2  .....................................     796
```

**Figure 3.28B: Nucleotide sequence alignments of the recombinant pKM173_VP2 vector containing the ORF of genome segment 6 (VP6) using the *in silico* clone of pKM173_VP2/6.** The *in silico* clone of pKM173_VP2/6 is indicated as pKM173_VP2/6 and forward (forward primer 2) sequence is indicated as pKM173_VP2/6_F2 nucleotide sequence and the reverse (reverse primer 2) sequence is indicated as pKM173_VP2/6_R2 nucleotide sequence. The part of the ORF of genome segment 6 that could not be sequenced with the forward 1 primer and reverse 1 primer is indicated in dark grey. The I-SceI restriction site is indicated in teal.

It is clear from Figure 23.8A that the cassette containing the open reading frame of genome segment 6 was cloned correctly into pKM173_VP2 and that no mutations are visible. The primers designed for sequencing made it possible to confirm the presence of key factors namely the restriction enzymes used for cloning as well as the start and stop codon of the ORF of genome segment 6. A part of the ORF of genome segment 6 (from 719 bp to 1206

bp) could not be sequenced with the forward 1 primer and reverse 1 primer designed for sequencing. However, the forward 2 and reverse 2 primers also designed for sequencing could sequence this part of the ORF of genome segment 6. Figure 3.28B shows the sequencing results of the forward 2 primer and reverse 2 primer indicating that the entire ORF of genome segment 6 was cloned into pKM173_VP2. Transformation of the pKM173_VP2/6 construct into several yeast strains and the possible co-expression of the proteins in these strains were performed using pKM173_VP2/6 colony 1.

### 3.3.4 Co-expression of proteins

In order for the recombinant plasmids to integrate efficiently into the yeast genomes, the recombinant plasmid had to be linearized before transformation could occur. Therefore, the recombinant plasmid sample (pKM177_VP2/6) was subjected to a restriction endonuclease digestion with the restriction enzyme NotI, as described in section 3.2.3.3. A 1.5 µl sample of each 10 µl restriction digestion was analysed on a 1% agarose gel, as shown in Figure 3.29. Two fragments (2210bp and 9994bp) were expected for pKM173_VP2/6, after restriction digestion.

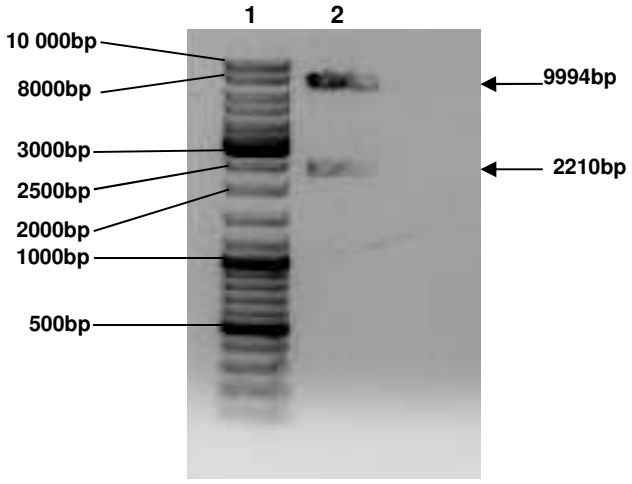


**Figure 3.29: Analysis by agarose gel electrophoresis of pKM173_VP2/6 plasmids, by means of NotI restriction enzyme analysis for transformation purposes.** Lane**:** 1) 10 000bp O'Generuler DNA marker (Fermentas). A 1.5 µl volume of 10 µl NotI digestion reaction was loaded as follows; lanes: 2) pKM173_VP2/6 recombinant plasmid.

Two bands were expected since the pKM173_VP2/6 plasmid contains two NotI restriction cutting sites. The expected results were obtained namely two bands (2210bp and 9994bp),

as seen in Figure 3.29. Transformation of the linearized recombinant plasmid, into the several yeast strains, was done as described in section 3.2.4.3. Twenty five recombinant plasmid colonies of each yeast strain were chosen for further experimental purposes. These colonies were plated out on fresh YPD (Yeast, Peptone, Dextrose) selective plates containng Hygromycin B and incubated overnight at 30°C for all the yeast strains except for *Kluyvermonyces marxianus* which was incubated at 37°C. Ten colonies of each yeast strain, that had the strongest growth, were chosen to screen for possible co-expression of the two rotavirus proteins (VP2 and VP6) in these yeast strains. The co-expression and analysis of the co-expressions of the proteins were done as described in section 3.2.4.4.

***Pierce ECL Western Blotting Substrate***

The analysis of the co-expression of rotavirus proteins VP2 (genome segment 2) and VP6 (genome segment 6) using the Pierce ELC western blotting substrate method was done as described in section 3.2.4.4. Ten recombinant colonies (containing the ORF of genome segment 2 and genome segment 6) for each individual yeast strain (*Kluyveromyces marxianus, Kluyveromyces lactis, Candida deformans, Saccharomyces cerevisiae, Yarrowia lipolytica, Arxula adeninivorans, Hansenula polymorpha* and *Debaryomyces hansenii)* were chosen to do protein analysis on. We expected to see two bands namely one at 48kDa (VP6) and one at 102kDa (VP2). However, no expression for either genome segment 2 (VP2) or genome segment 6 (VP6) was visible for any of the yeast strain colonies chosen for the co-expression analysis. For the Pierce ECL western blotting substrate method different developmental times (30 seconds, 1 min, 2 min and 10 min) were tested to determine whether or not that will make a difference, however, no expression was visible for any of the developmental times used (results not shown).

***BCIP-T/NBT Western Blotting substrate***

The analysis of the co-expression of rotavirus proteins VP2 (genome segment 2) and VP6 (genome segment 6) using the BCIP-T/NBT western blotting substrate was done as described in section 3.2.4.4. No expression of the rotavirus proteins could be obtained with the Pierce ELC western blotting substrate method using the NCDV antibodies that were previously successful with detecting baculovirus expressed VP2 and VP6 in our laboratory (M van der Westhuizen, MSc thesis and Dr K Jere, PhD thesis). The next step was to use the BCIP-T/NBT western blotting substrate with a different antibody, to try and confirm if the

individual expression of VP6, in yeast, can be obtained. The group specific/VP6 antibody was used for this study since it has been used before by the University of Stellenbosch in yeast expression experiments that they have conducted with VP6. The first step was to test the method on our positive control, which is a DLP VP2/6 insect cell lysate obtained from Dr Jere, to see whether or not we got expression with the positive control and if the conditions are optimal for the protein that we want to express. The analysis of the expression of the positive rotavirus control BCIP-T/NBT western blotting substrate (using the group specific/VP6 antibody) was done as described in section 3.2.4.4. We expected to see one band namely at 48kDa (VP6). The expression of the positive VP2 (genome segment 2) and VP6 (genome segment 6) insect cell lysate proteins was developed on a nitrocellulose membrane as seen in Figure 3.30.

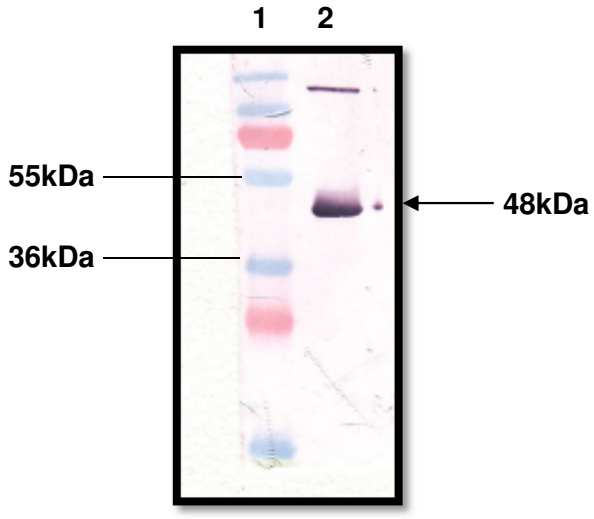


**Figure 3.30: Western Blot analysis of the expression of the positive DLP VP2/6 insect cell lysate control**. Lane: 1) Page Ruler protein marker (Fermentas). A 10 µl sample was used for analysis. Lane: 2) Positive DLP VP2/6 insect cell lysate.

The expected result was obtained namely a band at 48kDa, as seen in Figure 3.30. A non-specific band was also obtained at ± 100kDa which can be due to the fact that a DLP VP2/6 insect cell lysate was used that did not only contain VP6 but also the VP2 protein. Since the expression experiment with the DLP VP2/6 lysate positive control was successful, the next step was to see whether or not expression was possible in the several recombinant yeast strains colonies.

Sixteen recombinant colonies (containing the ORF of genome segment 2 and the ORF of genome segment 6) were chosen for each individual yeast strain (*Kluyveromyces marxianus, Kluyveromyces lactis, Arxula adeninivorans, Hansenula polymorpha* and*)* to do protein analysis on, except for four yeast strains, only seven recombinant colonies were chosen (*Yarrowia lipolytica*, *Debaryomyces hansenii*, *Saccharomyces cerevisiae and Candida deformans)* to do protein analysis on. The recombinant colonies co-expression was developed on nitrocellulose membranes as seen in Figure 3.31.

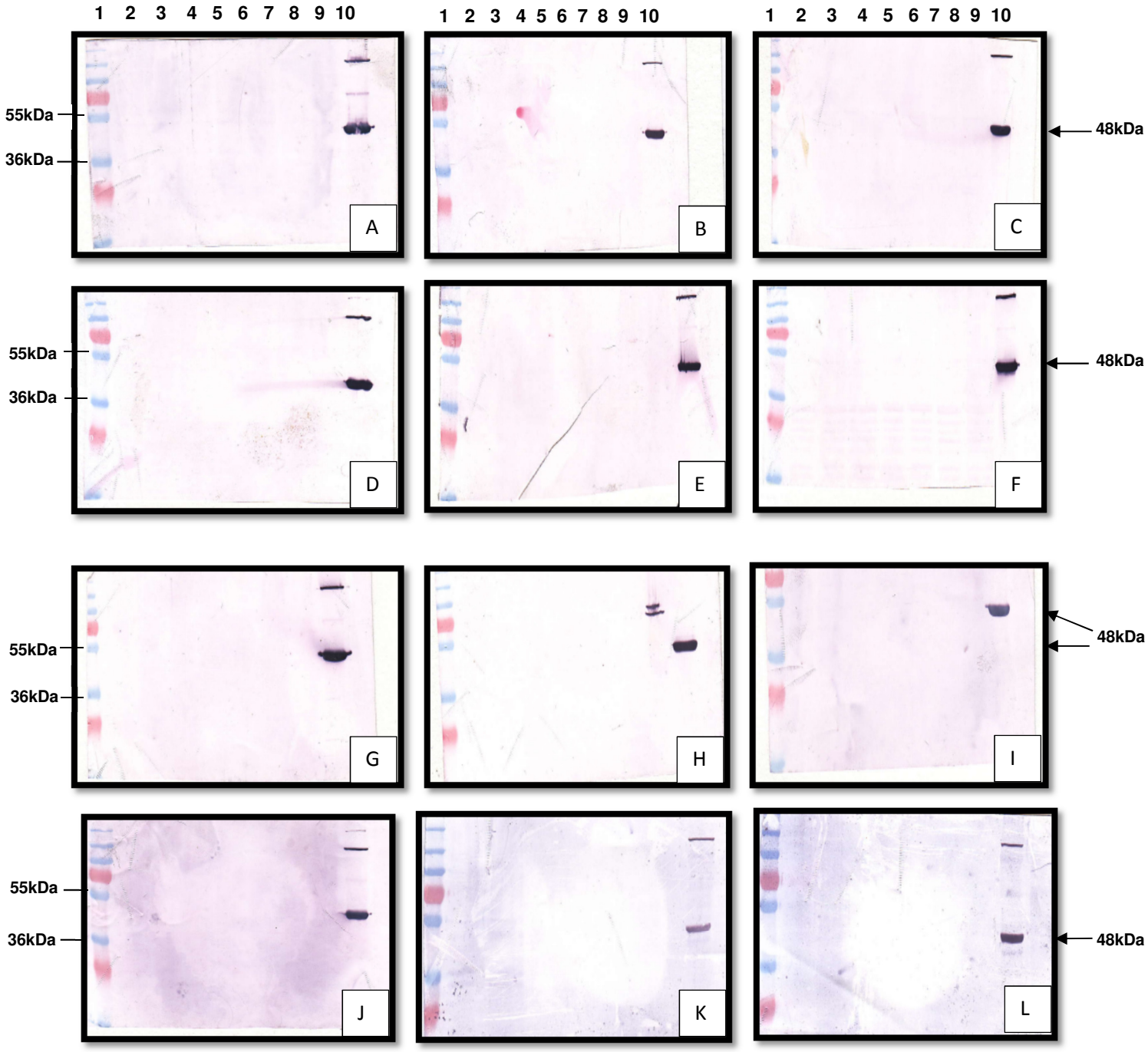


**Figure 3.31: Western blot analysis to detect VP6 expression in *Kluyveromyces marxianus, Kluyveromyces lactis, Candida deformans, Saccharomyces cerevisiae, Yarrowia lipolytica, Arxula adeninivorans, Hansenula polymorpha* and *Debaryomyces hansenii***. A 10 µl sample was used for analysis. All gels were loaded similarly. Lanes; 1) Page Ruler protein marker (Fermentas); 2-9) Colonies (1-7/colonies 9-14), 10) positive VP2/6 control. The different yeast strains tested were as follows: Blots A-B) *Arxula adeninivorans;* C-D) *Kluyveromyces marxianus; E-F) Kluyveromyces lactis; G-H), Hansenula polymorpha*; I) *Saccharomyces cerevisiae; J) Yarrowia lipolytica; K) Debaryomyces hansenii and L) Candida deformans.*

We expected to see one band namely at 48kDa (VP6) since the group specific/VP6 antibody was used that only detects VP6. Unfortunately, no expression for VP6 in yeast could be detected as seen in Figure 3.31. The next step was to determine whether or not the ORFs of genome segment 2 and genome segment 6 were present in these recombinant yeast colonies, since it was possible that the ORFs did not integrate correctly into the different yeast strains. In order to determine whether or not the ORFs were present the same colonies (from each yeast strain used for the western blot analysis) DNA were examined. This was achieved by plating the glycerol stocks of each recombinant yeast strain colony out on YPD agar medium and incubating overnight at 37°C (*Kluyveromyces marxianus)* and 30°C *(Kluyveromyces lactis, Candida deformans, Saccharomyces cerevisiae, Yarrowia lipolytica, Arxula adeninivorans, Hansenula polymorpha* and *Debaryomyces hansenii)*. All further experiments were carried out from these plates.

### 3.3.5 Genomic DNA isolation

Since the co-expression of VP2 (genome segment 2) and VP6 (genome segment 6), in all seven yeast strains, could not be achieved the next step was to see whether or not the ORF of the genome segments were actually present in the colonies used for expression. In order to achieve this, yeast genomic DNA extraction had to be done as described in section 3.2.4.5. Two colonies of each recombinant yeast strain were chosen to do DNA extraction on. A 20 µl sample of each 50 µl DNA extraction was analysed on a 1% agarose gel, as shown in Figure 3.32.

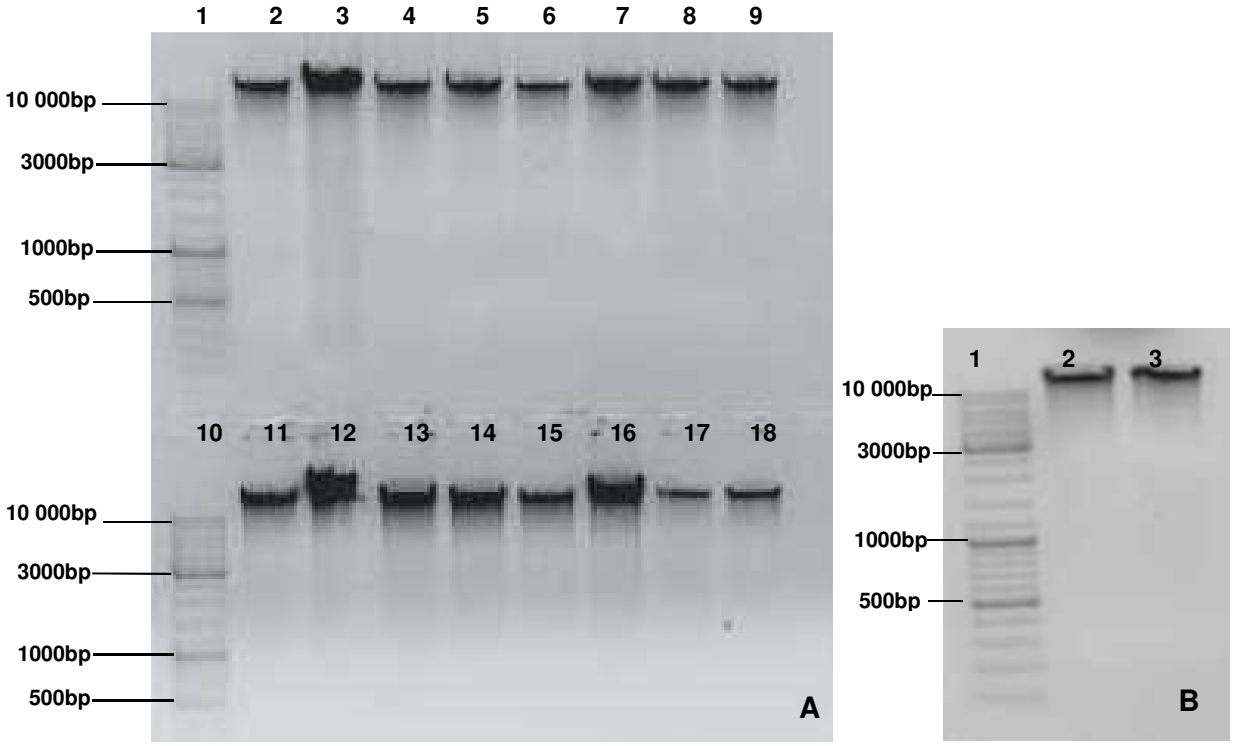


**Figure 3.32: Analysis by agarose gel electrophoresis of genomic DNA extraction of pKM173_VP2/6 containing colonies. (A)** Lanes: 1 and 10) 10 000bp O'Generuler DNA marker. A 20 µl volume of the 50 µl extraction was loaded as follows; lanes: 2) *A. adeninivorans* colony 1; 3) *C. derformans* colony 1; 4) *D. hansenii* colony 1; 5) *H. polymorpha* colony 1; 6) *K. lactis* colony 1; 7) *K. marxianus* colony 1; 8) *S. cerevisiae* colony 1; 9) *Y. lipolytica* colony 1; 11) *A. adeninivorans* colony 3; 12) *C. derformans* colony 2; 13) *D. hansenii* colony 2; 14) *H. polymorpha* colony 2; 15) *K. lactis* colony 2; 16) *K. marxianus* colony 3; 17) *S. cerevisiae* colony 2; 18) *Y. lipolytica* colony 2. **(B)** Lane: 1) 10 000bp O'Generuler DNA marker. A 20 µl volume of the 50 µl extraction was loaded as follows; lanes: 2) *A. adeninivorans* colony 2; 3) *K. marxianus* colony 2.

DNA was obtained for all eighteen colonies as seen in Figure 3.32. All eighteen colonies of pKM173_VP2/6 recombinants, in the different yeast strains, were then screened by means of PCR, as described in section 3.2.3.5. Two PCR reactions were set up for each recombinant colony, one containing primers for the ORF of genome segment 2 and a second PCR reaction containing primers for the ORF of genome segment 6. A colony was considered positive if the PCR reaction gave an amplicon of 2700bp (for the ORF of genome segment 2) and an amplicon of 1270bp (for the ORF of genome segment 6) respectively, that could be seen on a 1% agarose gel. Figure 3.33A and B show the results of the PCR screening for genome segment 2 and genome segment 6 respectively.

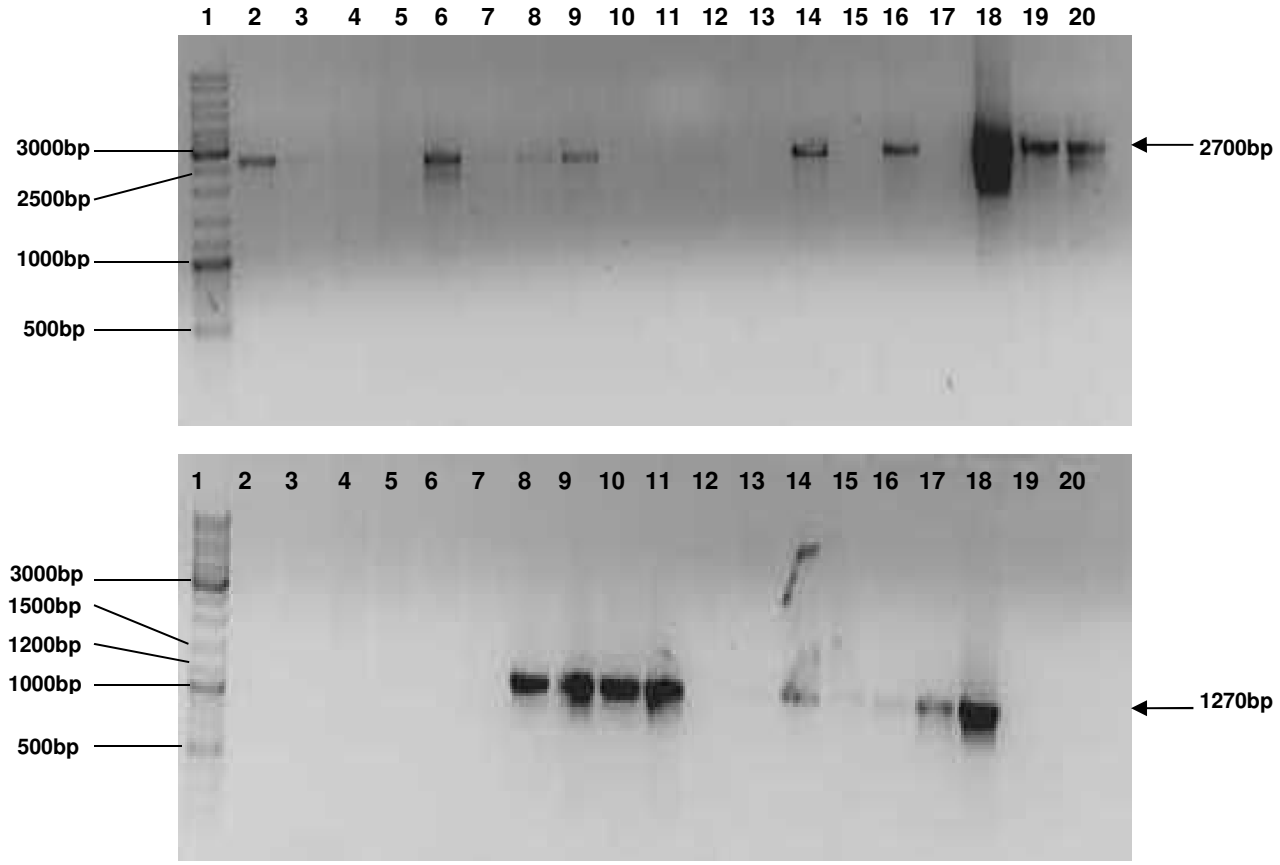


**Figure 3.33: Analysis by agarose gel electrophoresis of PCR colony screening for the ORF of genome segment 2 (VP2) and the ORF of genome segment 6 (VP6). (A)** Lane: 1) 10 000bp O'Generuler DNA marker. A 10 µl volume of each 20 µl colony screening reaction of the ORF of genome segmnet 2 was loaded as follows; lanes: 2) *A. adeninivorans* colony 1; 3) *A. adeninivorans* colony 3; 4) *C. deformans* colony 1; 5) ) *C. deformans* colony 2; 6) *D. hansenii* colony 1; 7) *D. hansenii* colony 2; 8) *H. polymorpha* colony 1; 9) *H. polymorpha* colony 2; 10) *K. lactis* colony 1; 11) *K. lactis* colony 2; 12) *K. marxianus* colony 1; 13) *K. marxianus* colony 3; 14) *S. cerevisiae* colony 1; 15) *S. cerevisiae* colony 2; 16) *Y.lipolytica* colony 1; 17) *Y.lipolytica* colony 2; 18) positive control (pKM173_VP2/6); 19) *A. adeninivorans* colony 2; 20) *K. marxianus* colony 2**. (B)** Lane: 1) 10 000bp O'Generuler DNA marker. A 10 µl volume of each 20 µl colony screening reaction of the ORF of genome segmnet 6 was loaded as follows; lanes: 2) *A. adeninivorans* colony 1; 3) *A. adeninivorans* colony 3; 4) *C. deformans* colony 1; 5) *C. deformans* colony 2; 6) *D. hansenii* colony 1; 7) *D. hansenii* colony 2; 8) *H. polymorpha* colony 1; 9) *H. polymorpha* colony 2; 10) *K. lactis* colony 1; 11) *K. lactis* colony 2; 12) *K. marxianus* colony 1; 13) *K. marxianus* colony 3; 14) *S. cerevisiae* colony 1; 15) *S. cerevisiae* colony 2; 16) *Y. lipolytica* colony 1; 17) *Y. lipolytica* colony 2; 18) positive control (pKM173_VP2/6); 19) *A. adeninivorans* colony 2; 20) *K. marxianus* colony 2.

The results obtained from the PCR colony screening analysis for the ORF of genome segment 2 and genome segment 6 are shown in Figures 3.33A and B, respectively. It was expected that one fragment (2700bp) should result from the PCR colony screening for the ORF of genome segment 2 and one fragment (1270bp) for the ORF of genome segment 6. The expected results were not obtained for all eighteen colonies.

Eight of the eighteen colonies screened for the presence of the ORF of genome segment 2 gave the expected result namely a band at 2700bp. These colonies were *A. adeninivornas* colony 1 and colony 2, *D. hansenii* colony 1, *H. polymorpha* colony 1 and colony 2, *S. cerevisiae* colony 1, *Y. lipolytica* colony 1 and *K. marxianus* colony 2, as seen in Figure 3.33 A. This indicated that the ORF of genome segment 2 encoding for VP2 was only present in these recombinant yeast constructs. Seven of the eighteen colonies screened for the presence of the ORF of genome segment 6 gave the expected result namely a band at 1270bp. These colonies were *H. polymorpha* colony 1 and colony 2, *K. lactis* colony 1 and colony 2, *S. cerevisiae* colony 1 and *Y. lipolytica* colony 1 and colony 2, as seen in Figure 3.33B. This indicated that the ORF of genome segment 6 encoding for VP6 was only present in these recombinant yeast constructs. From the eighteen recombinant yeast colonies screened for the presence of the ORF of genome segment 2 encoding for VP2 and the ORF of genome segment 6 encoding for VP6, only four colonies screened positive for the presence of both genes. These colonies were *H. polymorpha* colony 1 and colony 2, *S cerevisiae* colony 1 and *Y. lipolytica* colony 1.

It is clear from these experiments that the ORF of genome segment 2 and genome segment 6 was present in some of the recombinant yeast constructs, therefore, expression of these proteins should probably have been possible. The next step will be to eliminate the possibilities that could have been responsible for the reason why the rotavirus proteins (VP2 and VP6) did not express in the different yeast strains. An elimination experiment that can be conducted is to see whether or not enough RNA is produced for the proteins to express. An experiment can also be conducted to determine why some of the yeast constructs only contained one of the ORF of the genome segments and not both of the ORF of genome segment 2 and 6.

## 3.4 Summary

This part of the study had three main objectives. Firstly, to clone the ORFs of the wild type genome segment 2 and 6 of the human strain GR10924 G9P[6] genotype, into the pKM173 and pKM177 vectors respectively. Secondly, to clone the cassette containing genome

segment 6, into the pKM173_VP2 vector, to obtain the pKM173_VP2/6 construct. Thirdly, to investigate the possible co-expression of the wild type rotavirus structural proteins VP2 and VP6 in several yeast strains.

The various clones (pKM177_VP6, pKM173_VP2 and pKM173_VP2/6) were successfully constructed and the yeast strains engineered for the expression of the proteins were successfully engineered as well. The co-expression of the recombinant proteins VP2 and VP6, in the pKM173 vector could not be achieved in any of the several yeast strains (*Kluyveromyces marxianus, Kluyveromyces lactis, Candida deformans, Saccharomyces cerevisiae, Yarrowia lipolytica, Arxula adeninivorans, Hansenula polymorpha* and *Debaryomyces hansenii)*. Two different types of western blot analyses were used to verify protein expression namely the Pierce ECL Western Blotting Substrate method and the BCIP-T/NBT western blotting substrate method. Two types of antigens were used for the western blotting substrate methods namely 1) the group specific/VP6 antigen and 2) the goat polyclonal anti-rotavirus antibody with the secondary antibody donkey horseradish peroxidase conjugated anti-goat IgG. Using DNA extraction experiment and PCR analyses it was determined that the ORFs encoding for the proteins (VP2 and VP6) were present in some of the yeast strains. However, only four of the colonies screened for the presence of the ORFs encoding for the proteins contained both ORFs (genome segment 2 and genome segment 6), while the other only contained one of the two ORFs encoding for the proteins.

The conclusion of the work described in this chapter is: the recombinant proteins VP2 and VP6 could not be expressed in several yeast strains. However, it was determined that the DNA encoding for the proteins was present in some of the yeast strains indicating that they probably should have been able to be expressed in these yeast strains. Further studies can now be conducted to see where we can improve to make the co-expression of the recombinant proteins in these yeast strains possible. These studies include firstly, determining why both ORFs encoding for the proteins are not present in all the yeast strains. This can be done by conducting studies to determine if some of the yeast strains do not discard one of the ORFs encoding for the proteins during expression experiments. Secondly, to determine why the proteins did not express in the yeast strains. This can be done by determining if enough RNA is produced for protein expression to occur, if the expression levels of these proteins are not very low and lastly if the expression conditions are optimal for the expression of these two proteins (VP2 and VP6).

# Chapter 4:

# Concluding remarks and future prospects

Rotavirus causes gastroenteritis worldwide with children in developing countries, such as Sub-Saharan Africa and India, being affected the most. The two licensed rotavirus live attenuated vaccines (Rotarix and RotaTeq) have various shortcomings as described in Chapter 1 (section 1.9.3). Therefore, safer and better locally effective alternative vaccine candidates should be investigated such as non-live vaccine candidates since they have advantages that live attenuated vaccines do not have. One such an advantage is the fact that they contain the same antigenic properties as live vaccines but cannot replicate because they do not contain any genetic material. It has also been shown that the currently licensed rotavirus vaccines are not as efficacious in developing countries such as in Africa where their combined efficiency is only 55.4%. Factors that can affect the efficacy might be the frequent occurrence of breastfeeding, malnutrition and the prevalence of HIV (immuno-compromised) infections in these regions. It can also be due to the fact that different serotypes are prevalent in these regions (Seheri *et al.*, 2014). Other disadvantages of the currently licensed vaccines include the high cost of the vaccines and the fact that the vaccine manufacturing companies do not have the production capacity to produce enough vaccines for global distribution. In this study the local GR10924 G9P[6] strain was used, since the G9 and P6 serotypes are prevalent in Sub-Saharan Africa.

The first experimental chapter (Chapter 2) of this study reported the cloning and expression of the bacterial codon optimised open reading frames (ORFs) of genome segment 2 (VP2) and genome segment 6 (VP6) in bacteria. The ORFs of the bacterial codon optimised genome segment 2 (VP2) and genome segment 6 (VP6) were sub-cloned into the dual expression vector, pETDuet-1, from their original vectors, respectively, using compatible restriction enzymes. Three constructs were prepared for the bacterial codon optimised genome segments namely, pETDuet-1_VP6, pETDuet-1_VP2 and pETDuet-1_VP2/6, to evaluate the individual and co-expression of VP2 (genome segment 2) and VP6 (genome segment 6) in bacteria. The individual expression of VP6 resulted in insoluble proteins, as observed by a Coomassie stained SDS-PAGE. In order to investigate if the VP6 protein can be expressed soluble, the construct was expressed with the chaperone pGro7 and analysed with three different lysis buffers namely (i) Phosphate buffer solution (PBS) containing 0.5% nonyl phenoxyl-polyethoxyl-ethanol (NP40), (ii) 10 mM Tris containing 0.1 mM EDTA, 1%

dissolved organic carbon (DOC), 0.1% SDS and (iii) 10 mM Tris containing 0.1 mM EDTA, 1% SDS. Results were obtained that showed that the VP6 protein could be soluble in bacteria with the second lysis buffer (10 mM Tris; 0.1 mM EDTA, 1% DOC; 0.1% SDS). The expression of the bacterial codon optimised VP2 as well as the co-expression of the bacterial codon optimised VP2/6 was not possible. To eliminate the possible reasons for the failure to express, the pColdTF_VP2 construct, a cold shock based expression vector, was used to determine if the VP2 protein could be expressed. Since, the bacterial codon optimised VP2 could be expressed in the pColdTF vector but not in the pETDuet-1 vector, it is possible that the problem may be with the second multiple cloning site of pETDuet-1. Two points support this possibility. Firstly, the VP2 protein (genome segment 2) could be expressed in the pColdTF vector. Secondly, sequencing confirmed that no mutations were introduced into the open reading frame of genome segment 2 (VP2) during cloning. However, it has to be taken into consideration that the VP2 protein was expressed in the presence of a trigger factor in the pColdTF vector. Therefore, future studies still have to be conducted to determine exactly what caused VP2 not to express using the pETDuet-1 vector and if the expression of VP2 is only possible in the presence of a chaperone such as a trigger factor.

The second experimental chapter (Chapter 3) reports the engineering of several yeast strains that contain the ORFs of the wild type genome segment 2 and genome segment 6 that could express the rotavirus proteins (VP2 and VP6). The ORF of the wild type genome segment 2 (VP2) and genome segment 6 (VP6) was sub-cloned from their original plasmids, respectively, using compatible restriction enzyme sites. Two constructs were obtained namely pKM173_VP2 and pKM177_VP6. The expression cassette containing genome segment 6 as well as the yeast promoter and terminator regions were sub-cloned from pKM177_VP6 into the pKM173_VP2 construct, using the compatible restriction enzyme I-SceI. This yielded the construct pKM173_VP2/6 which was used for all expression purposes. The pKM173_VP2/6 construct was transformed into several yeast strains namely *Kluyveromyces marxianus, Kluyveromyces lactis, Yarrowia lipolytica, Debaryomyces hansenii, Candida deformans, Hansenula polymorpha, Arxula adeninivorans* and *Saccharomyces cerevisiae.* Co-expression of VP2 (genome segment 2) and VP6 (genome segment 6) could not be observed using western blot analysis. Two types of antibodies and western blot analysis were used. Firstly, the Pierce ECL western blotting substrate method was used with the goat polyclonal anti-rotavirus antibody and the secondary antibody donkey horseradish peroxidase conjugated anti-goat IgG. Secondly, the BCIP-T/NBT western blotting substrate method was used with the group specific (VP6) antibody. Co-expression of VP2 and VP6 was not detected in the several yeast strains screened, using

both western blot methods and both antibodies. To determine the possible reason for the proteins not expressing in the yeast strains, DNA extraction from the recombinant yeast strains was done. Two colonies of each of the various yeast strains, which were used for the expression experiments, were chosen for DNA extraction to determine whether or not the constructs contained the two genome segments of interest. It was observed that some of the constructs only contained one of the genome segments of interest and that only four of the eighteen constructs screened contained both genome segments 2 and 6. Therefore, further studies will have to be conducted to determine why the co-expression of VP2/6 was not possible in several yeast strains. Future experiments that could be conducted are to i) determine why some of the constructs only contained one of the two genome segments of interest and ii) determine if RNA was produced for the translation of proteins for expression.

When future studies are conducted and the co-expression of VP2/6, in bacteria and yeast, can be achieved, the next step will be to determine if double-layered particles (DLPs) can be produced in bacteria and yeast, by means of transmission electron microscopy. If DLPs can be produced based on local rotavirus strains, in these expression systems, it might become possible to produce safer and more efficacious alternatives to the currently licensed rotavirus vaccines or complement the current live rotavirus vaccines in developing countries.

# Appendix A

# Materials utilized in study

| Item | Product number | Supplier |
|---|---|---|
| 4 x Laemmli sample buffer | 161-0737 | Biorad |
| 6 x Orange loading dye | R0631 | Fermentas |
| Acetic acid | 1.0063.2500 | Merck |
| Acrylamide | 1.00209.10 | Merck |
| AGFA rapid developing solution | G153 | AGFA |
| AGFA rapid fixer solution | G354 | AGFA |
| Ammonium persulfate | A3678 | Sigma |
| Ampicillin | A0166 | Sigma |
| Antartic Phosphatse | M02895 | Biolabs |
| BCIP-T | R0821 | Thermo Scientific |
| BgIII | ER0081 | Fermentas |
| Bicine | B3876 | Sigma |
| Buffer O | BO5 | Fermentas |
| Buffer R | BR5 | Fermentas |
| Bugbuster | 70584-4 | Novagen |
| Coomassie Brilliant Blue R250 | B8467 | Sigma |
| DMSO | D2650 | Sigma |

| | | |
|---|---|---|
| dNTP | BG3001A | Takara |
| Donkey-anti-goat IgG | Ab97120 | Abcam |
| DTT | R0862 | Thermo Scientific |
| Eco47III | ER0321 | Fermentas |
| EcoRI | ER0271 | Fermentas |
| EDTA | G3685 | Fluka analytical |
| Ethanol | 1.00983.2500 | Merck |
| Ethidium bromide | 160539 | Sigma |
| Ex-Taq buffer | RR001A | Takara |
| Ex-Taq enzyme | RR001AM | Takara |
| Fermentas pageruler | SM1811 | Fermentas |
| GoTaq | M891A | Promega |
| GoTaq green buffer | M830B | Promega |
| Glucose | AP008331.500 | Merck |
| Glycerol | 49780 | Fluka |
| Glycine | 1.04169.1000 | Merck |
| Hygromycin B | 400052 | CalBiochem |
| IPTG | R0391 | Fermentas |
| I-SceI | ER1771 | Fermentas |
| Isopropanol | 1.09634.2500 | Merck |
| Kanamycin | 11815-032 | Gibco |
| Lithium acetate | L6883 | Sigma |

| | | |
|---|---|---|
| Magnesium Chloride (PCR purpose) | F510MG | Finnzymes |
| Magnesium Chloride | 8.14733.0500 | Merck |
| Magnesium Sulfate | M4643 | Merck |
| Methanol | 1.03009.2500 | Merck |
| Rotavirus group specific antigen | sc-52198 | Santa Cruz |
| NBT | R0841 | Thermo scientific |
| NcoI | ER0571 | Fermentas |
| N'N, Bis-acrylamide | 130672 | Sigma |
| NP40 | 74385 | Fluka |
| NotI | ER0592 | Fermentas |
| Nucleospin extract II kit | 636972 | Clonetech |
| O'Generuler DNA ladder mix | SM1773 | Fermentas |
| pETDuet-1 | TB337 | Novagen |
| PEG-4000 | 8.07490.1000 | Merck |
| Phosphate buffer solution (PBS) | P4417 | Sigma |
| Polyclonal goat primary antibody | Ab69560 | Abcam |
| Ponceou S | 101195165 | Sigma |
| Potassium acetate | 204.8222 | Sigma |
| Potassium chloride | AB004936.500 | Merck |
| Promega Pureyield plasmid midi-prep system | A2495 | Promega |

| | | |
|---|---|---|
| Protein loading buffer x4 dual colour | R1011 | Fermentas |
| QIAGEN® Plasmid Mini, Midi, Maxi, Mega and Giga kit | 12145 | Qiagen |
| SDS | L-4390 | Sigma |
| Sodium Chloride | S3014 | Sigma |
| Sodium hydroxide | S5881 | Sigma |
| Sorbitol | S1876 | Sigma |
| T4 DNA ligase | EL0011 | Thermo Scientific |
| T4 DNA ligase buffer | B69 | Thermo Scientific |
| Tango buffer | BY5 | Fermentas |
| TEMED | 1.10732.026 | Merck |
| Trisbase | 11814273001 | Roche |
| Tryptone | 1.10676.0500 | Merck |
| Tween | 1001331191 | Sigma |
| XbaI | ER0681 | Fermentas |
| XhoI | ER091 | Fermentas |
| XmaJI (AvrII) | ER1561 | Fermentas |
| Yeast extract | HG00BX6.500 | Merck |
| ZR Fungal/Bacterial DNA MiniPrep | D6005 | Zymo Research |

# Appendix B

# List of Figures

# Appendix C

# List of Tables

# Appendix D

# Abbreviations

***Chapter 1:***

**ACIP:** Advisory committee on immunization practice

**BBIL:** Bharat Biotech International Ltd

**Bp:** Base pair

**BVES:** Baculovirus expression system

***B.subtillus:*** *Bacillus subtilis*

**Ca:** Calcium

**$CD8^+$ T-cells:** Cytotoxic T-cells

**$CD4^+$ T-cells:** T-helper cells

**CDC:** Center for Disease control and prevention

**CT:** Cholera toxin

**DLP:** Double-layered particle

**DNA:** Deoxyribonucleic acid

**dsRNA:** Double stranded ribonucleic acid

***E.coli:*** *Escherichia coli*

**EDIM:** Epizootic diarrhoea of infant mice

**eIF4G:** Eukaryotic translation initiation factor 4G

**ER:** Endoplasmic reticulum

**FDA:** Food and Drug administration

**GAVI:** Global Alliance for vaccine and immunisation

**G-type:** Glycosylated structure

**HIV:** Human immunodeficiency virus

**IFN:** Interferon

**IgA/G/M:** Immunoglobulin A/G/M

**IPV:** Inactivated polio vaccine

**IRF:** Interferon regulatory factor

**IRV:** Inactivated rotavirus vaccine

**ISG:** Interferon stimulated gene

**LT:** Heat labile toxin

**MBP:** Maltose binding protein

**mRNA:** Messenger ribonucleic acid

**NF:** Nuclear factor

**NIAID:** National Institute of Allergy and infectious diseases

**NIH:** National Institute of Health

**NSP:** Non-structural proteins

**NTPase:** Nucleoside triphosphatases

**OmpA:** Outer membrane protein A

**OPV:** Oral polio vaccine

**PABP:** Poly (A) binding protein

**PATH:** Program for Appropriate Technology in Health

**PCV:** Porcine circovirus

**P-type:** Protease sensitive

**PRR:** Pattern recognition receptors

**PRV:** Porcine rotavirus

**RNA:** Ribonucleic acid

**RRV-TV vaccine:** Rhesus rotavirus tetravalent vaccine

**SAGE:** Strategic advisory group of experts

**SCID:** Severe combined immunodeficiency

**ssRNA:** Single stranded ribonucleic acid

**TB:** Tuberculosis

**TLP:** Triple-layered particle

**US:** United States

**VAERS:** Vaccine Adverse events reporting system

**VLP:** Virus-like particle

**VP:** Virus protein

**WHO:** World Health Organization

***Chapter 2:***

**× g:** G-force

**µg:** Microgram

**µl:** Microliter

**µm;** Micromolar

**APS:** Ammoniumperoxodisulphate

**DMSO:** Dimethylsulfoxide

**dNTP:** Deoxynucleotide Triphosphate

**DOC:** Dissolved organic carbon

**EDTA:** Ethylenediaminetetraacetic Acid

**gor:** Glutathione reductase

**IDT:** Integrated DNA technologies

**IPTG:** Isopropyl-β-D-thio-galactoside

**kDa:** Kilo Dalton

**M:** Molar

**MCS:** Multiple cloning site

**mg:** Miligram

**ml:** Mililiter

**mM:** Milimolar

**NP40:** Nonyl phenoxyl-polyethoxyl-ethanol

**ng:** Nanogram

**nm:** Nano metre

**ORF:** Open reading frame

**Ori:** Origin of replication

**PBS:** Phosphate buffer solution

**PCR:** Polymerase chain reaction

**rbs:** Ribosome binding site

**rpm:** Rotation per minute

**SDS:** Sodium deodecyl sulphate

**SOB:** Super optimal broth

**TAE:** Tris acetate Ethylenediaminetetraacetic Acid

**TEMED:** Tetramethylethylenediamine

**trxB:** Thioredoxin reductase

**V:** Volt

**v/v:** Volume/Volume

**w/v:** Weight/ Volume

***Chapter 3:***

**AP:** Alkaline phosphatase

**DTT:** Diyhiothreitol

**hph:** Hygromycin B resistance gene

**LiAc:** Lithium acetate

**NBT:** Nitro blue tetrazolium

**NCDV:** Nebraska calf diarrhoea virus

**OSB:** One step buffer

**PEG:** Polyethylene glycol

**TEF:** Translation elongation factor

**TGS:** Tris glycine SDS (Sodium deodecyl sulphate)

**UFS:** University of the Free State

**YPD:** Yeast extract peptone dextrose

# Appendix E

# References


1. Albertyn, J., Labuschagne, M., Theron, C., Smith, M. S. (2011). Novel expression constructs. Patent 2011/04652. Filed 23 June 2011.
2. Anderson, E. J., & Weber, S. G. (2004). Review Rotavirus infection in adults, *Lancet Infectious Diseases 4*: 91–99.
3. Angel, J., Franco, M. a, & Greenberg, H. B. (2007). Rotavirus vaccines: recent developments and future considerations. *Nature reviews. Microbiology*, *5*(7): 529–539.
4. Angel, J., Franco, M. a, & Greenberg, H. B. (2012). Rotavirus immune responses and correlates of protection. *Current Opinion in Vvirology*, *2*(4): 419–425.
5. Armah, G. E., Sow, S. O., Breiman, R. F., Dallas, M. J., Tapia, M. D., Feikin, D. R., Binka, F. N., *Steele, A. D., Laserson, K. L., Ansah, N. A., Levine, M.M., Lewis, K. Coia, M. L., Attah-Poku, M., Ojwando, J., Rivers,S. B., Victor, J.C., Nyambane, G., Hodgson, A., Schödel, F., Ciarlet, M., Neuzil*, K. M. (2010). Efficacy of pentavalent rotavirus vaccine against severe rotavirus gastroenteritis in infants in developing countries in sub-Saharan Africa: a randomised, double-blind, placebo-controlled trial. *Lancet*, *376*(9741): 606–14.
6. Arnold, M. M., Sen, A., Greenberg, H. B., & Patton, J. T. (2013). The battle between rotavirus and its host for control of the interferon signaling pathway. *PLoS Pathogens*, *9*(1):1-8.
7. Attoui, H., Mertens, P. P. C., Becnel, J., Bellaganahali, S. & others (2011). Reoviridae. *In:* King, A., Lefkowitz, E., Adams, M. J. & Carstens, E. B. (eds.) *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses.* Elsevier: 603-639.
8. Azevedo, M. S. P., Gonzalez, A. M., Yuan, L., Jeong, K.-I., Iosef, C., Van Nguyen, T., Lovgren-Bengtsson, K., Morein, B., Saif, L. J. (2010). An oral versus intranasal prime/boost regimen using attenuated human rotavirus or VP2 and VP6 virus-like particles with immunostimulating complexes influences protection and antibody-secreting cell responses to rotavirus in a neonatal gnotobiotic pig mode. *Clinical and Vaccine Immunology, 17*(3), 420–8.

9. Barnes, G. L., Lund, J. S., Mitchell, S. V, De Bruyn, L., Piggford, L., Smith, A. L., Furmedge, J., Masendycz, P. J., Bugg, H. C., Bogdanovic-Sakran, N., Carlin, J. B., Bishop, R. F. (2002). Early phase II trial of human rotavirus vaccine candidate RV3. *Vaccine*, *20*(23-24): 2950–2956.
10. Bernstein, D. I. (2006). Live attenuated human rotavirus vaccine, Rotarix. *Seminars in Pediatric Infectious Diseases*, *17*(4): 188–194.
11. Bhandari, N., Rongsen-Chandola, T., Bavdekar, A., John, J., Antony, K., Taneja, S., Goyal, N., Kawade, A., Kang, G., Singh, S., Juvekar, S., Muliyil, J., Arya, A., Shaikh, H., Abraham, V., Vrati, S., Proschan, M., Kohberger, R., Thiry, G., Glass, R., Greenberg, H. B., Curlin, G., Mohan, K., Harshavardhan, G.V.J.A. Prasad, S., Rao, T.S., Boslego, J., Bhan. M. J. (2014). Efficacy of a monovalent human-bovine (116E) rotavirus vaccine in Indian children in the second year of life. *Vaccine*, *325*(1): 110–116.
12. Bhandari, N., Sharma, P., Taneja, S., Kumar, T., Rongsen-Chandola, T., Appaiahgari, M. B., Mishra, A., Singh, S., Vrati, S., and the Rotavirus Vaccine Development Group (2009). A dose-escalation safety and immunogenicity study of live attenuated oral rotavirus vaccine 116E in infants: a randomized, double-blind, placebo-controlled trial. *The Journal of Infectious Diseases*, *200*(3): 421–429.
13. Bishop, R. (2009). Discovery of rotavirus: Implications for child health. *Journal of gastroenterology and hepatology*, *24(3)*: 81–85.
14. Birnboim, H.C., Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* 7(6): 1513-1523.
15. Carter, G. R., & Wise, D. J., Flores E.F. (2005). A Concise Review of Veterinary Virology, *International Veterinary Information Service*, A3416.0805.
16. Chan, J., Nirwati, H., Triasih, R., Bogdanovic-Sakran, N., Soenarto, Y., Hakimi, M., Duke, T., Buttery, J. P., Bines, J. E., Bishop, R. F., Kirkwood, C. D., Danchin, M. D. (2011). Maternal antibodies to rotavirus: could they interfere with live rotavirus vaccines in developing countries? *Vaccine*, *29*(6): 1242–1247.
17. Choi, A. H., Basu, M., Neal, M. M. M. C., Clements, J. D., & Ward, R. L. (1999). Antibody-Independent Protection against Rotavirus Infection of Mice Stimulated by Intranasal Immunization with Chimeric VP4 or VP6 Protein. *Journal of Virology, 73*(9): 7574–7581.
18. Choi, A., Basu, M., McNeal, M. M., Bean, J.A., Clements, J. D., Ward, R. L. (2004). Intranasal administration of an *Escherichia coli* expressed codon-optimized rotavirus VP6 protein induces protection in mice. *Protein Expression and Purification*, 53(1): 238.

19. Ciarlet, M., & Schödel, F. (2009). Development of a rotavirus vaccine: clinical safety, immunogenicity, and efficacy of the pentavalent rotavirus vaccine, RotaTeq. *Vaccine*, *275(*6*)*: 72–81.
20. Clements-mann, M. Lou, Makhene, M. K., Mrukowicz, J., Wright, P. F., Hoshino, Y., Midthun, K., Sperber, E., Karron, R., Kapikian, A. Z. (1999). Safety and immunogenicity of live attenuated human ± bovine (UK) reassortant rotavirus vaccines with VP7-speci ® city for serotypes 1 , 2 , 3 or 4 in adults, children and infants 1, *Vaccine*, *17*: 2715–2725.
21. Clements-Mann, M. L., Dudas, R., Hoshino, Y., Nehring, P., Sperber, E., Wagner, M., Stephens, I., Karron, R., Deforest, A., Kapikian, A. Z. (2001). Safety and immunogenicity of live attenuated quadrivalent human-bovine (UK) reassortant rotavirus vaccine administered with childhood vaccines to infants. *Vaccine*, *19*(32): 4676–4684.
22. Conner, M. E., Zarley, C. D., Hu, B., Parsons, S., Drabinski, D., Greiner, S., Smith, R., et al. (1996). Virus-like particles as a rotavirus subunit vaccine. *The Journal of Infectious Diseases*, *174*(1): 88–92.
23. Cortes-Perez, N. G., Sapin, C., Jaffrelo, L., Daou, S., Grill, J. P., Langella, P., Seksik, P., Beaugerie, L., Chwetzoff, S., Trugnan, G. (2010). Rotavirus-like particles: a novel nanocarrier for the gut. *Journal of Biomedicine & Biotechnology*, *2010*, 317545.
24. Crawford, S. U. E. E., Labbe, M., Cohen, J., Burroughs, M. H., Zhou, Y., & Estes, M. K. (1994). Characterization of Virus-Like Particles Produced by the Expression of Rotavirus Capsid Proteins in Insect Cells, *Journal of Virology, 68*(9): 5945–5952.
25. Danchin, M.H., Buttery, J., Kirkwood, C., Watts, E., Bishop, R., Barnes, G., Carlin, J., Lee, K., Justice, F., Clifford, V., Strong, D., Bines, J. (2013). PhaseI safety trial of RV3-BB rotavirus vaccine - a human neonatal candidate rotavirus vaccine. *Vaccine*, *28;31*(26): 2610-2616.
26. Dennehy, P. H. (2008). Rotavirus vaccines: an overview. *Clinical Microbiology Reviews*, *21*(1): 198–208.
27. Desselberger, U., & Huppertz, H.-I. (2011). Immune responses to rotavirus infection and vaccination and associated correlates of protection. *The Journal of Infectious Diseases*, *203*(2): 188–195.
28. Desselberger, U., Manktelow, E., Li, W., Cheung, W., Iturriza-Gómara, M., & Gray, J. (2009). Rotaviruses and rotavirus vaccines. *British Medical Bulletin*, *90*: 37–51.
29. Donato, C. M., Cannan, D., Bogdanovic-Sakran, N., Snelling, T. L., & Kirkwood, C. D. (2012). Characterisation of a G9P[8] rotavirus strain identified during a gastroenteritis outbreak in Alice Springs, Australia post Rotarix™ vaccine introduction. *Vaccine*, *30(1)*: 152–158.

30. Dore, D. D., Turnbull, B. R., & Seeger, J. D. (2012). Vaccine discontinuation and switching following regulatory interventions in response to rotavirus vaccine contamination with porcine circovirus DNA fragments, *Pharmacoepidemiology and Drug Safety*, 21: 415–419.
31. Estes, M. K., & Cohen, J. (1989). Rotavirus gene structure and function. *Microbiological Reviews*, *53*(4): 410–449.
32. Estes, M.K. and Kapikian, A.Z. (2007). Rotaviruses. In Fields' Virology, Knipe, D.M.,Howley, P.M. (eds) 5th edition. Philadelphia: Lippincott, Williams & Wilkins. 1918–1974.
33. Franco, M. A., Angel, J., & Greenberg, H. B. (2006). Immunity and correlates of protection for rotavirus vaccines. *Vaccine*, *24*(15): 2718–2731.
34. Fu, C., Wang, M., Liang, J., He, T., Wang, D., & Xu, J. (2007). Effectiveness of Lanzhou lamb rotavirus vaccine against rotavirus gastroenteritis requiring hospitalization: a matched case-control study. *Vaccine*, *25*(52): 8756–8761.
35. Gentsch, J. R., Laird, A. R., Bielfelt, B., Griffin, D. D., Banyai, K., Ramachandran, M., Jain, V., Cunliffe, N. A., Nakagomi, O., Kirkwood, C. D., Fischer, T. K., Parasher, U. D., Breese, J. S., Jiang, B., Glass, R. I. (2005). Serotype diversity and reassortment between human and animal rotavirus strains: implications for rotavirus vaccine programs. *The Journal of Infectious Diseases*, *192(*1): 146–159.
36. Glass, R. I., Bresee, J. S., Parashar, U. D., Jiang, B., & Gentsch, J. (2004). The future of rotavirus vaccines: a major setback leads to new opportunities. *Lancet*, *363*(9420): 1547–1550.
37. Glass, R. I., Parashar, U. D., Bresee, J. S., Turcios, R., Fischer, T. K., Widdowson, M.-A., Jiang, B., Gentsch, J. R. (2006). Rotavirus vaccines: current prospects and future challenges. *Lancet*, *368*(9532): 323–332.
38. Greenberg, H. B., & Estes, M. K. (2009). Rotaviruses : From Pathogenesis to Vaccination. *Gastroenterology*. *136*(6): 1939–1951.
39. Greig, S. L., Berriman, J. a, O'Brien, J. a, Taylor, J. a, Bellamy, a R., Yeager, M. J., & Mitra, A. K. (2006). Structural determinants of rotavirus subgroup specificity mapped by cryo-electron microscopy. *Journal of Molecular Biology*, *356*(1): 209–221.
40. Groome, M. J., Moon, S., Velasquez, D., Jones, S., Koen, A., van Niekerk, N. (2014). Effect of breastfeeding on immunogenicity of oral live-attenuated human rotavirus vaccine : a randomized trial in HIV-uninfected infants in Soweto, South Africa. *Bull World Health Organ, 92*: 238-245.
41. Herrmann, J. E., Chen, S. C., Fynan, E. F., Santoro, J. C., Greenberg, H. B., Wang, S., & Robinson, H. L. (1996). Protection against rotavirus infections by DNA vaccination. *The Journal of Infectious Diseases*, *174:* 93–97.

42. Jayaram, H., Estes, M. K., & Prasad, B. V. V. (2004). Emerging themes in rotavirus cell entry, genome organization, transcription and replication. *Virus Research*, *101*(1): 67–81.
43. Jere, K. C., Mlera, L., Neill, H. G. O., Potgieter, A. C., Page, N. A., Seheri, M. L., van Dijk, A. A. (2011). Whole Genome Analyses of African G2 , G8 , G9 , and G12 Rotavirus Strains Using Sequence-Independent Amplification and 454 W Pyrosequencing. *Journal of Medical Virology*, *2042*(83): 2018–2042.
44. Jiang, B, Estes, M. K., Barone, C., Barniak, V., O'Neal, C. M., Ottaiano, a, Madore, H. P., Conner, M. E. (1999). Heterotypic protection from rotavirus infection in mice vaccinated with virus-like particles. *Vaccine*, *17*(7-8): 1005–1113.
45. Jiang, Baoming, Wang, Y., Saluzzo, J.-F., Bargeron, K., Frachette, M.-J., & Glass, R. I. (2008). Immunogenicity of a thermally inactivated rotavirus vaccine in mice. *Human Vaccines*, *4*(2): 143–147.
46. Kang, G. (2006). Rotavirus vaccines. *Indian Journal of Medical Microbiology*, 24(4): 252-257.
47. Kapikian, A. Z., Simonsen, L., Vesikari, T., Hoshino, Y., Morens, D. M., Chanock, R. M., La Montagne, J. R., et al. (2005). A hexavalent human rotavirus-bovine rotavirus (UK) reassortant vaccine designed for use in developing countries and delivered in a schedule with the potential to eliminate the risk of intussusception. *The Journal of Infectious Diseases*, *192(*1): S22–299.
48. Khoury, H., Ogilvie, I., El Khoury, A. C., Duan, Y., & Goetghebeur, M. M. (2011). Burden of rotavirus gastroenteritis in the Middle Eastern and North African pediatric population. *BMC Infectious Diseases*, *11*(1): 9.
49. Kim, H., Park, J. G., Alfajaro, M. M., Kim, D., Hosmillo, M., Son, K., Lee, J., Bae, Y., Park, S., Kang, M., Cho, K. (2012). Pathogenicity characterization of a bovine triple reassortant rotavirus in calves and piglets. *Veterinary microbiology*, *159*(1-2): 11–22.
50. Lappalainen, S., Rodríguez-Limas, W. a, Pastor, A. R., Esquivel-Soto, E., Esquivel-Guadarrama, F., Ramírez, O. T., Palomares, L. A. (2014). Immunogenicity and protective efficacy of yeast extracts containing rotavirus-like particles: a potential veterinary vaccine. *Vaccine*, 32(24): 2794–2798.
51. Linhares, A. C., Velázquez, F. R., Pérez-Schael, I., Sáez-Llorens, X., Abate, H., Espinoza, F., López, P., Macías-Parra, M., Ortega-Barría, E., Rivera-Medina, D. M., Rivera, L., Pavía-Ruz, N., Nuñez, E., Damaso, S., Ruiz-Palacios, G. M., De Vos, B., O'Ryan, M., Gillard, P., Bouckenooghe, A., and the Human Rotavirus Vaccine Study Group. (2008). Efficacy and safety of an oral live attenuated human rotavirus vaccine against rotavirus gastroenteritis during the first 2 years of life in Latin

American infants: a randomised, double-blind, placebo-controlled phase III study. *Lancet*, 371(9619): 1181–1189.

52. Liu, K., Yang, X., Wu, Y., & Li, J. (2009). Rotavirus strategies to evade host antiviral innate immunity. *Immunology Letters*, *127*(1): 13–8.
53. Ma, H., Shaheduzzaman, S., Willliams, D. K., Gao, Y., & Khan, A. S. (2011). Investigations of porcine circovirus type 1 (PCV1) in vaccine-related and other cell lines. *Vaccine*, *29*(46): 8429–8437.
54. Madhi, S. A., Bch, M. B., Bamford, L., Chb, M. B., Hons, B., Ngcobo, N., & Bch, M. B. (2014). Effectiveness of pneumococcal conjugate vaccine and rotavirus vaccine introduction into the South African public immunisation programme. *South African Medical Journal*, *104*(3): 1–3.
55. Madhi, S. A., Cuncliffe, N. A., Steele, D., Witte, D., Kirsten, M., Louw, C., Ngwira, B., Victor, J. C., Gilliard, P. H., Cheuvart, B. B., Han, H. H., & Neuzil, K. M. (2010). Effect of human rotavirus vaccine on severe diarrhea in African infants. *New England Journal of Medicine. 362*(4): 289–298.
56. Malik, J., Bhan, M. K., & Ray, P. (2008). Natural immunity to rotavirus infection in children. *Indian Journal of Biochemistry & Biophysics*, *45*(4): 219–228.
57. Marlow, R. D., & Finn, A. (2012). The promise of immunisation against rotavirus. *Archives of Disease in Childhood*, *97*(8): 736–740.
58. Martella, V., Bányai, K., Matthijnssens, J., Buonavoglia, C., & Ciarlet, M. (2010). Zoonotic aspects of rotaviruses. *Veterinary Microbiology*, *140*(3-4): 246–255.
59. Matson, D. O. (2006). The pentavalent rotavirus vaccine, RotaTeq. *Seminars in Pediatric Infectious Diseases*, *17*(4): 195–199.
60. Matthijnssens, J., Ciarlet, M., McDonald, S. M., Attoui, H., Bányai, K., Brister, J. R., Buesa, J., Esona, M. D., Estes, M. K., Gentsch, J. R., Iturriza-Go´mara, M., Johne, R., Kirkwood, C. D., Martella, V., Mertens, P. P. C., Nakagomi, O., Parren˜o, V., Rahman, M., Ruggeri, F. M., Saif, L. J., Santos, N., Steyer, A., Taniguchi, K., Patton, J. T., Desselberger, U., Van Ranst, M. (2011). Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). *Archives of Virology*, *156*(8): 1397–1413.
61. Matthijnssens, J., Ciarlet, M., Rahman, M., Attoui, H., Estes, M. K., Gentsch, J. R., Iturriza-gómara, M., Kirkwood, C., Martella, V., Mertens, P. P. C., Nakagomi, O., Patton, J. T., Ruggeri, F. M., Saif, L. J., Santos, N., Teyer, A., Taniguchi, K., Desselberger, U., Van Ranst, M. (2008). *Archive Virology, 153*(8): 1621–1629.
62. Matthijnssens, J., Heylen, E., Zeller, M., Rahman, M., Lemey, P., Van Ranst, M. (2010*a)*. Phylodynamic analyses of rotavirus genotypes G9 and G12 underscore

their potential for swift global spread. *Molecular Biology and Evolution*, *27*(10): 2431–2436.

63. Matthijnssens, J., Joelsson, D. B., Warakomski, D. J., Zhou, T., Mathis, P. K., Van Maanen, M.-H., Ranheim, T. S., Ciarlet, M. (2010*b*). Molecular and biological characterization of the 5 human-bovine rotavirus (WC3)-based reassortant strains of the pentavalent rotavirus vaccine, RotaTeq. *Virology*, *403*(2): 111–127.
64. Midgley, S. E., Bányai, K., Buesa, J., Halaihel, N., Hjulsager, C. K., Jakab, F., Kaplon, J., Larsen, L. E., Monini, M., Poljsak-Prijatelj, M., Pothier, P., Ruggeri, F. M., Steyer, A., Koopmans, M., Bottiger, B. (2012). Diversity and zoonotic potential of rotaviruses in swine and cattle across Europe. *Veterinary Microbiology*, *156*(3-4): 238–245.
65. Moon, S., Wang, Y., Dennehy, P., Simonsen, K. a, Zhang, J., & Jiang, B. (2012). Antigenemia, RNAemia, and innate immunity in children with acute rotavirus diarrhea. *FEMS Immunology and Medical Microbiology*, *64*(3): 382–391.
66. Moon, S.-S., Tate, J. E., Ray, P., Dennehy, P. H., Archary, D., Coutsoudis, A., Bland, R., Newell, M., Glass, R. I., Parashar, U., Jiang, B. (2013). Differential profiles and inhibitory effect on rotavirus vaccines of nonantibody components in breast milk from mothers in developing and developed countries. *The Pediatric Infectious Disease Journal*, *32*(8): 863–870.
67. Murphy, T.V., Gargiullo, P.M., Massoudi, M. S., Nelson, D. B., Jumaan, A. O., Okoro, C. A., Zanardi, L. R., Setia, S., Fari, E., LeBarron, C. W., Wharton, M., & Livengoo, J. R. Rotavirus Intussusception Investigation Team. (2001). Intussusception among infants given an oral rotavirus vaccine. *New England Journal of Medicine*. 344(8): 564–572.
68. Mwenda, J. M., Ntoto, K. M., Abebe, A., Enweronu-Laryea, C., Amina, I., Mchomvu, J., Kisakye, A., Mpabalwani, E. M., Pazvakavambwa, I., Armah, G.E., Seheri, L.M., Kiulia, N. M., Page, N., Widdowson, M., Steele, A.D. (2010). Burden and epidemiology of rotavirus diarrhea in selected African countries: preliminary results from the African Rotavirus Surveillance Network. *The Journal of Infectious Diseases*, *202(*1): S5–S11.
69. O'Neal, C. M., Clements, J. D., Estes, M. K., & Conner, M. E. (1998). Rotavirus 2/6 viruslike particles administered intranasally with cholera toxin, Escherichia coli heat-labile toxin (LT), and LT-R192G induce protection from rotavirus challenge. *Journal of Virology*, *72*(4): 3390–3393.
70. O'Neil, C. M., Crawford, S. E., Estes, M. K., & Conner, M. E. (1997). Rotavirus virus-like particles administered mucosally induce protective immunity. *Journal of Virology.* 71(11): 8707-8717.

71. Parashar, U. D., Bresee, J. S., Widdowson, M.-A., & Gentsch, J. R. (2006). New breath for rotavirus vaccines. *Drug Discovery Today: Therapeutic Strategies*, *3*(2): 159–165.
72. Parez, N. (2008). Rotavirus gastroenteritis: why to back up the development of new vaccines? *Comparative Immunology, Microbiology and Infectious Diseases*. *31*(2-3): 253–269.
73. Patel, M. M., Clark, A. D., Glass, R. I., Greenberg, H., Tate, J., Santosham, M., Sanderson, C. F. B., Steele, D., Cortese, M., Parasher, U. D. (2009). Broadening the age restriction for initiating rotavirus vaccination in regions with high rotavirus mortality: benefits of mortality reduction versus risk of fatal intussusception. *Vaccine*, *27*(22): 2916–2922.
74. Patel, M., Shane, A. L., Parashar, U. D., Jiang, B., Gentsch, J. R., & Glass, R. I. (2009). Oral rotavirus vaccines: how well will they work where they are needed most? *The Journal of Infectious Diseases*, *200*(1): 39–48.
75. Patel, M., Steele, a D., & Parashar, U. D. (2012). Influence of oral polio vaccines on performance of the monovalent and pentavalent rotavirus vaccines. *Vaccine*, *30*(1): 30–35.
76. Patel, N. C., Hertel, P. M., Estes, M. K., De la Morena, M., Petru, A. M., Noroski, L. M., Revell, P. A., Hanson, C., Paul, M. E., Rosenblatt, M. D., Abramson, S. L. (2010). Vaccine-acquired rotavirus in infants with severe combined immunodeficiency. *The New England Journal of Medicine*, *362*(4): 314–319.
77. Patton, J.T. (2008). Segmented Double-stranded RNA viruses: structure and molecular biology. Laboratory of Infectious Diseases, NIAID, NIH, Bethesda, MD 20892-8026, US.
78. Patton, J. T. (1995). Structure and function of the rotavirus RNA-binding proteins. *Journal of General Virology*, *76*: 2633–2644.
79. Payne, D. C., Edwards, K. M., Bowen, M. D., Keckley, E., Peters, J., Esona, M. D., Teel, E. N., Kent, D., Parashar, U. D., & Gentsch, J. R. (2010). Sibling transmission of vaccine-derived rotavirus (RotaTeq) associated with rotavirus gastroenteritis. *Pediatrics*, *125*(2): 438–441.
80. Pesavento, J. B., Crawford, S. E., Estes, M. K., & Prasad, B. V. V. (2006). Rotavirus proteins: structure and assembly. *Current Topics in Microbiology and Immunology*, *309*: 189–219.
81. Potgieter, A. C., Page, N. a, Liebenberg, J., Wright, I. M., Landt, O., & Van Dijk, A. A. (2009). Improved strategies for sequence-independent amplification and sequencing of viral double-stranded RNA genomes. *The Journal of General Virology*, *90(*6):1423–1432.

82. Reyes-Ruiz, J. M., Barrera-Saldaña, H. A. (2006). Proteins in a DNA world: expression systems for their study. *Review Investigation Clinic, 58*(1): 47–55.
83. Ruiz-Palacios, G. M., Pérez-Schael, I., Velaquez. F. R., Abate, H., Breuer, T., Clemens, S. C., Cheuvart, B.,Espinoza, F., Gillard, P., innis, B. L., Vervantes, Y., Linhares, A. C., Lopez, P., Macias-Parra, M., Ortega_barria, E., Richarson, V., Rivera-Medina, D. M., Rivera, L., Salinas, B., Pavia-Ruz, N., Salmeron, J., Ruttimann, R., Tinoco, J. C., Rubio, P., Nunez, E., Guerrero, M. L., Yarzabal, J. P.,Damaso, S., Tornieporth, N., Saez- Llorens, X., Vergara, R. F., Vesikari, T., Bouchenooghe, A., Clemens, R., De Vos, B., O'Brian, M. (2006). Safety and efficacy of an attenuated vaccine against severe rotavirus gastroenteritis. *New England Journal of Medicine. 354:* 11-22.
84. Rippinger, C. M., Patton, J. T., & McDonald, S. M. (2010). Complete genome sequence analysis of candidate human rotavirus vaccine strains RV3 and 116E. *Virology*, *405*(1): 201–213.
85. Rodríguez-Limas, W. a, Tyo, K. E. J., Nielsen, J., Ramírez, O. T., Palomares, L. A. (2011). Molecular and process design for rotavirus-like particle production in Saccharomyces cerevisiae. *Microbial Cell Factories*, *10*(1), 33.
86. Rodríguez-Limas, W. A., Pastor, A. R., Esquivel-Soto, E., Esquivel-Guadarrama, F., Ramírez, O. T., Palomares, L. A. (2014). Immunogenicity and protective efficacy of yeast extracts containing rotavirus-like particles: a potential veterinary vaccine. *Vaccine*, *32*(24): 2794–1798.
87. Sambrook, J., Russell, D. W. (2001). Molecular cloning – A laboratory manual. *ColdSpring Harbour Laboratory Press*. Third edition.
88. Steele, A. D., Neuzil, K. M., Cunliffe, N. a, Madhi, S. a, Bos, P., Ngwira, B., Witte, D., Todd, S., Louw, C., Kirsten, M., Aspinall, S., Van Doorn, L. J., Bouckenooghe, A., Suryakiram, P. V., Han, H. H. (2012). Human rotavirus vaccine Rotarix™ provides protection against diverse circulating rotavirus strains in African infants: a randomized controlled trial. *BMC Infectious Diseases*, *12*(1): 213.
89. Seheri, M., Nemarude, L., Peenze, I., Netshifhefhe, L., Nyaga, M., Ngobeni, H., Maphalala, G., Maake, L. (2014). Update of Rotavirus Strains Circulating in Africa from 2007 through 2011. *The Pediatric Infectious Disease Journal, 33*(1): 76–84.
90. Suguna, K., & Rao, C. D. (2010). Rotavirus nonstructural proteins : a structural perspective. *A special section: Biology and phatogenesis of viruses*, *98*(3): 352-359.
91. Surendran, S. (2008). Review article : Rotavirus infection: Molecular changes and pathophysiology. *EXCLI Journal* (7): 154–162.
92. Tate, J. E., Burton, A. H., Boschi-Pinto, C., Steele, a D., Duque, J., & Parashar, U. D. (2012). 2008 Estimate of Worldwide Rotavirus-Associated Mortality in Children

Younger Than 5 Years Before the Introduction of Universal Rotavirus Vaccination Programmes: a Systematic Review and Meta-Analysis. *The Lancet Infectious Diseases*, *12*(2): 136–141.

93. Tate, J. E., Patel, M. M., Steele, a D., Gentsch, J. R., Payne, D. C., Cortese, M. M., Nakagomi, O., Cunliffe, N. A., Jiang, B., Neuzil, K. M., de Oliveria, L. H., Glass, R. I., Parasher, U. D. (2010). Global impact of rotavirus vaccines. *Expert Review of Vaccines*, *9*(4): 395–407.
94. Tate, J. E., Steele, a D., Bines, J. E., Zuber, P. L. F., & Parashar, U. D. (2012). Research priorities regarding rotavirus vaccine and intussusception: a meeting summary. *Vaccine*, *30 (1*): A179–84.
95. Shadman, R. (2000). The withdrawl of the Rotashield rotavirus vaccination due to an association with intussusception: fact or fiction?. *Vaccine Revolution*, Stanford University: 1–5.
96. Tom-Revzon, C. (2007). Rotavirus live, oral, pentavalent vaccine. *Clinical Therapeutics*, *29*(12): 2724–2737.
97. Trask, S. D., McDonald, S. M., & Patton, J. T. (2012). Structural insights into the coupling of virion assembly and rotavirus replication. *Nature reviews. Microbiology*, *10*(3): 165–77.
98. Vesikari, T, Karvonen, a, Prymula, R., Schuster, V., Tejedor, J. C., Cohen, R., Meurice, F., Han, H. H., Damaso, S., Bouckenoogle, A. (2007). Efficacy of human rotavirus vaccine against rotavirus gastroenteritis during the first 2 years of life in European infants: randomised, double-blind controlled study. *Lancet*, *370*(9601): 1757–1763.
99. Vesikari, Timo, Karvonen, A. V, Majuri, J., Zeng, S., Pang, X., Kohberger, R., Forrest, B. D., Hoshino, Y., Chanock, R. M., Kapikian, A. Z. (2006). Safety, efficacy, and immunogenicity of 2 doses of bovine-human (UK) and rhesus-rhesus-human rotavirus reassortant tetravalent vaccines in Finnish children. *The Journal of Infectious Diseases*, *194*(3): 370–376.
100. Vesikari, T., Matson, D. O., Dennehy, P., Van Damme, P., Santosham, M., Rodriguez, Z., Dallas, M. J., Heyse, J. F., Goveia, M. G., Black, S. B., Shinefield, H. R., Christie, C.D.C., Ylitalo, S., Itzler., R. F., Coia, M. L., Onorato, M. T., Adeyi, B. A., Marshall, G. S., Gothefors, L., Campens, D., Karvonen, A., Watt., J. P., O'Brien, K. L., DiNubile, M. J., Clark, H. F., Boslego, J. W., Offit, P. A., Heaton, P. M. (2006). Safety and efficacy of a pentavalent human-bovine (WC3) reassortant rotavirus vaccine. *The New England Journal of Medicine*. 354(1): 23–33.

101. Victoria, J. G., Wang, C., Jones, M. S., Jaing, C., McLoughlin, K., Gardner, S., & Delwart, E. L. (2010). Viral nucleic acids in live-attenuated vaccines: detection of minority variants and an adventitious virus. *Journal of Virology*, *84*(12): 6033–6040.
102. Wang, L., Huang, J., Nagesha, H. S., Smith, S. C., Phelps, A., Holmes, I., Martyn, J. C., Coloe, P. J., Reeves, P. R. (2006). Bacterial expression of the major antigenic regions of porcine rotavirus VP7 induces a neutralizing immune response in mice. *Vaccine, 17*(1999): 2636–2645.
103. Wang, Y., Azevedo, M., Saif, L. J., Gentsch, J. R., Glass, R. I., & Jiang, B. (2010). Inactivated rotavirus vaccine induces protective immunity in gnotobiotic piglets. *Vaccine*, *28*(33): 5432–5436.
104. Ward, R. L., & McNeal, M. M. (2010). VP6: A candidate rotavirus vaccine. *The Journal of Infectious Diseases*, *202(1):* 101–107.
105. Ward, R. L., Mcneal, M. M., & Steele, A. D. (2008). Why does the world need another rotavirus vaccine ?, *Therapeutics and Clinical Risk Management*, *4*(1): 49–63.
106. Werther, R. L., Crawford, N. W., Boniface, K., Kirkwood, C. D., & Smart, J. M. (2009). Rotavirus vaccine induced diarrhea in a child with severe combined immune deficiency. *The Journal of Allergy and Clinical Immunology*, *124*(3): 600.
107. Zaman, K., Dang, D. A., Victor, J. C., Shin, S., Yunus, M., Dallas, M. J., Podder, G., *Thiem, V.D., Mai,L., Luby, S. P., Tho, L. H., Coia, M.L., Lewis, K., Rivers, S., Sack, D. A., Schödel, F., Steele, A. D., Neuzil, K. M., Ciarlet, M.* (2010). Efficacy of pentavalent rotavirus vaccine against severe rotavirus gastroenteritis in infants in developing countries in Asia: a randomised, double-blind, placebo-controlled trial. *Lancet*, *376*(9741): 615–23.
108. Zhu, J., Yang, Q., Cao, L., Dou, X., Zhao, J., Zhu, W., Ding, F., Bu, R., Suo, S., Ren, Y., Li, G., Ren, X. (2013). Development of porcine rotavirus vp6 protein based ELISA for differentiation of this virus and other viruses. *Virology Journal, 10*(91): 2–9.