Chapter 2
Literature review

2.0 Introduction

Rotaviruses are the most common cause of mild to severe dehydrating gastroenteritis in children under the age of five years, worldwide (Cook et al., 1990, Parashar et al., 2006, Parashar et al., 2009). More than 80% of the approximately 453 000 annual rotavirus related deaths occur in Asia and Africa where access to rehydration therapy and proper medical care is limited (Parashar et al., 2003, Tate et al., 2012). Studies have shown that natural rotavirus infection generally protects against severe disease, although this depends on the time between exposures, the viral strains involved and the immune status of the human host (Anderson and Weber, 2004, Velázquez et al., 1996, Gladstone et al., 2011). Following rotavirus infection, a short incubation period of 1–2 days precedes sudden illness. Symptoms of rotavirus disease include diarrhoea for 4-7 days, vomiting as well as fever, nausea and abdominal cramping. Diagnosis of rotavirus infection is by immunoassays such as ELISA, passive agglutination tests as well as RT-PCR (Flewett and Woode, 1978, Gouvea et al., 1990, Desselberger et al., 2009). There is currently no specific treatment, but the most effective support involves oral, subcutaneous or intravenous rehydration (Desselberger, 1999, Desselberger et al., 2009). Other therapies such as oral immunoglobulin apparently reduce the duration of diarrhoea, but are not routinely used (Desselberger, 1999).

The discovery of rotaviruses probably began with the description of “agents” causing “epidemic diarrhoea” in calves and infant mice (Light and Hodes, 1943, Kraft, 1957). The simian agent 11 (SA11) which was isolated by Dr. Hubert Malherbe in 1958 from the rectum of an overtly healthy Cercopithecus monkey at the National Institute of Virology (Johannesburg, South Africa), is among the first rotavirus strains to be described (Malherbe and Strickland-Cholmley, 1967, Malherbe and Harwin, 1963). The importance of rotaviruses in human health emerged in 1973 when rotavirus particles were associated with severe diarrhoea in infants following the electron microscopic examination of duodenal mucosa of nine children (Bishop et al., 1973,
Davidson et al., 1975). Similar virus particles were subsequently identified in association with human gastroenteritis (Flewett et al., 1973, Flewett and Woode, 1978). Due to the ability of rotavirus SA11 to propagate readily in cell culture, the strain became a prototype for rotavirus biological studies such as virus replication and genome segment function (Estes et al., 1979a, Small et al., 2007, Matthijnssens et al., 2010b). In the middle of the 1970s other rotavirus strains were isolated from human stools. These include the prototype rotavirus DS-1 and Wa strains (Wyatt et al., 1983, Kalica et al., 1981). The rotavirus particle measures approximately 70 nm in diameter and has a wheel-like appearance when viewed under the electron microscope, hence the Latin name “rota” for wheel (Flewett et al., 1974).

2.1 Rotavirus particle structure, genome and genome segment-protein assignment

2.1.1 Genome structure, organisation and encoded protein functions
The rotavirus genome is composed of 11 dsRNA segments that are 667–3302 bp in size (based on rotavirus SA11) (Ramig, 1997, Estes and Cohen, 1989). The dsRNA can be extracted and the genome segments can be completely separated with polyacrylamide gel electrophoresis (PAGE) to produce a unique separation profile that is specific to various strains (Figure 2.1).
Each genome segment encodes a single protein except genome segment 11 which is bicistronic in most strains (Mitchell and Both, 1988), resulting in six structural proteins (VP1–VP4, VP6 and VP7) and six non-structural proteins (NSP1-NSP6; Figure 2.1). The individual genome segment sizes, encoded proteins and functions are indicated in Table 2.1. Each dsRNA genome segment contains a 5'-terminal cap i.e., m^7GpppG^{(m)} (Imai et al., 1983, Pizarro et al., 1991). Flanking the open reading frame (ORF) of each genome segment are 5'- and 3'-untranslated regions (UTRs) (Patton, 1995). The UTRs are of varying sequences and length. At the 5'-end the UTRs are 9–49 bp, and 17–182 bp at the 3'-end (Patton, 1995).
<table>
<thead>
<tr>
<th>GS</th>
<th>Size (bp)</th>
<th>Protein</th>
<th>Size (kDa)</th>
<th>Location</th>
<th>Genotype*</th>
<th>Protein function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3302</td>
<td>VP1</td>
<td>125</td>
<td>Core</td>
<td>R</td>
<td>RNA-dependent RNA polymerase; 3'-mRNA binding; forms transcription complex with VP3</td>
</tr>
<tr>
<td>2</td>
<td>2690</td>
<td>VP2</td>
<td>102</td>
<td>Inner capsid</td>
<td>C</td>
<td>Inner capsid structural protein; non-specific ssRNA and dsRNA binding; myristoylated; required for replicase activity of VP1.</td>
</tr>
<tr>
<td>3</td>
<td>2591</td>
<td>VP3</td>
<td>98</td>
<td>Core</td>
<td>M</td>
<td>Guanylyltransferase and methyltransferase; non-specific ssRNA binding; part of transcription complex with VP1.</td>
</tr>
<tr>
<td>4</td>
<td>2362</td>
<td>VP4</td>
<td>VP4: 88 VP5*: 60 VP8*: 28</td>
<td>Outer capsid</td>
<td>P</td>
<td>Outer capsid spike protein; P-type neutralization antigen; virulence determinant; cell-attachment protein; trypsin cleavage into VP5* and VP8* enhances infectivity; VP5* permeabilises membranes; VP8* contains the heamagglutinin domain (in some strains)</td>
</tr>
<tr>
<td>5</td>
<td>1611</td>
<td>NSP1</td>
<td>59</td>
<td>Cytoplasm</td>
<td>A</td>
<td>Associates with cytoskeleton; high degree of sequence variation; role in suppression of host interferon response</td>
</tr>
<tr>
<td>6</td>
<td>1356</td>
<td>VP6</td>
<td>48</td>
<td>Middle capsid</td>
<td>I</td>
<td>Major virion protein; middle capsid structural protein; homotrimeric structure; group- and subgroup-specific antigen</td>
</tr>
<tr>
<td>7</td>
<td>1105</td>
<td>NSP3</td>
<td>35</td>
<td>Cytoplasm</td>
<td>T</td>
<td>Homodimer; virus-specific 3'-mRNA binding; binds elongation factor eIF4G1 and circularises mRNA on translation initiation complex; involved in translational regulation and host shut-off</td>
</tr>
<tr>
<td>8</td>
<td>1059</td>
<td>NSP2</td>
<td>37</td>
<td>Cytoplasm</td>
<td>N</td>
<td>NTPase and helicase activity; non-specific ssRNA binding; major component of the viroplasm; binds NSP5 and VP1; essential for dsRNA synthesis and formation of infectious viral progeny</td>
</tr>
<tr>
<td>9</td>
<td>1062</td>
<td>VP7</td>
<td>37</td>
<td>Outer capsid</td>
<td>G</td>
<td>Outer capsid structural glycoprotein; G-type neutralization antigen; N-linked high mannose glycosylation and trimming; RER transmembrane calcium-binding</td>
</tr>
<tr>
<td>10</td>
<td>751</td>
<td>NSP4</td>
<td>20</td>
<td>Cytoplasm</td>
<td>E</td>
<td>Viral enterotoxin; receptor for double-layer particle budding through ER membrane; N-linked high mannose glycosylation; modulates intracellular calcium levels essential for viral RNA replication and formation of infectious viral progeny</td>
</tr>
<tr>
<td>11</td>
<td>667</td>
<td>NSP5/6</td>
<td>NSP5: 22 NSP6: 11</td>
<td>Cytoplasm</td>
<td>H</td>
<td>NSP5: viroplasm formation; multimerizes; O-linked glycosylation; hyper-phosphorylated; autokinase activity; enhanced by NSP2 interaction; binds ssRNA; component of viroplasm NSP6: interacts with NSP5 and localises to viroplasms</td>
</tr>
</tbody>
</table>

*Function-dependent genotype annotation (Table 2.4; Matthijnssens et al., 2008b)

GS, genome segment; ds, double-stranded; ss, single-stranded; ER, endoplasmic reticulum; RER, rough ER
Rotaviruses are divided into eight serogroups (A–H) and conserved group-specific terminal end nucleotides are present in all genome segments (Matthijnssens et al., 2012). The conserved 5′-terminal end sequences of rotavirus SA11 are 5′-GGC(A/U)₇- at the 5′-end, and -GACC-3′ at the extreme 3′-terminal end (Imai et al., 1983, Chizhikov and Patton, 2000, Tortorici et al., 2006, Patton, 1995, Wentz et al., 1996a). The 5′-terminal end sequences of other strains such as the new adult diarrhoea rotavirus strain (ADRV-N) starts with 5′-GG- but some genome segments contain variations following the second G (Table 2.2). For instance genome segments 1–4 (VP1–VP4) of ADRV-N start with the sequence 5′-GGCA₉U- while genome segments 5 (NSP1), 8 (NSP3) and 11 (NSP5/6) start with 5′-GGA-, and the initial sequences of genome segment 7 (NSP2) are 5′-GUA- (Jiang et al., 2008). The ADRV-N contains the sequence -ACC-3′ at the 3′-terminal end of genome segments 5 (NSP1), 6 (VP6) and 7 (NSP2) (Yang et al., 2004b). The terminal end sequences of the avian rotavirus HS-58 genome segments are also conserved with the 5′-end sequence 5′-GUUUUAAGU- and -UGUGAC-3′ at the 3′-end (Table 2.2) (Trojnar et al., 2010).

### Table 2.2. Variations between the 5′- and 3′-terminal end sequence of selected rotavirus strains of the different groups.

<table>
<thead>
<tr>
<th>Group</th>
<th>Strain</th>
<th>5′-terminal sequence</th>
<th>3′-terminal sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>SA11</td>
<td>5′-GGC(A/U)₇-</td>
<td>-A/GCC-3′</td>
</tr>
<tr>
<td>B</td>
<td>IDIR</td>
<td>5′-GGCC(U)₉-</td>
<td>-ACCC-3′</td>
</tr>
<tr>
<td>C</td>
<td>Bristol</td>
<td>5′-GGCC(A/U)₉-</td>
<td>-GGCU-3′</td>
</tr>
<tr>
<td>D</td>
<td>HS-58</td>
<td>5′-GGU₉(A)₃-</td>
<td>-GACC-3</td>
</tr>
<tr>
<td>H</td>
<td>ADRV-N</td>
<td>5′-GGCA₉U-</td>
<td>-ACCCC-3′</td>
</tr>
</tbody>
</table>

*Matthijnssens et al., 2012

The dsRNA genome segments are thought to be packed in conical spirals of which the pointed ends converge at the centre of the particle (Figure 2.2) (Pesavento et al., 2001, Prasad et al., 1996). This model of genome organisation is perceived to allow reversible condensation of the genome and independent transcription with minimal interference from the neighbouring enzyme-RNA complexes (Pesavento et al., 2001). However, the different genome segment lengths cannot allow central convergence.
RNA viruses, in general, have an unstable genome due to high mutation rates of $10^{-3}$–$10^{-5}$ errors per nucleotide per replication cycle (Elena et al., 2006). Early studies suggested that the rotavirus genome could be more stable than that of influenza and polioviruses (Flores et al., 1988). However, these studies were mainly based on genome segment 4 (VP4) of rotavirus strains isolated from asymptomatic neonates. The application of molecular methods in the characterisation of rotaviruses has shown that rotavirus genetic diversity is very wide (Iturriza-Gomara et al., 2004, Iturriza-Gomara et al., 2011).

Due to the segmented nature of the rotavirus genome, reassortment (exchange of genome segments between strains) can occur between strains of unrelated genogroups during co-infection (Ward et al., 1990, Ramig, 1997). Reassortment can lead to the generation of novel phenotypes in rotaviruses. In addition, the rotavirus genome is prone to rearrangements of sequences in the 3′-UTR after the ORF. This was shown to occur in cases where patients had chronic rotavirus infection, in immunodeficient children or when rotaviruses were serially cultured at high multiplicity of infection (Hundley et al., 1987, Pedley et al., 1983). Genome segment
11 (NSP5/6) is most prone to rearrangement and genome segments 5 (NSP1), 6 (VP6), 7 (NSP3), 8 (NSP2), 9 (VP7), and 10 (NSP4) are less frequently affected (Schnepf et al., 2008, Gault et al., 2001). The mechanism of genome segment rearrangement is thought to involve re-entry of the 3′-end of the negative strand into the catalytic core (forming a loop) and the RNA-dependent RNA polymerase making a mistake by switching template (Matthijnssens et al., 2006a). Genome segment recombination, where a region from a genome segment from one strain can combine with another region of a different strain to form a single full-length genome segment, has also been reported (Parra et al., 2004, Jere et al., 2011). The phenomenon was shown to occur between two strains of the same genotype and also between strains of a different genotype (Parra et al., 2004, Jere et al., 2011). Reports on the detection of genome segment recombination that generates rotavirus diversity has been rare so far, probably due to limited methods.

2.1.2 Rotavirus particle architecture
A mature rotavirus particle is non-enveloped and composed of three concentric capsid layers (Figure 2.3A) (Ludert et al., 1986, McClain et al., 2010). The inner and middle layers comprise of the structural proteins VP2 and VP6, respectively (Figure 2.3A). The outer layer is composed of VP4 and VP7 (Figure 2.3A and B). The complete virus particle is referred to as a triple-layered particle (TLP). The outer layer can be removed chemically to expose a transcriptionally active double-layered particle (DLP) (Cohen et al., 1979, Cohen and Dobos, 1979, Sandino et al., 1986, Estes et al., 1979b). When both the outer and middle layers are removed, the particle is known as the inner capsid and it is replicase active (Bican et al., 1982). The rotavirus inner capsid is composed 120 copies of VP2 (McClain et al., 2010). Most of the VP2 polypeptide folds into a comma-shape and five monomers converge into a star-like complex (McClain et al., 2010). On the inside, VP2 interacts with the RNA-dependent RNA polymerase (VP1) and methyltransferase (VP3) to form the inner virion (Figure 2.3B). However, the exact position of VP3 inside the VP2 core is not known (Trask et al., 2012c). Externally, VP2 interacts with VP6 which comprises 260 trimeric units and forms the intermediate capsid layer with a T=13 icosahedral symmetry. The principal interaction between VP2 and VP6 was shown to occur via the inward projecting β2-α3 loop of VP6 (which spans amino acid residues 64–72)
and the VP2 pocket formed by the amino acid residues M228, M839, M841, F244, and A220 (McClain et al., 2010, Charpilienne et al., 2002).

Figure 2.3. Rotavirus particle architecture. A, A cut-away model of a rotavirus triple-layered particle. The inner layer is composed of the structural protein VP2 (blue), the middle layer is composed of VP6 (green) and the outer layer is made up of VP4 (red) and VP7 (yellow-orange). B, Schematic illustration of rotavirus structure from cryoEM reconstruction, and the location of the structural protein components. C, Depiction of the interaction between VP4 (red), VP6 (green) and VP7 (yellow). Figure redrawn from McClain et al., 2010, Patton 2012 and with permission from the publishers.
The outer layer is made of 260 trimers of the glycoprotein VP7 and 60 trimers of VP4 (Yeager et al., 1994). VP6 interacts with the arm-like projections of the N-terminal regions of VP7 (Trask et al., 2012b). The arm-like extensions include amino acid residues at positions 51–77 and account for nearly all interactions between VP6 and VP7 (Chen et al., 2009a). The VP4 spikes have ‘head’, ‘body’, ‘stalk’ and ‘foot’ regions that are formed by the two trypsin cleavage products VP5* and VP8* (Dormitzer et al., 2004). The base domain of VP4 (VP5*) interacts with VP6 and three VP7 trimers (Figure 2.1C) which locks VP4 onto the assembled virus particle (Li et al., 2009b, Chen et al., 2009a). Part of VP5* and the whole of VP8* protrude >100Å over the outer capsid surface (Figure 2.1C) (Li et al., 2009b, Yeager et al., 1994). Above the capsid surface, VP4 was visualised as dimeric (2-fold symmetry) but a trimeric appearance is maintained below the outer capsid surface (Li et al., 2009b).

2.2 Classification of rotaviruses

The Rotavirus genus belongs to the Reoviridae family which comprises two subfamilies and a total of 15 genera (Table 2.3). The family is characterised by viruses that contain 9–12 linear segments of dsRNA (Attoui et al., 2011, Mertens, 2004). The two subfamilies are called Sedoreovirinae and Spinareovirinae. Rotavirus is classified under the subfamily Sedoreovirinae which also contains the genera Seadornavirus, Orbivirus, Phytoreovirus, Cardoreovirus and Mimoreovirus. The subfamily Spinareovirinae contains nine genera, i.e., Orthoreovirus, Aquareovirus, Oryzavirus, Fijivirus, Mycoreovirus, Cypovirus, Idnoreovirus, Dinovernaviruses and Coltivirus.
Table 2.3. Classification of dsRNA viruses within the *Reoviridae* family. The genera of interest in this PhD project are highlighted with gray shading and bold text. Data in the table was extracted from Mertens *et al.*, 2004.

<table>
<thead>
<tr>
<th>Subfamily</th>
<th>Genus</th>
<th>Number of genome segments</th>
<th>Particle diameter (genome size)</th>
<th>Selected examples of members of the virus species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinareovirinae</td>
<td>Aquareovirus</td>
<td>11</td>
<td>75 nm (30,500 bp)</td>
<td>Aquareovirus A (Chum Salmon reovirus CS)</td>
</tr>
<tr>
<td></td>
<td>Coltivirus</td>
<td>12</td>
<td>60 – 80 nm (29, 000 bp)</td>
<td>Colorado tick fever virus (strain Florio)</td>
</tr>
<tr>
<td></td>
<td>Cypovirus</td>
<td>10</td>
<td>65 nm (25, 000 bp)</td>
<td>Lymantria dispar cypovirus 1</td>
</tr>
<tr>
<td></td>
<td>Dinovernavirus</td>
<td>9</td>
<td>49 – 50 nm (unknown)</td>
<td>Aedes pseudoscutellaris reovirus</td>
</tr>
<tr>
<td></td>
<td>Fijivirus</td>
<td>10</td>
<td>65 – 70 nm (27, 000 – 30, 000 bp)</td>
<td>Fiji disease virus</td>
</tr>
<tr>
<td></td>
<td>Idnoreovirus</td>
<td>10 – 11</td>
<td>70 nm (27, 000 – 30, 000 bp)</td>
<td>Idnoreovirus 1: Diadromus pulchellus</td>
</tr>
<tr>
<td></td>
<td>Mycoreovirus</td>
<td>11 or 12</td>
<td>80 nm (~23, 000 bp)</td>
<td>mycoreovirus 1: Cryphonectria parasitica mycoreovirus-1</td>
</tr>
<tr>
<td>Orthoreovirus</td>
<td>10</td>
<td>80 nm (~23, 500 bp)</td>
<td></td>
<td>Mammalian orthoreovirus</td>
</tr>
<tr>
<td></td>
<td>Oryzavirus</td>
<td>10</td>
<td>70 nm (26, 000 bp)</td>
<td>Rice ragged stunt virus</td>
</tr>
<tr>
<td>Sedoreovirinae</td>
<td>Cardoreovirus</td>
<td>11</td>
<td>55 nm (Not known)</td>
<td>Eriocheir sinensis reovirus</td>
</tr>
<tr>
<td>Orbivirus</td>
<td>10</td>
<td>80 nm (19, 200 bp)</td>
<td></td>
<td>African horse-sickness virus; bluetongue virus</td>
</tr>
<tr>
<td>Minoreovirus</td>
<td>11</td>
<td>90 nm – 95 nm (25, 400 bp)</td>
<td></td>
<td>Micromonas pusilla reovirus</td>
</tr>
<tr>
<td>Phytoreovirus</td>
<td>12</td>
<td>70 nm (26, 000 bp)</td>
<td></td>
<td>Wound tumor virus</td>
</tr>
<tr>
<td>Rotavirus</td>
<td>11</td>
<td>80 nm (~18, 500 bp)</td>
<td></td>
<td>Rotavirus A</td>
</tr>
<tr>
<td>Seadornavirus</td>
<td>11</td>
<td>60 nm – 70 nm (~21, 000 bp)</td>
<td></td>
<td>Banna virus</td>
</tr>
</tbody>
</table>
Rotaviruses are classified into eight groups, A–H, based on VP6 (McNulty et al., 1984, Kohli et al., 1992, Matthijnssens et al., 2012). A VP6 amino acid sequence identity of 53% has been proposed as a cut-off value for defining rotavirus serogroups (Matthijnssens et al., 2012). Group A rotaviruses are important enteric viruses in humans and animals including birds. Serogroups B and C are also important in human and animal health. Group E rotaviruses were detected in pigs, while groups D, F and G are associated with infections in chickens (Pedley et al., 1986, Trojnar et al., 2010, Devitt and Reynolds, 1993, McNulty et al., 1984, McNulty et al., 1981). Group H rotaviruses have been isolated in humans and pigs (Yang et al., 2004c, Alam et al., 2007, Matthijnssens et al., 2012). Specific VP6 epitopes confer sub-group (SG) specificity on which rotaviruses were previously classified using monoclonal antibodies into SGI, SGII, SG I and II or non SGI and II (Nakagomi et al., 1985, Greenberg et al., 1983). Agarose gel electrophoresis migration patterns of genome segment 11 are used to classify rotaviruses into those with a short electropherotype (821 bp; long genome segment 11) and strains with a long electropherotype (667 bp; short genome segment 11). Historically, rotaviruses were classified using RNA-RNA hybridisation studies into genogroups that are represented by prototype strains (Nakagomi et al., 1989). The rotavirus Wa strain is the prototype of the Wa-like genogroup which contains subgroup II rotaviruses with a long electropherotype (Nakagomi et al., 1989). Strains in the DS-1-like genogroup are represented by the prototype rotavirus DS-1 strain and the strains are classified in subgroup I, with a characteristic short electropherotype (Nakagomi et al., 1989). The rotavirus AU-1 strain is the prototype of AU-1like strains which have a long electropherotype and are placed in subgroup I. The protease-sensitive VP4 and the glycoprotein VP7 induce neutralising antibodies and are used to classify rotaviruses into P (VP4) and G (VP7) serotypes. However, serotyping is limited due to a lack of antibodies against VP4 and VP7 as well as technical difficulties (Falcone et al., 1999). Nucleotide sequences encoding VP4 and VP7 have been used to determine P and G genotypes. The genotypes are nowadays determined using sequence specific RT-PCR (Falcone et al., 1999, Gentsch et al., 1992, Gouvea et al., 1990).

Reassortment of the rotavirus genome (Ramig, 1997) which contributes to rotavirus diversity occurs and participates in the complex genetic make-up and inter-species relationships between rotavirus strains. A full genome-based classification system
was proposed to allow the determination of genetic relationships between rotavirus strains (Matthijnssens et al., 2008a, Matthijnssens et al., 2008b, Matthijnssens and Van Ranst, 2012). The classification strategy compares nucleotide sequences of the open reading frames and the analysis is based on sequence identity cut-off limits. The full genome classification is depicted by the annotation Gx-P[x]-Ix-Rx-Cx-Mx-Ax-Nx-Tx-Ex-Hx which respectively represents the genotype for genome segments encoding VP7-VP4-VP6-VP1-VP2-VP3-NSP1-NSP2-NSP3-NSP4-NSP5/6 (Table 2.4) (Matthijnssens et al., 2008b). The capital letters in the genotype were derived from the function associated with the protein i.e., glycoprotein, protease-sensitive, inner capsid, RNA-dependent RNA polymerase, core and methyltransferase. Similarly, the annotation for the non-structural proteins is based on their function i.e., interferon antagonist, NTPase, translation enhancer, enterotoxin and phosphoprotein (Table 2.4) (Matthijnssens et al., 2008b). Currently, 27 G, 35 P, 16 I, 9 R, 9C, 8 M, 16, A, 9 N, 12 T, 14 E and 11 H genotypes were identified by April 2011 (Matthijnssens et al., 2011, Matthijnssens and Van Ranst, 2012). As more whole-genome sequencing projects are undertaken, it is expected that the genotype numbers may increase.

### Table 2.4. The whole-genome genotype constellation of selected prototype rotavirus strains. Reassortants are visualised by a non-homogenous constellation colour.

![Table 2.4](image)

*The colour scheme is used to enhance the visualisation of certain patterns or genome segment constellations. Green, red, and orange depict the human strains (Hu) Wa-like, DS-1-like, and AU-like genome segments, respectively. Yellow, blue, and purple respectively indicate the avian (Av) PO-13-like rotavirus genome segments; some typical porcine (Po) VP4, VP7, and VP6 genotypes; and the SA-11-like genome segments. Table adapted from Matthijnssens et al., 2008b with permission from the publisher.

Evolution of rotaviruses is partly driven by high mutations rates due to the error-prone RNA dependent RNA polymerase (Flores et al., 1988, Taniguchi and Urasawa, 1995, Chen et al., 1994, Elena et al., 2006, Sanjuán et al., 2010). Inter- or intragenic recombination and rearrangement downstream of the ORF also contribute
towards rotavirus diversity (Pedley et al., 1984, Hundley et al., 1987, Desselberger, 1996, Parra et al., 2004, Jere et al., 2011). The mechanism of rearrangement is thought to involve the re-entry of the 3’-end of the negative strand into the catalytic core (forming a loop) and the RNA dependent RNA polymerase making a mistake by switching template (Matthijnssens et al., 2006a). Due to their ability to readily reassort, various combinations of VP7 (G) and VP4 (P) types have been reported (Maunula and Von Bonsdorff, 2002, Matthijnssens et al., 2008a, Ramig, 1997). Global molecular epidemiological data indicates that G1-G4, G9, P[4] and P[8] rotaviruses are more prevalent overall (Santos and Hoshino, 2005, Santos et al., 2005, Widdowson et al., 2009, Matthijnssens et al., 2008a). G1P[8] rotaviruses are responsible for at least 70% of infections in North America, Europe and Australia. However, these strains only account for 23% and 30% of infections Africa and South America, respectively. A high occurrence of G8 rotaviruses with P[6] types has been reported in Africa (Santos and Hoshino, 2005, Mwenda et al., 2010). Rotavirus strains with G9 and G12 genotypes emerged worldwide in the last decade (Santos and Hoshino, 2005, Matthijnssens et al., 2010a, Page et al., 2010).

The nomenclature of rotavirus strains has been revised recently to include as much information about the strain as possible. The Rotavirus Classification Working Group recommends that strains be named using a notation that states the rotavirus (RV) group, species of origin, country of identification, common name, year of identification, G- and P-type (Matthijnssens et al., 2011). For instance, the simian rotavirus type A commonly known as the rotavirus SA11 strain, isolated in 1958 is named RVA/Simian-tc/ZAF/SA11/1958/G3P[2].

### 2.3 Rotavirus replication

Rotavirus is mainly transmitted via the faecal-oral route. Villous cells of the small intestine are the primary target of infection. The entry of rotavirus into the cells is a multi-step process. While the general steps in the replication cycle are known (Figure 2.4), there are several critical mechanisms and details that are not precisely known.
Figure 2.4. Schematic representation of the rotavirus replication cycle. The red dots indicate some of the mechanisms which are not well established how they occur. These include how DLPs are activated to transcribe mRNA, translocation of mRNA from the cytoplasm into the viroplasm as well as RNA assortment, packaging and replication. Figure from Trask et al., 2012b with permission from the publisher.
2.3.1 Cell attachment and entry

Attachment of rotavirus to cells is mainly through VP4 (Ludert et al., 1996a). Human and animal rotaviruses require sialic acid as the cellular receptor for viral attachment (Ciarlet and Estes, 1999, Haselhorst et al., 2009, Haselhorst et al., 2011). Sequence analysis of VP4 and VP7 showed the presence of integrin ligand motifs which were implicated in rotavirus cell entry (Coulson et al., 1997). Subsequently, it was shown that integrins α2β1 and α4β1 mediate the attachment of rotavirus SA11 to cells (Hewish et al., 2000). However, α2β1 was later found to be not necessary for attachment although it promotes cell entry (Ciarlet et al., 2002). At a post attachment step, the integrin αvβ3 interacts with rotavirus particles to mediate cell entry (Guerrero et al., 2000). This conclusion was based on observations that monoclonal antibody-blocking of αvβ3 in MA104 cells inhibited rotavirus infectivity (Guerrero et al., 2000). In addition to interacting with integrins, rotaviruses were also shown to interact with the heat shock protein 70 (Lopez and Arias, 2004). The low Ca++ levels in the cytoplasm causes the solubilisation of the outer capsid proteins VP4 and VP7 (Ruiz et al., 2005, Ruiz et al., 2000, Ruiz et al., 1996, Ruiz et al., 2007). The solubilisation of VP4 and VP7 leads to the exposure of transcriptionally active DLPs (Charpilienne et al., 2002, Estes et al., 1979a).

Rotavirus infectivity of cells is enhanced by, and probably dependent on, the proteolytic cleavage of VP4 spikes (Estes et al., 1981, Espejo et al., 1981, Arias et al., 1996). Cleavage results in two fragments i.e. VP5* and VP8*, but the two fragments remain in a non-covalent association on the virion (Ludert et al., 1996b, Trask et al., 2012b). Cleavage of VP4 was demonstrated to occur between R230 and N231 (Arias et al., 1996). Variation was reported for the rotavirus SA11 strain in which cleavage sites are present at R241 and R247 (Gorziglia et al., 1986). Trypsin cleavage is important for the conformational change of VP4 required for cell entry (Yoder et al., 2009). The VP4 conformational change involves rearrangement which requires the VP8* heads to dissociate, the trimeric coiled chains zip together and the body folds back to expose hydrophobic regions (Figure 2.5) (Yoder et al., 2009, Dormitzer et al., 2004, Trask et al., 2012b). Mutations in VP5*, especially V139D, demonstrated that rotavirus particles then become 10 000-fold less infectious (Kim et al., 2010a). The amino acid position 139 is located in the hydrophobic apex of VP5.
which is responsible for permeabilisation of the cell membrane (Kim et al., 2010a). Furthermore, the mutagenesis or truncation of amino acids in the VP5* hydrophobic domain (residues 385–404) abolishes membrane permeability (Dowling et al., 2000). This indicates that the hydrophobic apex of VP5*, and its exposure as a result of trypsin cleavage is required for membrane disruption during cell entry. However, the exact mechanism of rotavirus and cell membrane fusion is not understood due to the fact that rotaviruses lack a lipid envelope (Trask et al., 2012b).

![Figure 2.5](image)

**Figure 2.5. Model showing the rearrangement of VP4 during priming and cell entry.** A, Non-cleaved state with possible, flexible, free play indicated by wavy lines. The brown line attached to the purple line represents the non-cleaved trypsin cleavage site. B, Trypsin cleaved state. C, Folded-back state which exposes the hydrophobic regions during membrane penetration. D, Colour coded, linear VP4 indicating the location of amino acid regions in VP5* and VP8*. The head region is coloured in yellow (the VP8* core), notched at the sialoside-binding cleft. The body includes the red domain (VP5* antigenic region) and the orange cap (the potential membrane interaction loop), the purple appendage, and part of the grey and cyan tubes. The stalk is the lower part of the cyan tubes and the foot is shown in green (Dormitzer et al., 2004). The figure was used with permission from the publisher.

### 2.3.2 Rotavirus genome transcription, translation and replication

During the process of viral entry, VP4 and VP7 are shed to result in transcriptionally-active DLPs (Figures 2.4 and 2.5). The loss of VP4 and VP7 is thought to occur due to solubilisation in the low calcium level in the cytoplasm (Dowling et al., 2000, Gilbert et al., 2001). The mechanism which activates transcription following the shedding of VP4 and VP7 is not known, but the low calcium levels are also thought
to activate the viral polymerase (VP1) (Ruiz et al., 2000). Since VP4 interacts with VP6, it may be that its loss may cause a conformational change in VP6 which is relayed to VP2 and finally induce the transcriptase activity of the RNA-dependent RNA polymerase. Transcription activity is dependent on VP6 since the removal of VP6 results in loss of transcriptase activity (Sandino et al., 1986).

The RNA-dependent RNA polymerase uses the minus-strand of dsRNA as a template for the synthesis of capped, non-polyadenylated plus-sense (+) ssRNA (Patton, 1986, Guglielmi et al., 2010). Rotavirus transcripts are detectable 1 hour post infection (hpi) (Patton et al., 2004). The (+) ssRNA transcripts serve a dual role as templates for viral protein translation mRNAs and templates for dsRNA synthesis (Chen et al., 1994). Using RNA interference, it was shown that the amount of transcript during infection was size-dependent (Ayala-Breton et al., 2009). At 12 hpi the amount of genome segment 10 transcript was 4.5 and 2 times more abundant when compared to the amounts of transcript of genome segments 1 and 6, respectively. These observations were consistent with the results reported previously (Stacy-Phipps and Patton, 1987).

NSP3 was shown to specifically bind to the 3'-terminal end of rotavirus transcripts and interacts with the eukaryotic initiation factor 2-α (eIF2-α) (Piron et al., 1998, Piron et al., 1999). The binding of NSP3 to the 3'-terminal end follows the manner in which the poly-A binding protein (PABP) binds to the poly-A tail of eukaryotic transcripts. The binding of NSP3 to the 3'-terminal end is thought to result in circularisation (although not proven) of rotavirus transcripts followed by enhanced translation of rotavirus transcripts (Vende et al., 2000). However, the silencing of the expression of NSP3 with RNA interference did not affect the translation levels of other rotavirus proteins (Montero et al., 2006). In the same study it was found that NSP3 shuts off the synthesis of host cell proteins (Montero et al., 2006). This mechanism probably provides an advantage for rotavirus propagation.

NSP2 and NSP5 localise to form electron-dense, membrane-free, cytoplasmic inclusions (Figure 2.4) known as viroplasms (Fabbretti et al., 1999, Contin et al., 2010). Synthesis of minus-sense strand (-) ssRNA and genome replication occurs in these cytoplasmic inclusions. Although it is known that inner particle is replicase
active (Bican et al., 1982), it is not clear if rotavirus transcripts are replicated into (-) ssRNA in the viroplasms. In addition to NSP2 and NSP5 that are required for viroplasm formation, the viroplasms contain structural proteins VP1, VP2, VP3 and VP6, and the viral (+) ssRNA segments (Eichwald et al., 2004). Furthermore, cellular tubulin was demonstrated to be sequestered into the viroplasm and NSP2 depolymerises the tubulin as a potential mechanism of evading the host-cell’s antiviral mechanisms (Martin et al., 2010). For the (+) ssRNA to be used as dsRNA-synthesis templates, they have to be recruited into the viroplasms. There must be a mechanism for trafficking rotavirus transcripts and viral proteins from the cytoplasm into the membrane-free viroplasms. However, no mechanism for transcript trafficking has been identified and the steps involved in transcript movement are not known. The sequence -UGUGACC-3′ was shown to contain the essential replication signals (Patton et al., 1996). Recently, soaking VP1 apoenzyme crystals with RNA nucleotides representing the consensus sequence 5′-UGUGACC-3′ showed that VP1 specifically recognises -UGUG-3′ in the terminal 3′ end of (+) ssRNA for packaging and dsRNA synthesis (Lu et al., 2008).

Genomic dsRNA can be detected 2–4 hpi (Patton et al., 2004). Equimolar amounts of dsRNA for each genome segment are detectable despite size variations (Ayala-Breton et al., 2009). After 12 hours, the ratio of (+) ssRNA to (-) ssRNA was estimated at approximately 1:1 for segment 1, 2:1 for segment 4, 3:1 for segment 5, and 3:1 for segment 6 (Stacy-Phipps and Patton, 1987). The viroplasm serves as a factory in which dsRNA is synthesised, proteins assemble and package dsRNA into DLPs before addition of VP4 and VP7 in the ER (Figure 2.4). Viroplasm formation is most likely to be an evolutionary mechanism that protects rotavirus dsRNA from being sensed by the cellular innate immune system.

Several aspects of rotavirus replication are not known or there are controversies in literature (Figure 2.4). For instance, the precise order, control and regulation mechanisms of replication are not clearly understood. It is not established whether (+) ssRNA is recruited into the viroplasm alone or in the company of VP1. Specific signals or a recognition mechanism are required to transport (+) ssRNA and proteins into the viroplasm. The source for the transcripts used for dsRNA is yet to be established. It is thought that (+) ssRNA made by DLPs that are in close proximity to
viroplasms could be incorporated into the viroplasm (Silvestri et al., 2004). This idea has been disputed by experiments in which DLPs were labelled and followed in the cytoplasm relative to viroplasm localisation (Carreno-Torres et al., 2010). A suggestion that there is no (+) ssRNA trafficking pathway was indicated by transfection of BrdU-labelled T7-polymerase-derived RNA and determination of the location of RNA and viroplasms (Silvestri et al., 2004). The transfected RNA did not correspond to viroplasms locations suggesting that there was no mechanism of (+) ssRNA trafficking into viroplasms and that templates for dsRNA synthesis were made within the viroplasm (Silvestri et al., 2004). These conclusions seem to be experimentally valid and suggest that some infecting DLPs would have translocated into the viroplasm to synthesise (+) ssRNA to be used as templates for dsRNA synthesis. However, a recent study showed that DLPs labelled in vitro with Cy5 do not co-localise with viroplasms (Carreno-Torres et al., 2010). This observation disputes the conclusion that DLPs associate with viroplasms and produce the (+) ssRNA that is assorted and replicated into dsRNA. However, rotavirus viroplasms co-localise with lipid droplet-associated proteins perilipin A and ADRP, as well as the lipids of lipid droplets (Cheung et al., 2010). This association with lipid droplets is thought to be required for the formation of viroplasms and the generation of infectious viral progeny (Cheung et al., 2010).

2.3.3 Genome packaging and particle morphogenesis
Each rotavirus particle contains equimolar amounts of dsRNA segments (McDonald and Patton, 2011, Stacy-Phipps and Patton, 1987, Patton, 1990). This shows that packaging is a selective and well coordinated process. How the dsRNA for each genome segment is recognised is not known. It is currently thought that cis-acting signals contained in each genome segment may play a role in assortment and packaging (McDonald and Patton, 2011). Assortment was previously perceived to occur in a core-filling manner similar to that of the bacteriophage Φ6 which has a genome comprising of three dsRNA segments (Mindich, 1999). However, the concerted assortment is now considered as the preferred model (McDonald and Patton, 2011).
The preformed \( \Phi 6 \) procapsid packaging of the genome segments is dependent on sequences of approximately 200 nucleotides that are located near the 5'-terminal end (Qiao et al., 1995). The 5'-UTR in rotavirus genome segments are only 9–49 bp and also contain signals for promoting replication (Patton et al., 1996, Mitzel et al., 2003, Patton, 1995). If present, such packaging signals have not been identified and may overlap into the ORF. The genome packaging in \( \Phi 6 \) was also shown to be highly ordered i.e., segment S is packaged alone followed by segments M and L in that order (Qiao et al., 1995). An alternative packaging model for rotaviruses assumes a concerted process in which (+) ssRNA form complexes with VP1 followed by the sequential assembly of VP2 onto the (+) ssRNA-polymerase complexes (McDonald and Patton, 2011). The concerted model of packaging is based on the influenza packaging model (Hutchinson et al., 2010). In this model, it is thought that assortment is mediated by RNA-RNA interactions which occur before genome packaging (McDonald and Patton, 2011). Rotavirus polymerase complexes (PCs) are thought to form first, based on the affinity of VP1 for (+) ssRNA, followed by assortment of the different 11 PCs in the absence of VP2 (McDonald and Patton, 2011). The assembly of VP2 onto the PCs activates VP1 to synthesise the (-) ssRNA strand resulting in a replication intermediate (Patton et al., 1997).

The multi-functional NSP4 accumulates in the endoplasmic reticulum and appears to regulate the addition of the outer capsid onto the DLP (Figure 2.6). The transmembrane intracellular NSP4 (iNSP4) was shown to bind DLPs and VP4 in a chaperone-like manner (Trask et al., 2012b). The silencing of the expression of NSP4 in MA104 cells, using siRNA, reduced rotavirus yields by about 75% (Lopez et al., 2005a). This indicates the critical role NSP4 plays in the addition of the outer capsid (Hu et al., 2012). DLPs formed in the viroplasms bud off to the adjacent endoplasmic reticulum (ER; Figure 2.6). While the mechanism utilised by DLPs to exit viroplasms is not known, it is thought that NSP4 recruits DLPs into an outer capsid assembly pathway (Trask et al., 2012b, Au et al., 1989, Berkova et al., 2006). Through an unknown mechanism and control in the ER, VP4 is thought to be added first followed by VP7 which locks VP4 in place (Figure 2.6) (Trask et al., 2012b). The efficient addition of VP4 and VP7 onto the DLP is facilitated by the higher calcium level in the ER relative to the cytoplasm (Waldron et al., 1994, Ruiz et al., 2000,
Kuum et al., 2012). The rotavirus triple-layered particles are released at the apical surface of polarised cells before cell lysis (Cuadras et al., 2006).

Figure 2.6. A schematic illustration of DLP budding and penetration of the ER during the addition of the outer capsid. Interaction of the DLP with VP4 and NSP4 results in ER membrane deformation and entry into the ER. Here, a temporary envelope containing VP7 is acquired and VP7 assembles onto the particle. The mechanism involving the loss of the envelope is not known. Figure adapted from Trask et al., 2012b with permission from the publisher.

2.4 Pathogenesis of rotavirus infection

Following the process of infection, rotavirus induces epithelial necrosis of the small intestine which leads to villous atrophy (Johnson et al., 1986, Ramig, 2004, Desselberger et al., 2009). Subsequently, enterocyte atrophy results in the loss of intestinal enzymes, low absorption capacity and increased osmotic pressure (Desselberger et al., 2009). In response to damaged intestinal epithelium, crypt cells become hyperplastic which is accompanied by increased fluid secretion (Ramig, 2004). The product of genome segment 10, NSP4, is an enterotoxin which disrupts cellular electrolyte balance by direct stimulation the enteric nervous system cells leading to an increase in intracellular Ca²⁺ levels (Ball et al., 1996, Dong et al., 1997, Ramig, 2004, Lundgren et al., 2000). VP4 is also a major contributor to pathogenicity, but VP3, VP7, NSP1 and NSP2 have also been implicated (Desselberger et al., 2009).
2.5 Immune responses to rotavirus infection

The innate immune response is the first line of defence against virus infection (García-Sastre and Biron, 2006). Part of the innate immune system is composed of pattern recognition receptors (PRRs) that detect viral pathogen associated molecular patterns (PAMPs) (Levy et al., 2011). The PRRs include the cytoplasmic retinoic-acid inducible gene I-like receptors (RLRs), membrane-bound Toll-like receptors (TLRs) and the cytoplasmic nucleotide oligomerisation domain (NOD)-like receptors (Akira et al., 2006, Mogensen, 2009). The RLRs include the retinoic acid-inducible gene I (RIG-I), the melanoma differentiation associated gene 5 (MDA5) and the laboratory of physiology and genetics gene 2 (LPG2) (Saito et al., 2007, Pippig et al., 2009, Loo et al., 2008). Viral PAMPs include replicative intermediates generated during viral replication, especially dsRNA (Thompson and Locarnini, 2007). Rotaviruses induce the secretion of cytokines such as type I interferon (IFN-β) and type III interferon (IFN-λ) (Pott et al., 2011, Deal et al., 2010, Frias et al., 2012, Sen et al., 2011). The specific innate immune response to rotavirus infection depends on the strain and host cell type (Sen et al., 2011, Frias et al., 2012). Through different pathways, the stimulation of the innate immune response results in a cascade of events that ultimately results in the establishment of an anti-viral state (Figure 2.7) (Pott et al., 2012, Pott et al., 2011, Angel et al., 2012). Therefore, rotaviruses have evolved to develop anti-IFN mechanisms that enable rotaviruses to evade the innate immune response by inhibiting the IFN system via NSP1 (Figure 2.7) (Barro and Patton, 2007, Arnold and Patton, 2011, Sen et al., 2009). However, the immune response to rotavirus transcripts, if any, has not been described.
Figure 2.7. Overview of the general innate immune response to rotavirus infection in the cell. Upon infection, rotavirus ligands such as dsRNA stimulate the PRRs RIG-I or MDA5, and PKR, leading to the activation of IFN regulatory factor 3 (IRF3) or NF-κB. Activation of the molecules results in their translocation into the nucleus where they stimulate the production of IFN as well as the induction of interferon stimulated genes (ISG). The viral NSP1 inhibits IRF3 in a strain-dependent manner (Sen et al., 2009), and therefore prevents the formation of an antiviral state. Figure from Angel et al., 2012 with permission from the publisher.

The cellular immune response to rotavirus infection has been studied in several animal models namely, mouse, rat, rabbit and pig (Franco and Greenberg, 1997, Yuan et al., 1998, Conner et al., 1993, Knipping et al., 2011). Adult mice do not develop disease upon infection, but IgA and IgG-producing B lymphocytes are required for viral clearance (Franco and Greenberg, 1999). T-cells were also found to be important in rotavirus clearance (Franco and Greenberg, 1999). However, T-cell deficient mice are able to clear rotavirus infection, but viral clearance is slow (Franco and Greenberg, 1997, Franco and Greenberg, 1995). The early response to rotavirus infection in mice was associated with the presence of activated B-cells and was found to be T-cell independent (Blutt et al., 2002). A recent report indicated that the supplementing of milk with IgY antibodies against the human rotavirus Wa strain raised in chicken passively conferred significant protection rates against the Wa
strain in gnotobiotic piglets in a dose-dependent manner (Vega et al., 2012). The piglets incidentally develop anti-IgY antibodies of the IgG type in serum and local IgA and IgG (Vega et al., 2012). Furthermore, llama-derived anti-VP6 antibodies have been shown to possess broad neutralising activity in vitro and confer protective immunity against diarrhoea in mice (Garaicoechea et al., 2008). However, results obtained from animal models do not always predict what happens in humans.

In humans, rotavirus infection induces systemic and intestinal antibodies that confer protection against severe diarrhoea upon re-infection (Davidson et al., 1983, Ward and Bernstein, 1994, Velázquez et al., 1996). Protection levels varied from 40% protection against any infection, 75% against diarrhoea and 88% against severe diarrhoea (Velázquez et al., 1996, Gladstone et al., 2011). While the exact correlates of protection are not precisely clear, intestinal IgA antibodies against the outer capsid seem to correlate best with protection (Desselberger and Huppertz, 2011, Desselberger et al., 2009, Franco and Greenberg, 1999, Franco et al., 2006). Recently, it has been shown that mice lacking mucosal IgA fail to develop protective immunity against repeat multiple exposures to rotavirus (Blutt et al., 2012). However, rotavirus-specific IgA antibodies are not optimal correlates of protection following vaccination in children in developing countries (Angel et al., 2012). In young children, lymphoproliferative responses to rotavirus occurred simultaneously with increase in antibody titre in approximately 40% of the cases (Makela et al., 2004). The T-cell response in children was found to be weaker and declined shortly after infection while adults respond with a strong T-cell response (Makela et al., 2004). Strong and consistent cellular responses in children occur after several infections (Makela et al., 2004).

2.6 Rotavirus vaccines

The first live oral rotavirus vaccine to be licensed was RotaShield® (Hochwald and Kivela, 1999). The vaccine was a reassortant containing the rhesus rotavirus (RRV) genetic backbone and VP7-encoding genome segments from rotavirus D, DS-1 and ST3 strains. RotaShield® was withdrawn from the market following an unacceptable incidence of 1:10 000 association with intussusception, a form of bowel obstruction (CDC, 1999, Abramson et al., 1999, Murphy et al., 2001). The pathogenic
mechanism associated with intussusception could however, not be found (Lynch et al., 2006). Furthermore, the absence of seasonal intussusception associated with seasonal variation of rotavirus infection suggests a non-significant intussusception-rotavirus association (Rennels et al., 1998, Bines, 2005).

Currently, RotaTeq® (Merck) and RotarixTM (GlaxoSmithKline) are the licensed live attenuated vaccines in use worldwide (Angel et al., 2007). RotarixTM is a monovalent vaccine containing a human G1P[8] vaccine strain (Ruiz-Palacios et al., 2006). RotaTeq® is a pentavalent vaccine composed of VP4 or VP7 from G1–G4 and P[8] strains upon a bovine rotavirus WC3 strain genetic backbone (Vesikari et al., 2006b). The vaccines do not prevent infection but are effective against the development of severe disease.

In large vaccine safety trials, efficacies of 85–98% in the prevention of severe gastroenteritis were observed in the developed European and North American countries (Vesikari et al., 2006a, Ruiz-Palacios et al., 2006). However, in African countries the efficacy has been relatively lower, even when compared to developing countries in Latin America (Ruiz-Palacios et al., 2007). For instance, using RotarixTM, a combined efficacy of 61% was observed in a study conducted in Malawi and South Africa (Madhi et al., 2010). A study conducted in Ghana, Mali and Kenya indicated efficacy levels of just over 60% (Armah et al., 2010). The efficacy differences between the developed and developing countries may be due to differences in gastrointestinal microbial composition, malnutrition in poor countries, the maturity of the immune system and possibly immunosuppression due to infectious pathogens (Gladstone et al., 2011). A description of high vaccine failure in Nicaragua in partially or fully immunised (with RotaTeq®) children due to local wild-type and vaccine strain reassortment, resulting in atypical genome constellations has been reported (Bucardo et al., 2012). This report highlights the need for vaccines that are designed to contain region-specific strains. Furthermore, the rotavirus vaccines are a risk to children born with severe combined immunodeficiency (SCID) which has an incidence between 1/50 000 and 1/100 000 live births (Bakare et al., 2010, Chan and Puck, 2005). Rotavirus vaccines administered to SCID patients induce diarrhoea and persistent virus shedding (Werther et al., 2009, Uygungil et al., 2010, Patel et al., 2010). While high numbers of child deaths can be saved in African countries by
these vaccines, a significant proportion of the population may not be protected effectively. Furthermore, the demand could increase to more than 300 million doses of which only ~20% is currently possible (Atherly, 2010). The cost of rotavirus vaccines is relatively high to be able to adequately meet the demand the poorest countries (Kim et al., 2010b, Madsen et al., 2012). Novel approaches to vaccine design and strain composition may be required for the developing countries in Africa and Asia where the need is greatest.

In what may be a rare event, a RotaTeq® vaccine strain was transmitted from a two-month infant to an older, unvaccinated sibling (Payne et al., 2010). It is thought that reassortment occurred between vaccine strains with the G1P[5] and G1P[8] genotypes to result increased virulence in the reassortant strain (Payne et al., 2010). Considering the lower efficacy levels in African countries and in the event of an outbreak of a virulent strain, rational design and rapid production of vaccines will be required. A rapid response, characterised by a shorter vaccine-strain development time, can only be possible when a functional reverse genetics system is available such as is the case for influenza viruses (Webby et al., 2004).

2.7 Reverse genetics

The engineering of specific mutations into viral genomes and the generation (recovery/rescue) of viable virus from cloned cDNA or mRNA is known as reverse genetics (Figure 2.8). Virus may also be recovered from infectious viral transcripts transfected directly into permissible cells (Figure 2.8). This technology is one of the most powerful tools applied in modern molecular virology.
Figure 2.8. Illustration of reverse genetics strategies for (+) ssRNA viruses. Wild-type viral transcripts can be transfected into cells (A), or reverse transcribed, cloned and the plasmid(s) transfected into cells (B). cDNA can be linearised and used as a template for in vitro transcription followed by transfection of the transcripts into cells (C). Ideally, all these strategies should result in virus recovery. The approaches in B and C allow genome manipulation at cDNA level. The scissor indicates the potential for genome manipulation at cDNA level. Figure modified from (Zhang, 2007).

The first reverse genetics system was the recovery of a λ-phage and SV40 hybrid which was rescued from monkey kidney cells (Goff and Berg, 1976). For RNA viruses, the first reverse genetics system for positive-sense RNA viruses to be developed was that for the poliomyelitis virus (Racaniello and Baltimore, 1981a). In this system, a cDNA copy of whole-genomic RNA was cloned into a pBR322 plasmid. Transfection of the recombinant plasmid into cultured mammalian cells resulted in the recovery of infectious poliovirus (Racaniello and Baltimore, 1981a). Further developments of the poliomyelitis virus reverse genetics system included the use of the SP6 polymerase to generate poliovirus transcripts from cDNA templates (Kaplan et al., 1985). When synthetic transcripts derived from SP6 in vitro transcription were transfected into HeLa 3 cells, infectious poliovirus was recovered (Kaplan et al., 1985). The genomic RNA of the poliomyelitis virus has a positive sense and this may have contributed to the relatively easy achievement of a reverse
The power of reverse genetics in virology is best illustrated using the influenza virus reverse genetics system. The system was pioneered when ribonucleoprotein complexes (RNPs) were formed in vitro by viral genomic RNA in the presence of purified nucleoproteins and polymerase (Luytjes et al., 1989). RNPs and cDNA were transfected into eukaryotic cells followed by infection with helper influenza A virus. The helper virus would incorporate a genome segment from the cDNA for the recovery of a recombinant virus. The current reverse genetics system for influenza viruses is a result of many developments and improvements over the years (Neumann and Kawaoka, 2001, Neumann et al., 1999, Pleschka et al., 1996, Fodor et al., 1999). Initially, influenza virus recovery required 12 plasmids which were subsequently reduced to 8 plasmids (Hoffmann and Webster, 2000). Influenza virus can now be recovered entirely from the transfection of only 5 plasmids (Neumann et al., 2005).

The Spanish influenza pandemic of 1918–1919 tragically killed an estimated 50 million people (Taubenberger et al., 2001). “Resurrection” of the causative influenza virus bearing all 8 genome segments was achieved using reverse genetics (Tumpey et al., 2005a). Under biological safety level 3 conditions, the recovered influenza virus was shown to be lethal in mice, embryonated eggs and that it could propagate without trypsin (Tumpey et al., 2005a). Additional studies were carried out in which the haemaglutinin (HA) and neuraminidase (NA) encoding sequences were transferred into a genetic background of a different strain. The results showed that the recombinant viruses containing the 1918 HA and NA were lethal in mice (Tumpey et al., 2004, Tumpey et al., 2005b). In other studies on influenza virus biology, using reverse genetics, HA and NA encoding sequences in influenza virus were manipulated to show that N-glycans at the tip of HA are potent regulators of virus propagation in cell culture (Wagner et al., 2000). Furthermore, the recombinant H5N1 influenza vaccine, registered by the FDA, is produced using reverse genetics although currently egg-dependent (Satterlee, 2008, Tambyah, 2008). The generation of vaccines using reverse genetics helps to shorten the time between an outbreak and delivery of vaccines than conventional methods. Global distribution is also
expected to be safer if plasmids only and not infectious materials are transported. Nucleotide sequences can be retrieved from GenBank, eliminating the need for transporting biological material. However, not all GenBank nucleotide sequences are reliable.

Nucleotide sequences that are not truly representative of the majority of the virus population could be obtained due to skewed distribution of PCR products and cloning bias (Acinas et al., 2005). RNA viruses exist in a population containing minor variants of the same species (Domingo et al., 2012, Domingo et al., 2006). This is a population of the same virus in an equilibrium of mutational variations in which the consensus genetic material is the most predominant and viable (Domingo et al., 2006). Some of the mutations may be lethal or cause the presence of genetically unfit minor population variants that might interfere with infectivity (González-López et al., 2004). Therefore, some of the sequences in GenBank could potentially lack a genetic fitness if they were sampled from minor variants. Reliable nucleotide sequences can be obtained by determining the consensus sequences of the viral genome using next generation sequencing (Metzker, 2010).

2.7.1 Reverse genetics systems for dsRNA viruses
Viral reverse genetics involves the generation of infectious virus particles in cell culture from cDNA clones or ‘infectious’ (+) ssRNA transcripts. Using reverse genetics, the viral genome can be manipulated with recombinant DNA techniques to introduce directed mutations or generate chimaeric viruses by exchanging coding regions. The ability to engineer recombinant mutant viruses makes it possible to study the biology of the virus and also to generate rationally designed vaccines (Ebihara et al., 2005).

The first demonstration that viable orthoreovirus (reovirus) could be recovered from RNA was performed by transfecting ssRNA and/or dsRNA of the reovirus ST3 strain together with a reticulocyte lysate in which ssRNA or dsRNA had been translated (Roner et al., 1990). Following the transfection, murine fibroblast cells (L929) were infected with a helper reovirus ST2 strain. Reovirus was harvested within 24–48 h and the reovirus ST3 strain formed plaques in 5 days, while the helper reovirus ST2
only formed plaques after 12 days (Roner et al., 1990). However, only 4% of cells were able to produce reovirus in this system. Using this reverse genetics system the incorporation of a CAT gene, provided in cDNA form, into the reovirus genome was reported (Roner and Joklik, 2001).

A plasmid-based reverse genetics system was recently developed for reoviruses (Kobayashi et al., 2007). In this reverse genetics system, the complete set of reovirus genome segments were individually fused at their native 5′- termini to a T7 polymerase promoter and cloned into separate plasmids (Figure 2.9A). The construct contained a hepatitis D virus (HDV) ribozyme (Rib) at the 3′-end to enable the transcription of reovirus mRNAs that contained the exact 3′-end sequences. To recover reovirus, L929 cells were infected with a recombinant vaccinia virus (rDIs-T7pol) which provided the T7 polymerase for transcription and capping of reovirus transcripts. This was followed by transfecting the 10 plasmids and recovery of reovirus after up to 5 days of cell culture. The reverse genetics system was further improved by transfecting plasmids which contained more than one genome segment (Figure 2.9B) (Kobayashi et al., 2010).

**Figure 2.9. Reovirus reverse genetics strategies.** A, representation of reovirus cDNA (black) in a plasmid, flanked by T7 promoter (T7P; yellow) and hepatitis delta virus ribozyme (Rib; white). B, The 10 reovirus plasmids, each encoding a single genome segment, transfected into murine fibroblast L929 cells followed by reovirus recovery. C, An improved strategy in which only four plasmids are used to deliver at least two genome segments each. The reovirus genome segments are coloured black, T7 promoter gray, and the HDV ribozyme white. Figure redrawn from Kobayashi et al., 2007 and Kobayashi et al., 2010 with permission from the publishers.
To date, no reports describing the application of the plasmid-based system have been published yet. This suggests that the system is probably technically difficult to apply. This was recently confirmed by a report that in some instances, reovirus cannot be recovered (Boehme et al., 2011). The authors (Boehme et al., 2011) attribute the failure to recover reovirus to technical problems such as plasmid purity and concentration, transfection efficiency and T7 polymerase activity (Boehme et al., 2011). It is also thought that there may be lethal mutations which prevent virus recovery and this could be remedied by supplying wild type proteins of the mutants during the rescue procedure (Boehme et al., 2011). However, the provision of wild-type proteins does not correct the lethal mutation contained in the genome.

Bluetongue virus (BTV), another member of the Reoviridae family, can be recovered using the complete set of 10 pure viral mRNAs obtained *in vitro* either from transcriptionally active viral cores or using T7-polymerase transcription of viral cDNA (Taniguchi and Komoto, 2012). The first recovery of BTV was from ssRNA derived from core particles (Boyce and Roy, 2007). This showed that BTV ssRNA is sufficient for the recovery of viable BTV. In this system, the addition of dsRNA had no effect on the efficiency of BTV recovery. In the following year, Boyce and co-workers (2008) reported the recovery of reassortants containing genome segments from BTV serotype 1 (BTV-1) and BTV serotype 9 (BTV-9). This was achieved by the transfection of core-derived BTV-1 and BTV-9 ssRNA into BSR cells, a clone of BHK-21 cells. Furthermore, synthetic transcripts, of the entire genome, derived from *in vitro* transcription of cDNA templates with T7 polymerase were used to recover viable BTV (Boyce et al., 2008). It was possible to create reassortants using T7 polymerase derived transcripts. Recently, using the same approach, virulent and avirulent field strains of BTV have been recovered (van Gennip et al., 2012). It was also reported that BTV can be recovered from *in vitro* reconstituted BTV subcores in a cell-free reverse genetics system (Lourenco and Roy, 2011).

Similar to the recovery of BTV from core-derived ssRNA, African horsesickness virus (AHSV) was recently recovered from the transfection of core-derived transcripts (Matsuo et al., 2010). It was also possible to recover reassortants by transfecting ssRNA from two different AHSV serotypes. AHSV recovery was more efficient following two ssRNA transfections that were 18 hours apart, indicating that AHSV
genome replication occurred in two phases (Matsuo et al., 2010). The second transfection is thought to provide transcripts that are replicated and packaged as genomic dsRNA.

### 2.8 Rotavirus reverse genetics

Following the first development of a reverse genetics system developed for orthoreovirus, there were suggestions that the approach could also be easily applied to other Reoviridae family members (Roner and Joklik, 2001). This optimism was further increased by the development of a plasmid-only reverse genetics systems for the other dsRNA viruses in the Reoviridae family described above (Boyce and Roy, 2007, Boyce et al., 2008, Matsuo et al., 2010, Kobayashi et al., 2007). However, rotaviruses are currently refractory to a complete plasmid-based or transcript-based reverse genetics system. Three reverse genetics systems that utilise a helper virus have recently been described for rotaviruses. The three systems allow single genome segment manipulations. The first was developed to manipulate the VP4-encoding genome segment 4 (Figure 2.9) (Komoto et al., 2006). In this system, plasmids containing the full length genome segment 4 of rotavirus SA11 were designed to be flanked by a T7 promoter sequence at the 5'-end and a hepatitis delta virus ribozyme (HDV) and T7 terminator sequences at the 3'-end. COS-7 cells were infected with a recombinant vaccinia virus (rDIs-T7pol) followed by transfecting with the genome segment 4-containing plasmid after 1 hour (Komoto et al., 2006).
Figure 2.10. A plasmid-based reverse genetics system for manipulation of rotavirus genome segment 4 (VP4). COS-7 cells are infected with a T7-expressing vaccinia virus followed by transfection with a plasmid containing a rotavirus genome segment 4 (VP4). The cells are then infected with a helper virus. The helper virus is inhibited using specific antibodies (selection) followed by the recovery of a recombinant virus. Figure from Taniguchi and Komoto (2012), with permission from the publisher.

Twenty hours later, a helper rotavirus KU strain was used to infect the cells. Following another 24 hours, a mixture of two monoclonal neutralising antibodies directed against VP4 of the KU RV strain was used to suppress the KU helper virus and allowing the recovery of the recombinant rotavirus containing genome segment 4 (VP4) from rotavirus SA11 (Figure 2.9). Using this system, Komoto and co-workers recovered recombinant rotavirus with a chimeric VP4 (Komoto et al., 2008). They were able to introduce a DS-1 epitope II into the VP4 of rotavirus SA11 (Komoto et al., 2008). The use of a helper rotavirus and antibodies for selection of recombinants limits the system to manipulation of antibody-accessible viral proteins such as VP4. The helper virus reverse genetics system cannot be applied to non-structural proteins. Furthermore, application of such a system to the other outer capsid protein VP7 has not been reported.

In a second approach, Troupin and co-workers (2010) used genome segment rearrangement in an attempt to overcome some of the limitations inherent in the
helper virus reverse genetics system for rotaviruses, as an alternative system (Troupin et al., 2010). The reverse genetics system is a modification of the one described by Komoto and co-workers (2006). It is based on observations that rearranged genome segments are preferentially packaged into rotavirus particles (Troupin et al., 2010, Hundley et al., 1987, Alam et al., 2008). Recombinant rotavirus expressing rearranged genome segment 7 (NSP3) from in vitro-modified cDNA was obtained (Troupin et al., 2010). The plasmid was designed such that the rearranged rotavirus cDNA encoding NSP3 is flanked by T7 polymerase promoter at the 5′-end, a ribozyme and T7 terminator sequences at the 3′-end. At 16 hours post transfection, COS-7 cells were infected with a recombinant vaccinia virus to provide the T7 polymerase. Two hours later, the cells were infected with a helper bovine rotavirus RF strain. While there was no selection pressure applied, this reverse genetics system requires multiple serial passage to eliminate the wild type rotavirus. Preferential packaging of rearranged genome segments does not, however, confer any growth advantage (Troupin et al., 2011). This may explain the need for multiple serial passages in culture. Application of the rearrangement-dependent reverse genetics is limited to rearranged genome segments. Rearrangement mostly occurs in genome segment 11 and less frequently in genome segments 5 (NSP1), 6 (VP6), 7 (NSP3), 8 (NSP2), 9 (VP7) and 10 (NSP4) (Gault et al., 2001).

The third report on rotavirus reverse genetics involved a single genome segment reverse genetics system that is applied to NSP2. A plasmid containing the complete genome segment 8 cDNA of rotavirus SA11 is transfected into COS-7 cells followed by infection with a temperature-sensitive rotavirus mutant (Figure 2.10) (Trask et al., 2010b). The cDNA is located downstream of a T7 polymerase and upstream of HDV and T7 terminator sequences. To facilitate T7 polymerase transcription, COS-7 cells are transfected with a plasmid expressing the vaccinia virus T7 polymerase (Trask et al., 2010b). This system is also limited in terms of its applicability. The helper virus must have some known temperature sensitivity. However, temperature sensitivity has been mapped in only four rotavirus SA11 and seven RRV strains (Criglar et al., 2011).
In the absence of a true transcript-based or plasmid-only reverse genetics system for rotaviruses, some researchers have resorted to RNA interference (RNAi) for studying rotavirus genome segment functions. For instance, siRNA was shown to effectively silence genome segment 4 and that VP4 may not be required for the budding of DLPs into the endoplasmic reticulum (Dector et al., 2002). In another study, the silencing of genome segment 11 (NSP5/6) using RNAi caused a reduction in the synthesis of all viral (+) ssRNA, proteins and dsRNA (Lopez et al., 2005b). This could be expected due to the important role of NSP5 in viroplasm formation (Contin et al., 2010, Campagna et al., 2005, Arnoldi et al., 2007). However, the
application of RNAi is limited to transient silencing genome segment expression and cannot be used to manipulate the rotavirus genome.

From all the above, it is clear that a complete, robust and user-friendly reverse genetics system is urgently required for total rotavirus genome segment manipulation and better understanding of rotavirus biology. The reverse genetics system could also be used in the development of rationally designed vaccines with regionally relevant rotavirus strains. An understanding of how cells respond to rotavirus RNA introduced by transfection, in addition to knowledge about rotavirus replication, may be strategically important in the development of a rotavirus reverse genetics system.