Specific RNA- and protein-binding characteristics of the nucleoprotein of a South African rabies virus isolate

by

Jeanette A. Jacobs

Submitted in partial fulfillment of the requirements for the degree
of

PHILOSOPHIAE DOCTOR
(MICROBIOLOGY)

in the

Department of Microbiology and Plant Pathology,
Faculty of Natural and Agricultural Sciences,
University of Pretoria,
Pretoria, South Africa

Supervisor: Prof. L.H. Nel
Co-supervisor: Dr. J. Theron

January 2002
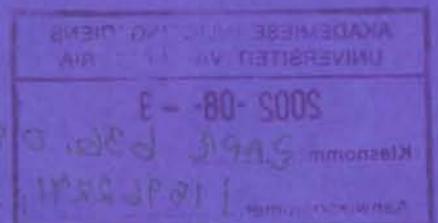
I declare that the thesis, which I hereby submit for the degree Ph.D. at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at another university.

J Jacobs

BEDANKINGS

- **Prof. L. H. Nel.** Ek het as 'n rou voorgraadse student by Prof begin en 6 jaar later het ek my Ph.D. graad voltooi steeds met Prof as my promoter. Baie dankie vir die vertroue wat Prof van die begin af in my geplaas het. Dankie vir die vryheid wat Prof my gegee het om as navorser te ontwikkel. Alle voorspoed met die toekoms.

- **Dr. J. Theron.** Dankie vir die harde werk en intellektuele insette wat jy in hierdie projek ingesit het. Die vertroue wat jy in my gehad het as ek in myself begin twyfel het, en jou aan moedinging in moeilike tye was van onskatbare waarde. Baie dankie!!

- **Marinda Oosthuizen, Bridgitta Steyn, Julian Jaftha, Marietjie Meiring, Jonathan Keytel.** Ons het vir jare saam gewerk en julle het in die proses wonderlike vriende geword. Marinda, baie dankie vir die hulp met GelCompare en die scan van die proteïen gelle, jy was dikwels ongelooflik besig, maar het altyd 'n tydjie afgestaan om te help – dankie.

- **Personeel en studente van die Departement Mikrobiologie en Plantpatologie – Fanus Venter, Volker Brözel, Antoinette, Francois, Boet, Livio, Jacques, Maree, Carinne, Marleen, Jaco, Chris, Sonya, Raynardt.** Ons het saam lang ure gewerk, baie goeie vriende geword. Baie dankie vir julle belangstelling in my projek en die lekker kuiertjies wat ons so tussen die werk deur kon inpas.

- **Daleen, Corrie, Louis en Ronald.** Julle is my steunpilare. Lief vir julle!!

- **My Ouers.** Baie dankie vir julle volgehoue ondersteuning oor die baie jare van student wees. Ek is nou uiteindelik klaar en hoop dat ek met my lewe en loopbaan voorentoe julle naam hoog sal hou. Mag die Here julle seën.

INDEX

CHAPTER 3: Expression of the nucleo- and phosphoproteins in an eukaryotic and prokaryotic expression system.

CHAPTER 4: Role of rabies virus nucleoprotein phosphorylation during interaction with single stranded RNA and the phosphoprotein

CHAPTER 5: Concluding remarks

CHAPTER 6: References

Specific RNA- and protein-binding characteristics of the nucleoprotein of a South African rabies virus isolate

by

Jeanette A. Jacobs

Supervisor: Prof. L.H. Nel
Co-supervisor: Dr. J. Theron
Department: Microbiology and Plant Pathology
Degree: PhD (Microbiology)

SUMMARY

Rabies is a highly fatal nervous disease of humans and all other warm-blooded vertebrates, and is generally transmitted by the bite of diseased animals, most commonly dogs and other carnivores. The aim of this investigation was to investigate the functional role that nucleoprotein phosphorylation might play in the ability of this protein to bind single-stranded RNA and to form complexes with the phosphoprotein (P). Towards achieving these goals, full-length cDNA copies of the N and P genes of a South African viverrid rabies virus isolate were cloned and characterized by nucleotide sequencing. Comparison to cognate rabies virus sequences indicated high levels of homology and a high degree of conservation with regard to functional domains. Analysis of the N and P proteins furthermore indicated the presence of potential consensus recognition sites for both protein kinase C and casein kinase II.

Both the N and P genes were subsequently expressed in the BAC-to-BAC™ baculovirus expression system. Expression of the P protein by the recombinant baculovirus yielded a soluble protein, but the recombinant baculovirus expressed N protein was insoluble. During dual expression of the N and P proteins, which has been reported to aid in N protein solubility by N-P protein complex formation, no complex formation could be shown with the assays used. A contributing factor may have been the low expression

levels for the respective proteins. To overcome this, the N and P genes were consequently expressed in *Escherichia coli* using recombinant pGEX expression vectors. Using this approach, both the N and P proteins were expressed as soluble GST fusion proteins that could be purified by glutathione affinity chromatography to a level of homogeneity.

By making use of a commercially available PKC enzyme an *in vitro* phosphorylation kinase assay was developed whereby the bacterial-expressed unphosphorylated recombinant rabies virus proteins could specifically be phosphorylated. This is the first report of the rabies virus N protein being phosphorylated by PKC as previously identified phosphorylated amino acid residues are located within the consensus recognition motifs correspond to casein kinase II.

The ability of the *in vitro* phosphorylated rabies virus N protein to bind to ssRNA was investigated by poly(U)- and poly(A)-Sepharose binding assays. It was shown that the *in vitro* phosphorylated N protein displayed a higher affinity for ssRNA when compared to the unphosphorylated version of the protein. The phosphorylated rabies virus N protein did not exhibit a preference towards the A-rich artificial ssRNA substrate. To investigate the role of N protein phosphorylation on its ability to bind to P protein, different combinations of phosphorylated and unphosphorylated rabies virus N and P proteins were used in a protein-protein binding assay. The results indicated that unphosphorylated versions of these proteins formed complexes having a 2:1 stoichiometry of N to P protein. *In vitro* phosphorylation of the P protein resulted in the N-P protein complexes with a stoichiometry of 1:2 of N to P protein, whether the N protein independent of the phosphorylation status of the N protein. In the event of unphosphorylated P protein complexed to the phosphorylated N protein, a N:P ratio of 1:1 was found.

COMMUNICATIONS

Congress contributions:

1. J. Theron, J.A. Jacobs and L.H. Nel. The role of phosphorylation in RNA binding of the rabies virus nucleoprotein and in nucleoprotein-phosphoprotein interaction. XIth *Congress of Virology,* Sydney, Australia, August 1999. (Poster).

2. J.A. Jacobs, L.H. Nel and J.Theron. Characterization of the phosphorylation of rabies virus Nucleo- and Phosphoproteins. 17th *Congress of the South African Genetics Society,* Pretoria, South Africa, June 2000. (Paper).

3. J.A. Jacobs, L.H. Nel and J. Theron. Genetic characterization of the Nucleo- and Phosphoprotein of rabies virus. *BioY2K Millennium Congress,* Grahamstown, South Africa, January 2000. (Poster).

4. J.A. Jacobs, L.H. Nel and J. Theron. Analysis of the *in vitro* phosphorylation of the Nucleoprotein and Phosphoprotein of rabies virus. *BioY2K Millennium Congress,* Grahmstown, South Africa, January 2000 (Poster).

5. J.A. Jacobs, J. Theron and L.H. Nel. RNA- and protein-binding characteristics of the phosphorylated Nucleoprotein of rabies virus. *Journal of Virology.* (In preparation).

CHAPTER 1

Literature Review

1.1   Introduction

Rabies (*rabidus* being Latin for mad) is a highly fatal nervous disease of humans and all other warm-blooded vertebrates. The disease is generally transmitted by the bite of diseased animals, most commonly dogs and other carnivores (Swanepoel *et al.*, 1993). The causative agent of the disease is a member of the rod- or bullet-shaped Rhabdoviridae family (*rhabdos* being Greek for rod) which have a single-stranded, negative sense RNA genomes (Krebs *et al.*, 1995). It is divided into 2 main genera, *Vesiculovirus* and *Lyssavirus*.

Rabies virus is the prototype virus of the genus *Lyssavirus*, and has a large geographic distribution in Africa and Europe (King and Crick, 1988; Bourhy *et al.*, 1993). Control of the disease and epidemiological monitoring of the virus have historically been the main focus of rabies virus research. The epidemiology of the disease in South Africa has been studied in some detail over the last decade. Rabies was found to be epizootic in domestic dogs in KwaZulu-Natal, with little wildlife involvement. This is in contrast to the rest of the country where three wildlife host species, *viz.* the yellow mongoose, bat-eared fox and black backed jackal (each associated with a specific geographical region) act as reservoir species for rabies virus (von Teichman *et al.*, 1995). Based on serotypical and genotypical data, South African rabies viruses exist in two groups: the canid viruses that infect domestic dogs, jackals and bat-eared foxes and the viverrid virus group consisting of viruses circulating in viverrid hosts (mongooses and genets).

The genus *Vesiculovirus* is essentially comprised of the prototype vesicular stomatitis virus (VSV). The virus has served as a model for much of the research regarding members of the Rhabdoviridae family. This has been due to the high degree of similarity in genomic organization, as well as transcription and replication cycles between the members of this virus family. For VSV, the viral genes have been cloned and the regulation of gene expression has been extensively studied *in vitro*. These studies have been facilitated by the development of infective cDNA clones, which enabled the controlled manipulation of gene expression.

Despite their similar genomic organization (Abraham and Banerjee, 1976; Flamand and Delagneau, 1978), as well as patterns of RNA and protein synthesis during the infectious cycle

(Coslett *et al.*, 1980; Holloway and Obijeski, 1980), rabies virus and VSV exhibit notable differences in their biological activity.

- Rabies virus requires nuclear functions (Wiktor and Koprowski, 1974), but VSV is capable of infecting enucleated cells (Follet *et al.*, 1974).

- The rabies virus infectious cycle is very slow with viral protein synthesis not observed until 6 h post-infection (Hummeler *et al.*, 1967). In contrast, VSV infection is rapid and highly efficient with progeny virions emerging as early as 4 h post-infection (Simonsen *et al.*, 1979).

- Unlike rabies virus that continues to produce progeny virions for several days post-infection (Hummeler *et al.*, 1967; Matsumoto, 1974), VSV rapidly inhibits host cell macromolecular synthesis, (McGowan and Wagner, 1981; Weck and Wagner, 1978; Wertz and Youngner, 1972) and most of the infected cells are killed at 12 h post-infection (Marcus and Sekellick, 1975; Marcus *et al.*, 1977; Weck and Wagner, 1979).

1.2 Structure of Rhabdoviruses

Rabies virus has a bullet-shaped morphology (Fig.1A) and the virions consist of a nucleocapsid core surrounded by a host-derived lipid envelope. The nucleocapsid core contains all the elements necessary for viral transcription (Kawai, 1977). These include the single-stranded, negative sense RNA genome that is associated with the nucleoprotein (N), the large polymerase protein (L) and the phosphoprotein (P) (Kawai, 1977). In addition to these proteins, two membrane proteins, the matrix protein (M) and glycoprotein (G), are located within the virus envelope (Fig. 1.1B).

1.2.1 The viral genome

Rabies viruses contains a single-stranded, negative sense RNA genome which is 11 932 nucleotides in length, in the case of the Pasteur virus (Tordo and Poch, 1988). Transcription of the rabies virus genomic RNA, as with VSV, produces sequentially one leader RNA (Kurilla *et al.*, 1984) and five polyadenylated monocistronic mRNAs (Coslett *et al.*, 1980; Holloway and Obijeski, 1980).
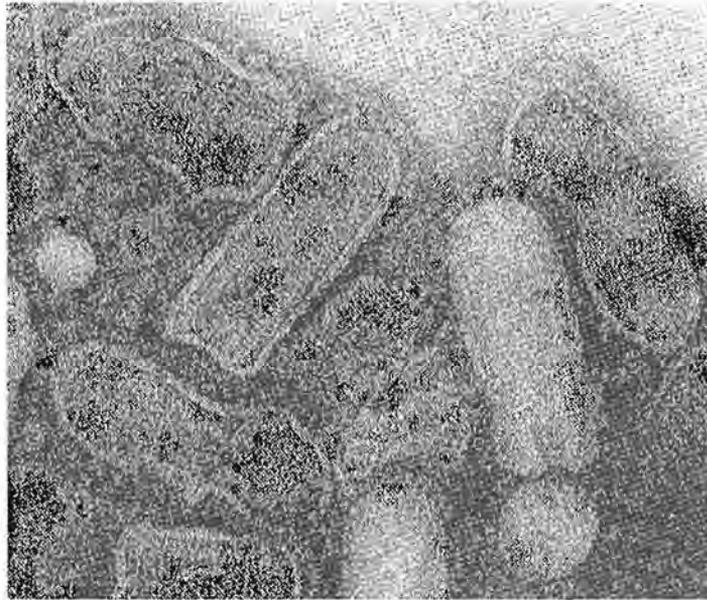
Fig. 1A. Electron micrograph of rhabdovirus particles (VSV) indicating the characteristic bullet-shaped virus particles (Fraenkel-Conrat *et al.*, 1988).

The gene order along the genome, from 3' to 5', is as follows: leader, N, P, M, G and L (Flamand and Delagneau, 1978) (Fig. 2). A conserved intergenic region consisting of the dinucleotide, GA, separates the five protein-encoding genes in the case of VSV (Rose, 1980). In contrast, the four intergenic regions found in the rabies virus genome are of variable length and composition. There are 2 nt separating the nucleoprotein gene from the phosphoprotein gene, 5 nt each separating the phosphoprotein gene from the matrix protein gene and the matrix protein gene and the glycoprotein gene, and 24 to 29 nt separating the glycoprotein gene from the large polymerase gene (Conzelmann *et al.*, 1990, Tordo *et al.*, 1986). These intergenic regions play an important role during transcriptional attenuation at the gene borders, resulting in the mRNA being produced in a sequential gradient (Finke *et al.*, 2000).
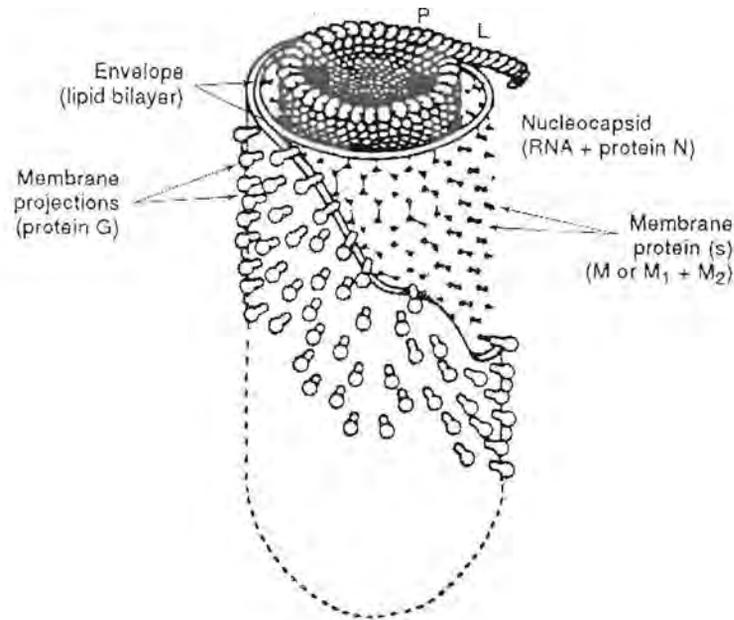
Fig. 1.1B. Schematic representation of the rabies virus virion showing the helical nucleocapsid surrounded by a host-derived lipid bilayer containing the matrix protein and the transmembrane glycoprotein (Wunner, 1991).

The 3' terminus of the VSV genome is 50 to 53 nt in length (Giorgi *et al.*, 1983) and plays a critical role in regulatory steps such as polymerase initiation (Emerson, 1982; Emerson and Yu, 1975; Keene *et al.*, 1978) and N protein encapsidation (Blumber *et al.*, 1983). It also acts as a switch from transcription to replication (Blumberg *et al.*, 1981). The rabies virus leader RNA is 56 nt in length (Kurilla *et al.*, 1984). Blumberg *et al.* (1983) have proposed that the first 14 nucleotides in the VSV leader RNA segment, which contains an A residue on the first and every third base (ACG AAN ACN ANN AAA), constitutes the encapsidation initiation signal. In rabies virus, the encapsidation initiation sequence has been identified as the A-rich sequence spanning nucleotides 20 to 30 in the leader RNA (5'-AAGAAAAAACA-3') (Yang *et al.*, 1998). A similar encapsidation initiation sequence (5'-AAAAATGAGA-3') is present at nucleotides 20 to 30 at the 5' end of the negative-strand genomic RNA (Conzelman *et al.*, 1990; Yang *et al.*, 1998).

```
3' /   N       P      M      G     IGR                    L                      5' /
     ┌─┬────┬──────┬─────┬──────┬─────┬──────────────────────────────────────┬─┐
     │ │1424│ 991  │ 805 │ 1675 │ 423 │               6475                   │ │
     └─┴────┴──────┴─────┴──────┴─────┴──────────────────────────────────────┴─┘

     0    1     2     3     4     5     6     7     8     9    10    11    12
```

Kilobases
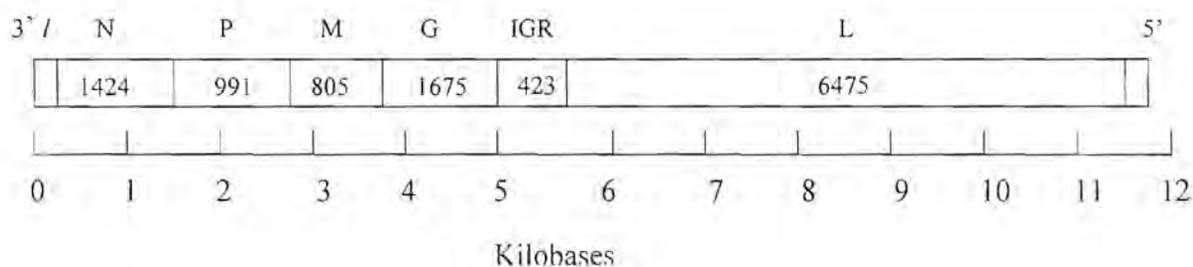
Fig. 1.2. A schematic representation of the single-stranded negative sense RNA genome of rabies virus. The order of each genomic region, in a 3' to 5' direction is: *l* – leader RNA; N – nucleoprotein gene; P – phosphoprotein gene; M – matrix gene; G – glycoprotein gene; IGR – intergenic region; L – large polymerase gene. The length of the respective genes (nucleotides) is indicated in each gene.

1.2.2 Rabies virus proteins

1.2.2.1 Glycoprotein (G)

Primarily as a result of its importance in vaccine-development, the glycoprotein has been extensively studied. The glycoprotein is composed of four distinct domains: the signal peptide that allows the translocation of the polypeptide through the membrane of the endoplasmic reticulum (Tordo and Kouknetzoff, 1993), the ectodomain that is exposed on the outer surface of the virion and includes the glycosylation, palmytolation and antigenic sites (Coulon *et al.*, 1993), the transmembrane peptide that anchors the protein within the viral envelope (Tordo and Kouknetzoff, 1993), and the cytoplasmic domain located in the inner part of the virion (Tordo and Kouknetzoff, 1993). Both the signal peptide and transmembrane peptide are hydrophobic, while the cytoplasmic domain is hydrophilic. The transmembrane glycoprotein stimulates the humoral and cell-mediated immunity against viral infection by not only inducing and binding the virus neutralizing antibodies, but also by stimulation of the T cells (Tordo and Poch, 1988). Furthermore, the G protein plays an important role in the virus-host cell interaction in that it mediates attachment of the virus to the host cells (Tordo and Poch, 1988).

1.2.2.2 Matrix protein (M)

The matrix protein is the most abundant protein present in the virions. The protein plays a central role in viral assembly and release of the viral progeny from infected cells by budding. Cells infected with a rabies virus deficient in M protein were shown to lack the typical bullet-shaped morphology of rhabdovirus particles. Instead it contained long, rod-shaped virions, thus

demonstrating severe impairment of the virus assembly process. Complementation of the M protein-deficient virus by recombinant plasmid rescued virus formation (Mebatsion *et al.*, 1999). In infected cells, 10% of the M protein population is stably associated with the plasma membrane, whereas most of the cellular M protein is found in soluble form in the cytoplasm up to the moment of incorporation into new virions (Barge *et al.*, 1996). The interaction between the G protein and the RNP complex is mediated by a layer of M protein molecules, allowing efficient release of infectious progeny virions (Mebatsion *et al.*, 1999).

1.2.2.3 RNA polymerase protein (L)

The L protein gene is 6 475 nucleotides long (HEP-Flury strain) and encodes for a polypeptide of 2 127 amino acids, with a molecular weight of 242 938 Da (Morimoto *et al.*, 1998). The L protein is a multifunctional enzyme and is the RNA-dependent RNA polymerase. This protein reportedly carries out all of the enzymatic steps required for transcription, including initiation and elongation of transcripts as well as co-transcriptional modifications of RNAs such as capping, methylation and polyadenylation (Banerjee, 1987).

1.2.2.4 Nucleoprotein (N)

The N protein of rabies virus is responsible for the encapsidation of the RNA genome, ensuring the genome's protection against nucleases. The N protein binds to the phosphate-sugar backbone of the RNA and exposes the nucleotide bases to the RNA polymerase during transcription and replication (Emerson, 1987; Iseni *et al.*, 2000; Keene *et al.*, 1981; Klumpp *et al.*, 1997). In addition, the N proteins of all Rhabdoviruses are thought to play a crucial role in transition of the RNA synthetic mode from transcription to replication (Patton *et al.*, 1984; Wertz *et al.*, 1987).

The N protein of rabies virus has been expressed in recombinant baculovirus-infected cells, purified and characterized in great detail (Iseni *et al.*, 1998). The baculovirus-expressed N protein had the same morphological characteristics as viral nucleocapsids when observed with negative stain electron microscopy (EM). The nucleocapsids were 84 Å long; 53 Å high and the spacing between the N protein monomers along the coil were 35 Å. The recombinant nucleocapsids contained cellular RNA and treatment with RNase yielded a 9 nucleotides per N

monomer stoichiometry, similar to that found in viral nucleocapsids. The same group of researchers also studied the rabies virus N-RNA rings by cryo-electron microscopy and calculated three-dimensional (3D) models from single-particle image analysis combined with back projection. It was found that the N protein has a bi-lobed shape, and each monomer has two sites of interaction with each neighbour [(Fig. 1.3) (Schoen *et al.*, 2001)]. Trypsin treatment of the N protein cuts at lysine-376 and removes a 17 kDa fragment at the C-terminal of the N protein (Iseni *et al.*, 1998; Kouznetzoff *et al.*, 1998). The 3-D structure after trypsin digestion indicated that it cuts off part of one of the lobes without shortening the protein or changing other structural parameters (Fig. 1.3). However, this digested protein was unable to bind to the P protein, indicating that the C-terminal part of the N protein is the only binding site for the P protein on the N-RNA.

RNA-binding studies have revealed that the N protein binds 5 to 10 times more leader RNA than to non-leader RNA (Yang *et al.*, 1998). Furthermore, N protein encapsidation of non-leader RNA, but not the leader RNA, was inhibited when the P protein was simultaneously added to the encapsidation reaction, indicating that the P protein helps confer the specificity of leader RNA encapsidation by the N protein (Yang *et al.*, 1998). A hyperconserved motif within rabies virus N proteins starting with the sequence NH_2 S-P-Y-S-S-N (position 298-352) has been identified as being primarily responsible for binding of the leader RNA (Kouznetzoff *et al.*, 1998). However, the conformation of the N protein may also be important in allowing N protein-RNA interaction (Kouznetzoff *et al.*, 1998). This has been supported by the following observations: temperature has an inhibitory effect on the RNA-binding specificity of the N protein, N protein solubilized from purified RNPs or in bacterial extracts following treatment with guaniduim HCl, resulted in a dramatic decrease in RNA-binding activity, and limited structural changes in the NH_2-terminal core had a negative effect on its RNA-binding activity.

50 Å
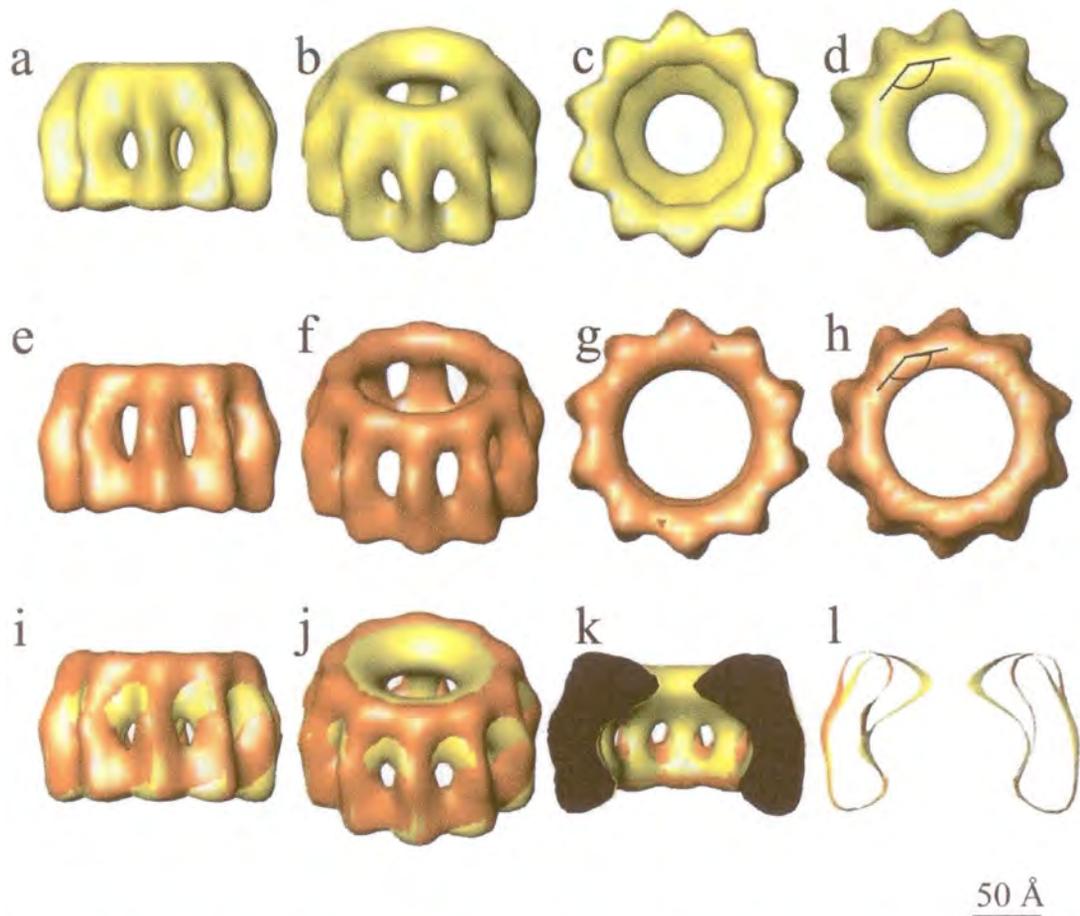
Fig. 1.3. Three-dimensional reconstruction of native (yellow) and trypsin digested (orange) N-RNA rings. The maps are shown in side views (a and e), tilted views (b and f), and end views from the bottom (c and g) and form the top (d and h). The lower panel shows a superposition of the native and digested rings in side view (i), tilted view (j), with the front half cut off (k), and in axial slice (l). The two structures are very similar except at the top end, where the inner part of the circle is missing for the digested rings. The angle between N-N-N shown in panels d and h measures 144° (reproduced from Schoehn *et al.*, 2001).

1.2.2.5. Phosphoprotein (P)

Although the rabies virus P protein is abundant in virus infected cells, it presents only 6% of the virion (Kawai, 1977). Chenik *et al* (1995) reported that the P protein exhibits a large degree of heterogeneity. The authors detected the P protein and at least four additional shorter products (P2, P3, P4 and P5) in infected BSR cells, in BHK cells transfected with a plasmid encoding the wild-type P protein and in purified virions of the CVS rabies virus strain. Mutagenesis of the 5'-terminal region of the P mRNA initiation codon indicated that these shorter proteins were the result of ribosomal leaky scanning (Chenik *et al.*, 1995). In the proposed translation model, ribosomes bind at the capped 5' end of the mRNA and scan downstream to the first AUG codon. Failure to initiate translation at this codon would result in continued scanning downstream for the next available AUG; a second failure allows recognition at a third AUG (Kozak, 1986). Immunofluorescence staining experiments performed on transfected cells have indicated that three of these shorter products (P3, P4 and P5) initiated from, respectively, the third, fourth, and fifth AUG codons. P3, P4 and P5 were mostly located within the nucleus, whereas the wild-type P protein and P2-form were mainly cytoplasmic. Thus, a truncation of at least the N-terminal 52 amino acids of P protein appears to result in the ability to enter the nucleus. A possible explanation is that the N-terminal deletions remove a cytoplasmic retention signal. It is unclear what the function of these truncated P proteins may be, but the presence of large amounts of P2 and P3 forms in purified virions suggests that they are involved in one of the steps of viral multiplication in infected cells (Chenik *et al.*, 1998).

Within non-segmented negative sense RNA viruses, nucleotide sequence analysis of the P protein reveals very little amino acid sequence similarity (Banerjee and Barik, 1992; Lamb and Kolakofsky, 1996). The protein also varies in size, ranging from 27 to 68 kDa between different viruses. Despite these differences, these proteins are all highly acidic and heavily phosphorylated. During viral gene expression the P protein plays a dual role by acting as an essential component of the polymerase (P-L complex), and as a chaperone to deliver soluble N protein (P-N complex) for the encapsidation of RNA during replication (De and Banerjee, 1997). Three functional domains have been identified in the P protein of *Rhabdoviridae*, using VSV as the prototype. Mutational and deletion analyses of the P protein of VSV have indicated at least three functional domains, which play distinct roles during transcription. Domain I is a

highly acidic region that is located at the amino-terminus of the protein (Gill and Banerjee, 1985; Hudson *et al.*, 1986) and must be phosphorylated by cellular casein kinase II (CKII) (Barik and Banerjee, 1992a) at specific sites to serve as a transcriptional activator (Chattopadhyay and Banerjee, 1988; Takacs *et al.*, 1992; Barik and Banerjee, 1992b; Gao and Lenard, 1995a; Chen *et al.*, 1997). VSV specifically packages CKII that remains strongly associated with the ribonucleoprotein complex during morphogenesis (Gupta *et al.*, 1995). Mutational analyses of the phosphorylated amino acids suggested that phosphorylation of this domain may play a critical role during transcription (Pattnaik *et al.*, 1997), but it is not essential for replication (Spadafora *et al.*, 1996; Pattnaik *et al.*, 1997). This view is further supported by the findings that phosphorylation-negative P mutants efficiently form soluble N-P complexes (Takacs *et al.*, 1993; Takacs and Banerjee, 1995) and can initiate the replication process *in vitro* in the presence of cellular factors (Gupta and Banerjee, 1997). This domain also interacts with the L protein to form the active RNA polymerase (Emerson and Schubert, 1987; Takacs and Banerjee, 1995). Domain II is phosphorylated by an L protein-associated kinase (LAK) (Chattopadhyay and Banerjee, 1997b) and is also involved in the binding of P protein to the L protein (Gill *et al.*, 1986; Paul *et al.*, 1988). Mutational analyses of this domain have indicated that phosphorylation is essential for transcriptional activity *in vitro* (Chattopadhyay and Banerjee, 1987b; Gao and Lenard, 1995b). Domain III is a basic and highly conserved 21-amino acid region at the C-terminal of the P protein. This domain is required for binding to the N-RNA template (Gill and Banerjee, 1985; Gill *et al.*, 1986; Emerson and Schubert, 1987) as well as for interaction with soluble N protein (Takacs *et al.*, 1993; Takacs and Banerjee, 1995), indicating its important role in both transcription and replication.

1.3 The processes of viral transcription and replication

From the preceding descriptions, it is apparent that the rabies virus N and P proteins play an important role in the viral transcription and replication processes. Therefore, these processes will be reviewed with reference to the role of the N and P proteins in these processes. Since the discovery of Baltimore *et al.* (1970) that VSV packages the RNA-dependent RNA polymerase that transcribes the genome RNA into messenger RNAs within the mature virion, VSV has became the prototype model for non-segmented negative sense RNA viruses such as members of the Rhabdoviridae and Paramyxoviridae families. While the gene order and overall

mechanism of gene expression in these two virus families are very similar, specific differences have been found and will be discussed in the following sections.

1.3.1 Viral Transcription

The VSV genome is tightly complexed with the nucleoprotein (N) to form the nucleocapsid structure. The active template for transcription and replication of the VSV genome is the nucleocapsid structure, which together with the large polymerase protein (L), and the phosphoprotein (P) serves as the active template for transcription and replication. Both transcription and replication is catalyzed by the viral RNA-dependent RNA polymerase complex, which is composed of the L and P proteins (De and Banerjee, 1985; Emerson and Yu, 1975; Naito and Ishihama, 1976). Biochemical and genetic studies suggest that the L protein is responsible for the enzymatic activities, including the polymerase activity (Hercyk et al., 1988; Hunt et al., 1984; Schubert et al., 1985). Although no enzymatic activities have been associated with the P protein, the phosphorylation status of this protein has been shown to influence the transcriptase and replicase functions of the polymerase protein (Emerson and Schubert, 1987; Pattnaik et al., 1997).

The promoter for VSV genome transcription consists of three separate elements, the 3' leader region, the non-transcribed leader RNA-N gene junction, and the N gene start signal (Whelan and Wertz, 1999a). Additional distinct sequences within the leader region (nucleotides 19 to 29 and 34 to 46) are required for transcription but not replication (Whelan and Wertz, 1999a; Li and Pattnaik, 1999). The RNA polymerase initiates transcription from the 3' terminal end of the genome (Emerson, 1987) and transcribes the genes in a sequential manner (Abraham and Banerjee, 1976; Ball and White, 1976). In doing so, the viral RNA polymerase first synthesizes a small 47-nucleotide long leader RNA that is uncapped and non-polyadenylated and does not encode any viral protein. Following the leader RNA synthesis, five capped and polyadenylated mRNAs, encoding each of the individual viral proteins, are transcribed sequentially in the same order that these genes are found in the virus genome (3'-N-P-M-G-L-5') (Abraham and Banerjee, 1976; Ball and White, 1976).

Schnell *et al.* (1996) have demonstrated that a sequence of 23 conserved nucleotides, which is found at the 3' and 5' junction of all VSV genes, could direct the expression of foreign genes within recombinant VSV genomes. This provided the first evidence that these sequences contained all the *cis*-acting signals necessary to polyadenylate and terminate transcripts from the upstream gene and then to reinitiate transcription at the adjacent downstream gene. The conserved sequence, 3'-AUAC(U)$_7$-5', which is found at the 3' ends of each gene is critical for both polyadenylation and termination of VSV transcription (Barr *et al.*, 1997a, Hwang *et al.*, 1998). When this sequence is encountered, the polymerase stutters over the seven U nucleotides to produce a poly(A) tail approximately 150 nucleotides in length (Schubert *et al.*, 1980). Immediately following the polyadenylation signal, there are two non-transcribed intergenic dinucleotides that may be required for transcription termination (Stillman and Whitt, 1997). However, other studies have suggested that it may function as a spacer element between the transcription termination and re-initiation signals (Barr *et al.*, 1997b; Hwang *et al.*, 1998). Following the intergenic dinucleotide is the sequence 3'-UUGUCnnUAC-5' (with n being any nucleotide). The first three nucleotides of this sequence are the most important for efficient gene expression, as mutations at these positions severely reduce the amount of mRNA levels from the mutated gene. This sequence may be important for re-initiation following polyadenylation and release of the upstream mRNA (Stillman and Whitt, 1997).

There exists a gradient in the molar amounts of the transcribed mRNAs that also follows the gene order from the 3' end of the genome so that the 3'-proximal N gene is transcribed most frequently and the 5'-distal L gene is transcribed least frequently. The gradient in the molar amounts of the mRNAs is believed to be due to attenuation at each of the gene junctions during transcription (Iverson and Rose, 1981). This may be the result of the inability of the polymerase to reinitiate transcription of the downstream gene following transcription termination and polyadenylation of the upstream mRNA.

1.3.2 Replication

1.3.2.1 Switch between transcription and replication

During replication, the transcriptase enzyme switches to the replicative mode and transcribes across the gene boundaries, synthesizing the full-length complement of the negative-strand

genome RNA. The P-L protein complex responsible for transcription is modified to a replicase complex when it reaches the leader-N gene junction (Vidal and Kolakofsky, 1989; Kolakofsky *et al.*, 1991). The modification is presumably triggered by binding of newly synthesized N protein, likely in the form of a soluble P-N protein assembly complex, to nascent leader RNA (Leppert *et al.*, 1979; Blumberg *et al.*, 1981; Peluso and Moyer, 1988). Since N and P genes are the most proximal to the 3' end of the transcribing genome, synthesis of N and P proteins results in immediate complex formation, thus keeping N protein in a soluble and replication-competent form (Bell *et al.*, 1984). The P protein, once complexed with the N protein, is not required to be phosphorylated in Domain I by casein kinase II for its replication function, whereas phosphorylation of the P protein is necessary to interact with the L protein to form the active RNA polymerase complex (Gupta and Banerjee, 1997). As the concentration of N-P protein complex rises in the cell following infection, it acts on the RNP complex to down-regulate transcription and in association with a putative host protein(s) switches transcription to replication (Gupta and Banerjee, 1997). This model is appealing in that all available polymerase activity is channeled toward synthesis of mRNAs until a sufficient amount of the capsid proteins accumulates for replication. The N protein in the complex enwraps the nascent RNA chains of both positive and negative sense genome-length RNAs. As the concentration of N-P protein complex falls below a threshold value, replication ceases and transcription ensues, and this cycle continues during the life cycle of the virus.

1.3.2.2 Replication

During replication, the RNA polymerase first synthesizes the full-length positive sense antigenome, which is enwrapped with the N protein, forming the N-RNA complex. This complex then serves as the template for the synthesis of the negative sense progeny genome RNA (Banerjee, 1987; Banerjee *et al.*, 1977). It is envisaged that the N protein complexes with the nascent leader RNA transcript to initiate encapsidation of the growing RNA chains, leading to replication (Banerjee, 1987; Blumberg *et al.*, 1983; Blumberg *et al.*, 1984; Blumberg *et al.*, 1981; Giorgi *et al.*, 1983).

It is generally believed that the 3'-terminal RNA sequence of the genome RNA is the binding site of the VSV RNA polymerase (Banerjee *et al.*, 1977; Isaac and Keene, 1982; Keene *et al.*,

1981) to initiate transcription. Thus, the 3'-terminal domain of the genome RNA and its complement (leader-sense) RNA are the two important *cis*-acting RNA sequences that are potential targets for cellular proteins to bind and promote the transcriptase and the replicase functions, respectively. It has been shown previously that both positive sense and negative sense leader RNA (the complement of the 3'-terminal sequence of the positive sense genome RNA) interact specifically with the nuclear autoantigen, La protein, in infected cell cytoplasm, raising the possibility that this interaction may have some specific role in the replicative pathway of the virus (Kurilla and Keene, 1983; Kurilla *et al.*, 1982; Wilusz *et al.*, 1983). Moreover, in view of the similarity in the sequences of RNA polymerase III products and the 3' end of the leader RNA, it seems that the interaction of La protein with the leader RNA may be mediated by a sequence motif which regulates VSV transcription and replication. In a separate series of studies, the leader RNA of VSV was implicated in inhibiting cellular RNA synthesis by its temporary localization inside the nucleus following infection (Kurilla *et al.*, 1982; Weck and Wagner, 1978), suggesting that it may interact with specific nuclear proteins involved in RNA synthesis. McGowan *et al.* (1982) tested this by using a soluble cell extract as the source of the RNA polymerase, and reported that purified leader RNA inhibits DNA-dependent transcription of adenovirus and simian virus 40 genes *in vitro*. Gupta *et al.* (1998) subsequently indicated binding of the leader RNA of VSV with hnRNP U and its localization in the cytoplasm. As a result, they speculated that the leader RNA, by its binding to hnRNP U, may be involved in VSV-mediated shut-off of host DNA and RNA metabolism, *via* structural alteration of the leader-N gene junction, enabling the RNA polymerase to read through the junction region and leading to the synthesis of full-length antigenome.

1.4 Aims of this study

From the literature review, it is clear that the nucleoprotein of negative-sense, single stranded RNA viruses is responsible for encapsidating the viral genome, consequently controlling host and viral protein access to the virus genome, and effectively regulating virus transcription and replication processes. Rabies virus differ from the prototype member of the *Rhabdoviridae* virus family, Vesicular stomatitis virus (VSV), in that the nucleoprotein (N) and the phosphoprotein (P) are both phosphorylated, in contrast to the P protein being the only phosphorylated protein of VSV. Despite the insoluble nature of the N protein, the

characterization and functional significance of phosphorylation of the rabies virus N and P proteins were the focus of many published reports in recent years.

The aims of this study can be summarized as follows:

1. Cloning and characterization of the N and P proteins of a South African rabies virus isolate.
2. Expressing both the N and P proteins, singularly and in combination in a BAC-to-BAC™ baculovirus expression system.
3. Expressing the N and the P proteins in *Escherichia coli* using a pGEX expression system.
4. Developing an *in vitro* phosphorylation strategy.
5. Investigating the influence of phosphorylation of the N protein on its ssRNA binding ability, as well as characterization of the specificity of N protein binding to ssRNA.
6. Characterize the influence of phosphorylation on complex formation between the N and P proteins.

CHAPTER 2

Genetic characterization of the nucleoprotein and phosphoprotein genes and encoded gene products of a South African viverrid rabies virus strain

2.1 Introduction

Many studies of viruses in the *Lyssavirus* genus, of which rabies virus is the prototype virus, have focused on the nucleo- and glycoproteins. This has been due to their high antigenicity, which has led to the production of monoclonal antibodies for use in of rabies virus strain differentiation. The phosphoprotein has traditionally received less attention. However, this situation has changed when the important role of the phosphoprotein, in combination with the nucleoprotein, during the process of viral transcription and replication was realized.

The rabies virus nucleoprotein is encoded by the N gene and is 1 424 nucleotides in length (Pasteur virus) (Tordo *et al.*, 1986b). The nucleoprotein plays an important role during virus transcription and replication. Not only is it capable of binding to the viral RNA genome thereby protecting it from nuclease digestion (Kouznetzoff *et al.*, 1998), but it also forms complexes with the phosphoprotein (Schoenh *et al.*, 2001), which plays an important regulatory role during virus propagation. Comparative sequence analyses between different rabies virus strains indicated high levels of homology, e.g. 97.5% between the N genes of Pasteur virus (PV) and CVS; and 99.1% between the N genes of PV and SAD-B19 (Tordo *et al.*, 1986a; Conzelmann *et al.*, 1990). This high level of sequence conservation may point to the importance of maintaining the nucleocapsid structure, which is believed to have been preserved during evolution of the Rhabdoviruses (Bourhy *et al.*, 1989).

The phosphoprotein-encoding gene is 900 nucleotides in length and is highly conserved between different rabies virus strains, displaying 99% homology between the P genes of the Pasteur virus, ERA and PM strains (Larson and Wunner, 1990). Differences between P protein amino acid sequences tend to cluster to specific regions (Nadin-Davis *et al.*, 1997). The N-terminal 50 amino acids is well conserved in all rabies virus strains, the next 35 amino acids (residues 51 to 85) is more variable, while a second variable region is located at position 130 to 180. The P protein is complexed with the N protein as well as the large polymerase protein and thus forms part of the nucleocapsid core (De and Banerjee, 1997). The P protein has been shown to bind to the L protein (Chenik *et al.*,

1998) thereby forming the active RNA polymerase in which the P protein serves as a non-catalytic co-factor (Emerson and Schubert, 1987).

Both the rabies virus N and P proteins are phosphorylated. Phosphorylation of the rabies virus N protein is one of the features that distinguishes it from the cognate VSV N protein. In contrast to the N protein that is phosphorylated at serine and threonine amino acid residues (Anzai *et al.*, 1997; Yang *et al.*, 1999), the P protein is phosphorylated only at serine amino acid residues (Gupta *et al.*, 2000). Although the kinases responsible for rabies virus P protein phosphorylation has been identified as protein kinase C (PKC) and rabies virus protein kinase (RVPK) (Gupta *et al.*, 2000), no information regarding the kinase(s) responsible for phosphorylation of the N protein is available. Phosphorylation of the N protein has been reported to result in less leader RNA being encapsidated and in up-regulation of viral transcription and replication processes (Yang et al., 1999). Phosphorylation of the P protein may play a role in determining the specific host range of the rabies virus (Prehaud *et al.*, 1992; Gupta *et al.*, 2000).

The objectives of this part of the study were to clone and sequence the full-length nucleoprotein and phosphoprotein genes of a South African viverrid rabies virus, and to compare the obtained sequences to the homologous genes and encoded proteins of other rabies virus strains in order to identify conserved features which may be related to the function of the respective proteins.

2.2 Materials and Methods

2.2.1 Monoclonal antibody typing of Rabies virus isolates

Viverrid rabies isolate m710/90 was obtained from the Rabies Unit of the Veterinary Institute at Onderstepoort. This virus was isolated from a rabid *Cynictis Penicillata* (Yellow mongoose) in 1990 and characterized as follows: monoclonal antibody (Mab) typing was performed using a Mab-N panel consisting of 2 Mab-Ns from the Wistar Institute, Philadelphia, and 4 Mab-Ns prepared at the Central Veterinary Laboratory (CVL), Weybridge. Wistar Institute Mab-Ns were used at the dilutions recommended by the donors and the Weybridge Mab-Ns were used at a concentration of four-fold their fluorescence end-point, as previously described (King *et al.*, 1993). Sequence analysis has also found this rabies virus isolate (m710/90) to belong to the viverrid biotype (Olivier, 1997). A 10% suspension of original brain material was inoculated into 21-day-old suckling mice and following virus multiplication, the rabies virus isolate was stored in a 20% lyophilized mouse brain suspension.

2.2.2 Design of nucleotide primers for amplification of the Nucleoprotein (N) and the Phosphoprotein (P) genes

To amplify the full-length N and P genes, oligonucleotide primers were designed based on nucleotide sequence alignments of the 5' and 3' terminal regions of the corresponding genes from various rabies virus strains. The ClustalX multiple sequence alignment program was used for this analysis (Thompson *et al.*, 1997). In order to facilitate cloning of the PCR-amplified genes, unique restriction endonuclease recognition sites were also included in the primers (Table 2.1).

2.2.3 Total RNA extraction

Total RNA was extracted using TRI REAGENT® (Molecular Research Center, Inc.) according to the manufacturer's instructions. Briefly, *ca.* 50 – 100 mg of lyophilized mouse brain material was homogenized in 1 ml of TRI REAGENT®. Following addition of 200 µl of chloroform, the homogenate was separated into aqueous and organic phases by centrifugation at 12 000 rpm for 15 min at 4°C. The RNA was precipitated from the aqueous phase by addition of 500 µl of isopropanol. Following centrifugation, as above,

the pellet was washed twice with 75% ethanol and then suspended in diethyl pyrocarbonate (DEPC) H_2O.

Table 2.1   Primers used to amplify and sequence the nucleoprotein (N) and phosphoprotein (P) genes.

Gene	Primer	Use	Sequence (5' – 3')	Reference
N gene	JJ1	cDNA, PCR, sequencing	<u>ggatcc</u>tacaatggatgccgac	This study
	N_3	PCR, sequencing	**gaattc**gatgtctggcgtcttgcc	This study
	N_2	Sequencing	cccatatagcatcctac	Sacramento *et al.*, 1991
	N_1	Sequencing	tttgagacagccccttttg	Sacramento *et al.*, 1991
	NucRev	Sequencing	gacagttccgtcatctgc	Nel *et al.*, 1997
P gene	PHF	cDNA, PCR, sequencing	tat**gaattc**atcccaaacatgagcaa	This study
	PHR	PCR, sequencing	tta**gaattc**ggttagcaagatgtatagc	This study

Oligonucleotide sequences in **Bold** indicate EcoR1 restriction enzyme recognition sites, and the <u>Underlined</u> sequence indicates a BamH1 restriction enzyme recognition site.

2.2.4   Reverse transcription

For first strand cDNA synthesis of the N and P genes, total RNA (*ca.* 0.5 to 4.0 μg) was annealed with 12 pMol of the positive sense primers, N_3 and PHF (Table 2.1) respectively, at 65°C for 5 min. The reaction mixtures were cooled on ice and the RNA was reverse-transcribed at 37°C for 90 min in a 10 μl reaction mixture containing 50 U of Moloney Murine Leukemia Virus reverse transcriptase (M-MuLV) (USB^{TM}), 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 1 mM of each deoxynucleotide triphosphate (dNTP) and 5 U of RNasin ribonuclease inhibitor

(Promega). After completion of the reaction, the reaction mixtures were each diluted to a final volume of 50 µl with UHQ H₂O.

2.2.5 Polymerase chain reaction

The reaction mixtures (100 µl) contained 10 µl of diluted first strand cDNA, 100 µM of each dNTP, 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 0.1% (v/v) Triton X-100 (supplied in *Taq* 10 × buffer) and 0.25 U of Taq DNA polymerase (Promega). For amplification of the N gene, 10 pMol of the JJ1 and N_3 primers were used, while for amplification of the P gene, 10 pMol of the PHF and PHR primers were used. The tubes were placed in a GeneAmp thermocycler (Model 2400; PE Applied Biosystems) for 30 cycles using the following temperature profile: denaturation at 94°C for 45 s, annealing at 50.5°C for 45 s, and elongation at 72°C for 90 s. The "hot start" protocol, with initial denaturation at 95°C for 5 min before addition of the enzyme, was carried out as described by Erlich *et al.* (1991).

2.2.6 Agarose gel electrophoresis

The PCR-amplified products were analyzed by agarose gel electrophoresis (Sambrook *et al.*, 1989). For this purpose, horizontal 0.8% (w/v) agarose gel slab gels were cast and electrophoresed at 100 V in 1 × TAE buffer (40 mM Tris-HCl, 20 mM Na.acetate, 1 mM EDTA; pH 8.5) using a Biorad Wide Mini Sub™ electrophoresis cell. The agarose gels were stained with ethidium bromide (0.5 µg/ml) and the DNA fragments were visualized by UV fluorescence. The DNA fragments were sized according to their migration in the gel as compared to that of a standard DNA molecular marker (Lambda DNA cut with EcoR1 and HindIII; Roche).

2.2.7 Purification of PCR products

The amplicons were purified from the agarose gel using a Geneclean™ kit (Bio 101 Inc.) according to the manufacturer's instructions. Briefly, the amplicons were excised from the agarose gel and mixed with *ca.* three volumes of a 3 M NaI solution. The agarose was dissolved at 55°C after which 5 µl of glassmilk® was added to the suspension. After incubation on ice for 5 min, the silica-bound DNA was pelleted by brief centrifugation,

and washed three times with ice-cold NEW Wash (a solution containing NaCl, Tris and EDTA in ethanol and water). The DNA was eluted from the silica matrix at 55°C for 2 - 3 min in a final volume of 12 µl of UHQ H₂O. A small aliquot of the eluate (1 µl) was electrophoresed on an agarose gel in order to determine the concentration of the DNA.

2.2.8 Cloning of the amplicons into the pGEM® T-Easy vector (Promega)

- *Ligation of DNA fragments*

Approximately 300 ng of purified amplicon and 50 ng of vector were ligated with 3 U of T4 DNA ligase and 1 µl of a 10 × ligation buffer (660 mM Tris-HCl, 10 mM DTT, 50 mM MgCl₂, 10 mM ATP; pH 7.5) at 4°C for 16 h in a final volume of 10 µl.

- *Preparation of competent cells*

The CaCl₂ method, as described by Sambrook *et al.* (1989), was used for the preparation of competent *E. coli* cells. A 1 ml overnight culture of *E. coli* JM109 cells was inoculated into 100 ml of preheated (37°C) sterile LB-broth, incubated at 37°C with agitation until an absorbency at 550 nm of 0.4 to 0.5 and then cooled on ice for 20 min. The cells from 30 ml of the culture were pelleted in Corex tubes by centrifugation at 4 000 rpm for 10 min at 4°C in a Sorvall® HB 4 rotor (Du Pont Instruments). The cells were suspended in 10 ml ice-cold 50 mM CaCl₂, incubated on ice for 1 h, pelleted as before and finally resuspended in 1 ml of the CaCl₂ solution. The competent cells were incubated on ice for at least 1 h before being used in transformation experiments.

- *Transformation of competent E. coli JM109 cells*

Competent cells were transformed using the heat-shock method as described by Sambrook *et al.* (1989). Competent cells (200 µl) were mixed with 5 µl of the respective ligation mixtures in sterile transformation tubes, and incubated on ice for 30 min. The cells were then incubated at 42°C for 90 s and rapidly chilled on ice for 2 min, after which 800 µl of pre-warmed LB-broth was added. After incubation at 37°C for 1 h with agitation, the recovered cells were plated in aliquots of 100 to 200 µl onto the surface of LB-agar plates supplemented with 100 µg/ml ampicillin. To allow for blue-white colour

selection, based on insertional inactivation of the *lac* Z gene, 40 µl of X-gal (2% (v/v) stock solution) and 10 µl of IPTG (100 mM stock solution) were also spread over the surface of the agar plates. The plates were then incubated overnight at 37°C and observed for the presence of recombinant transformants with a Gal⁻ phenotype. These were selected for further characterization and grown overnight in 3 ml of LB-broth, supplemented with ampicillin.

2.2.9 Plasmid DNA extraction

Recombinant plasmids were isolated using the alkaline lysis method as described by Birnboim and Doly (1979). The cells from 1.5 ml of the overnight cultures were harvested by centrifugation at 15 000 rpm for 1 min in a microfuge tube using a benchtop microfuge. The bacterial pellets were suspended in 100 µl of Solution 1 (50 mM glucose, 25 mM Tris-HCl (pH 8.0), 10 mM EDTA) and incubated on ice for 5 min. The resultant spheroplasts were lysed following addition of 200 µl freshly prepared Solution 2 (1% SDS, 0.2 M NaOH). The mixtures were neutralized with the addition of 150 µl of Solution 3 (3 M sodium acetate), resulting in the precipitation of chromosomal DNA, high molecular weight RNA and proteins. After incubation on ice for 15 min, the precipitate was removed by centrifugation at 15 000 rpm for 5 min. The plasmid DNA was precipitated from the recovered supernatants by the addition of 2.5 volumes of ethanol. After incubation at -20°C for 1 h, the DNA was pelleted by centrifugation and washed twice with 70% ethanol. The vacuum-dried DNA pellets were resuspended in 50 µl 1 × TE buffer (10 mM Tris, 1 mM EDTA; pH 7.4).

2.2.10 Purification of recombinant plasmid DNA

Plasmid DNA (50 µl) was incubated with 1 µl RNase A (10 ng/µl) at 37°C for 30 min and then precipitated with 30 µl of a PEG/NaCl solution (20% (w/v) PEG 6000, 2.5 M NaCl). After incubation on ice for 30 min, the DNA pellet was obtained by centrifugation at 15 000 rpm for 20 min, washed twice with 70% ethanol, vacuum-dried and resuspended in 25 µl 1 × TE. This method yielded highly purified plasmid DNA and

strongly selects for the super-coiled covalently closed circular (ccc) form of plasmid DNA.

2.2.11 Restriction endonuclease digestion

Purified plasmid DNA (1 µg) was typically digested with 5 U of restriction enzyme (Roche) in the appropriate concentration salt (using the 10 × buffer supplied by the manufacturer). Reaction volumes were typically 15 µl, and incubation was carried out at 37°C for 90 min, after which the reaction was terminated by inactivation of the enzyme at 65°C for 3 min. Following restriction of the plasmid DNA, the restriction fragments were analyzed by agarose gel electrophoresis as described in Section 2.2.6.

2.2.12 Nucleotide sequencing

Sequencing of the cloned insert DNA was performed using an ABI PRISM™ Big Dye™ Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems, Perkin-Elmer). Each reaction mixture contained 50 – 100 ng of PEG/NaCl-purified template DNA, 3.2 pMol primer, 2 µl Terminator Ready Reaction Mix, 2 µl of 5 × Buffer and UHQ water in a final reaction volume of 10 µl. Cycle sequencing was performed using the following program for 25 cycles: denaturation at 96°C for 30 s, annealing at 50°C for 15 s and extension at 60°C for 4 min. Following brief centrifugation, the extension products were precipitated by the addition of 8 µl UHQ water and 32 µl 95% ethanol (final concentration of 60%). The tubes were incubated at room temperature for 15 min, centrifuged at 15 000 rpm for 15 min and the supernatants carefully aspirated. The pellets were washed twice with 50 µl 70% ethanol, vacuum-dried for 10 - 15 min and then stored at -20°C. Prior to electrophoresis, the purified extention products were resuspended in 3.5 µl Blue dextran/EDTA loading buffer (PE Applied Biosystems), denatured for 2 min at 90°C and loaded onto the ABI Prism model 377 DNA sequencer.

2.2.13 Sequence analysis

The obtained nucleic acid sequences were edited with the ABI Prism Sequencing Analysis 3.1 and the ABI Prism Sequencing Navigator 1.0.1 software programs (PE

Applied Biosystems). The ClustalX program was used to perform multiple sequence alignment analyses of the South African viverrid rabies virus strain and several other rabies virus strains. The nucleic acid and amino acid sequences used in the alignment were extracted from GenBank[1] and had the following accession numbers: SAD, M31046; Pasteur virus, X03673; ERA, X55727; Pitman-Moore, X55729; RC-HL, AB009663; and Nishigahara, AB044824. The nucleic acid sequences were translated using Seqpup, and resulting amino acid sequences were used in PROSITE to identify possible protein kinase recognition sites. The sequence data is presented in multiple alignment formats generated by ClustalX, saved in postscript format and displayed in CorelDraw v. 9.0.

2.3 Results

2.3.1 cDNA synthesis of the N and P genes

Brain material of suspected rabid *Cynictis penicillata* (Yellow mongoose) were screened for rabies virus by means of fluorescence antibody typing (FAT) using monoclonal antibodies directed against the nucleoprotein. The viruses were serotyped further to distinguish between Canid and Viverrid serotypes. A viverrid rabies virus isolate (isolate 710/90) was identified and sequence analysis of the nucleoprotein and glycoprotein genes confirmed the viverrid biotype of the isolate (Olivier, 1997). This isolate was selected for further use in this investigation.

Following propagation of rabies virus isolate 710/90 in suckling mice, total RNA was extracted from the lyophilized brain material and used in reverse transcriptase-mediated PCR amplification of full-length cDNA copies of the N and P genes. To aid in the design of oligonucleotide primers for the specific amplification of the P gene, nucleic acid sequences of the P gene of ERA, PM and CVS-11 rabies virus strains (Larson and Wunner, 1990) as well as the Pasteur virus strain (Tordo *et al.*, 1986) were aligned using the ClustalX multiple sequence alignment program (Thompson *et al.*, 1997). Primers for the specific amplification of the full-length N gene were designed based on the

---

[1] The GenBank database website address: www.ncbi.nlm.nih.gov/GenBank/

Fig. 2.1 Agarose gel electrophoretic analyses of the RT-PCR cDNA amplicons of viverrid rabies virus isolate 710/90. Lane MW, Molecular weight marker; Lane 2, PHF-PHR amplicon of the P gene; Lane 3, JJ1-N, amplicon of the N gene. The sizes of molecular weight marker, phage lambda DNA restricted with EcoR1 and HindIII, are indicated to the left of the figure in base pairs.
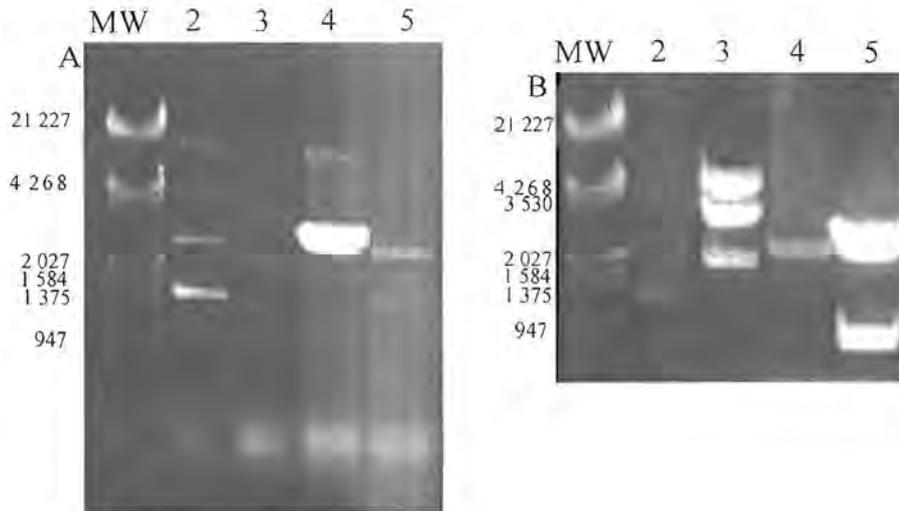
Fig. 2.2 Agarose gel electrophoretic analysis of the recombinant plasmids (A) NpGEM and (B) PpGEM, constructed by cloning of N and P gene cDNA into the pGEM-T Easy vector. (A) Lanes: MW, molecular weight marker; 2, uncut pGEM-T Easy; 3, uncut recombinant NpGEM; 4, EcoR1 and BamH1-digested pGEM-T Easy; 5, EcoR1 and BamH-digested recombinant NpGEM. (B) Lanes: MW, molecular weight marker; 2, uncut pGEM-T Easy; 3, uncut recombinant PpGEM; 4, EcoR1-digested pGEM-T Easy; 5, EcoR1-digested recombinant PpGEM. The sizes of molecular weight marker, phage lambda DNA restricted with EcoR1 and HindIII are indicated on the left of the figures in base pairs.

unpublished partial nucleotide sequence of a N gene of a viverrid rabies virus (Olivier, 1997). In order to facilitate cloning of PCR-amplified N and P genes, unique restriction endonuclease recognition sites were included in the primers. The oligonucleotide primer sequences and applications are listed in Table 2.1. Following agarose gel electrophoresis of the reaction mixtures, amplicons of the expected sizes were obtained when compared to the size of the DNA molecular weight marker. These corresponded to 1.4 kb for the full-length N gene (Fig. 2.1, lane 3) and *ca.* 900 bp for the full-length P gene (Fig. 2.1, lane 2).

2.3.2 Construction of recombinant pGEM® T-Easy vectors containing the full-length N and P genes

The use of *Taq* polymerase lacking a proofreading ability often result in amplicons containing a single non-template specific deoxyadenosine at the 5' ends of the amplified fragments. The pGEM®-T Easy Vector contains single 3'-thymidine overhangs at the insertion site, which results in a highly effective PCR-cloning system. Following gel-purification of the amplicons and ligation into pGEM®-T Easy, competent *E. coli* JM109 cells were transformed and colonies with a white phenotype were selected and cultured in LB-broth supplemented with ampicillin. Following plasmid DNA extractions, the plasmid DNA was analyzed for N and P gene-specific inserts by using restriction endonucleases of which the recognition sites have been incorporated during the design of the primers. The putative recombinant NpGEM clones were digested with BamH1 and EcoR1 and resulted in excision of a 1.4 kb DNA fragment, corresponding to the size of the expected full-length N gene (Fig. 2.2A, lane 5). Digestion of the selected putative recombinant PpGEM clones with EcoR1 resulted in excision of a DNA fragment of *ca.* 900 bp (Fig. 2.2B, lane 5). Recombinant plasmid DNA containing either the N or P gene, was selected and used for further DNA manipulations.

2.3.3 Sequencing of the viverrid rabies virus N and P genes

To verify the integrity of the cloned N and P genes, nucleotide sequences were determined by automated sequencing as described in Section 2.2.12. In addition to using the universal pUC/M13 forward and reverse sequencing primers, internal gene-specific

```
              ************* ***** ****    ******** ********* *** *** *** ******   ***********
Nishigra     ATGGATGCCGACAGGATTGTATTCAGATCTAATAATCAGGTGGTCTCTTTGAGGCCTGAGATTATCGCTGATCAATATGA   80
RCHL         ATGGATGCCGACAGGATTGTATTCAGAGCTAATAATCAGGTGGTCTCTTTGAGGCCTGAGATTATCGCTGATCAATATGA   80
CVS          ATGGATGCCGACAAGATTGTGTTCAAAGTCAATAATCAGGTGGTCTCTTTGAGGCCTGAGATTATCGTGGATCAATATGA   80
Pasteur      ATGGATGCCGACAAGATTGTGTTCAAAGTCAATAATCAGGTGGTCTCTTTGAAGCCTGAGATTATCGTGGATCAATATGA   80
SAD          ATGGATGCCGACAAGATTGTATTCAAAGTCAATAATCAGGTGGTCTCTTTGAAGCCTGAGATTATCGTGGATCAATATGA   80
ViverridSA   ATGGATGCCGACAAGATTGTATTCAAGGTCAATAATCAAGTGGTCTCTCTGAAGCCCGAGGTTATCGTGGATCAATATGA   80

              ******** ***** ***** ** *   ** *****  * ** ******* ** ** ***** ** ** **********
Nishigra     GTACAAGTACCCTGCTATCAAAGATTTGAAAAAGCCCTGTATAACCCTAGGGAAAGCTCCCGACTTGAACAAAGCATACA  160
RCHL         GTACAAGTACCCTGCTATCAAAGATTTGAAAAAGCCCTGTATAACCCTAGGGAAAGCTCCCGACTTGAACAAAGCATACA  160
CVS          GTACAAGTACCCTGCCATCAAGGATTTGAAAAAGCCTTGTATCACCCTAGGGAAAGCCCCCGACTTGAACAAAGCATACA  160
Pasteur      GTACAAGTACCCTGCCATCAAGGATTTGAAAAAGCCTTGTATCACCCTAGGGAAAGCCCCCGACTTGAACAAAGCATACA  160
SAD          GTACAAGTACCCTGCTATCAAAGATTTGAAAAAGCCCTGTATAACCCTAGGGAAAGGCTCCCGATTTAAATAAAGCATACA 160
ViverridSA   GTACAAGTATCCTGCTATCAAGGACTCAAAGAAGCCTAGCATAACCCTAGGGAAAAGCCCCCGACTTAAACAAAGCATACA 160

              * ** ***** ***** ****  ***** ****  * ** ** ******** *****  ******* ***********
Nishigra     AATCAGTTTTATCAGGCATGAATGCCGCCAAACTTGACCCCGACGATGTATGTTCCTACTTGGCAGCAGCAATGCAGTTC  240
RCHL         AATCAGTTTTATCAGGCATGAATGCCGCCAAACTTGACCCCGACGATGTATGTTCCTACTTGGCAGCAGCAATGCAGTTC  240
CVS          AATCAGTTTTATCAGGCATGAATGCCGCCAAACTTGATCCGGATGTGTATGTCTCCTACTTGGCAGCAGCAATGCAGTTC  240
Pasteur      AATCAGTTTTATCAGGCATGAATGCCGCCAAACTTGATCCGGATGTGTATGTCTCCTACTTGGCAGCAGCAATGCAGTTC  240
SAD          AGTCAGTTTTGTCAGGCATGAGCGCCGCCAAACTTAATCCTGACGATGTATGTTCCTATTTGGCAGCGGCAATGCAGTTT  240
ViverridSA   AGTCGGTTTTGTCAGGTATGAACGCCGCAAAACTCGACCCCGATGATGTATGTTCCTACCTGGCAGCGGCAATGCAGTTC  240

              ******* *   ***** ***** ** ******* *********** *******  * ******** * ********  **
Nishigra     TTTGAGGGGACGTGTCCGGAAGACTGGACCAGCTATGGAATCCTGATTGCAAGAAAAGGAGACAAGATCACCCCCAAATTC 320
RCHL         TTTGAGGGGACGTGTCCGGAAGACTGGACCAGCTATGGAATCCTGATTGCAAGAAAAGGAGACAAGATCACCCCCAAATTC 320
CVS          TTTGAGGGGACATGTCCGGAAGACTGGACCAGCTATGGAATCCTGATTGCACGAAAAGGAGATAGGATCACCCCCAAACTC 320
Pasteur      TTTGAGGGGACATGTCCGGAAGACTGGACCAGCTATGGAATCCTGATTGCACGAAAAGGAGATAGGATCACCCCCAAACTC 320
SAD          TTTGAGGGGACGTGTCCGGAAGACTGGACCAGCTATGGAATTGTGATTGCACGAAAAGGAGATAAGATCACCCCCAGGTTC 320
ViverridSA   TTTGAGGGAATATGTCCAGAAAACTGGACTAGCTATGGAATCTTGATTGCGAGGAAGGAGACAAGATCACCCCCGGATTC  320

              * * ***** ***** ** * *************** ***** ********* ***** **** *  ******** *
Nishigra     TCTGGTGGAAATAAAACGCAATGATGTAGAAGGGAATTGGGCTCTGACAGGAGGTATGGAAATGACAAGGGACCCCACTG  400
RCHL         TCTGGTGGAAATAAAACGCAATGATGTAGAAGGGAATTGGGCTCTGACAGGAGGTATGGAACTGACAAGGGACCCCACTG  400
CVS          TCTAGTGGAGATAAAGCGTACTGATGTAGAAGGGAATTGGGCTCTGACAGGAGGCATGGAATTGACAAGGGACCCCACTG  400
Pasteur      TCTAGTGGAGATAAAGCGTACTGATGTAGAAGGGAATTGGGCTCTGACAGGAGGCATGGAATTGACAAGGGACCCCACTG  400
SAD          TCTGGTGGAGATAAAACGTACTGATGTAGAAGGGAATTGGGCTCTGACAGGAGGCATGGAACTGACAAGAGACCCCACTG  400
ViverridSA   TATTGTGGATATAAAGCGTACTGATGTGAGAAGGGAACTGGGCTTTGACAGGAGGCATGGAATTGACGAAGGACCCCACCG 400

              ** **** ***** ** ** *********** ******* ** ********* ***** ** ** ***** ** **
Nishigra     TCTCTGAGCATGCATCTTTGGTCGGTCTTCTCCTGAGTCTGTATAGGTTGAGCAAAATATCAGGACAAAACACCGGTAAC  480
RCHL         TCTCTGAGCATGCATCTTTGGTCGGTCTTCTCCTGAGTCTGTATAGGTTGAGCAAAATATCAGGACAAAACACCGGTAAC  480
CVS          TCTCTGAACATGCATCTTTAGTCGGTCTTCTCCTGAGTCTGTACAGGTTGAGCAAAATATCAGGACAGAACACTGGTAAC  480
Pasteur      TCTCTGAACATGCATCTTTAGTCGGTCTTCTCCTGAGTCTGTACAGGTTGAGCAAAATATCAGGACAGAACACTGGTAAC  480
SAD          TCCCTGAGCATGCGTCCTTAGTCGGTCTTCTCTTGAGTCTGTATAGGTTGAGCAAAATATCGGGCAAAACACTGGTAAC   480
ViverridSA   TCTCTGAGCATGCGTCCTTGGTCGGTCTTCTCTTGAGTCTATATCGGTTGAGCAAGATATCAGGCCAAAACACCGGCAAT  480

              **** ************* ***** ******************* ** ***** ***** ** ** ** ***** *
Nishigra     TATAAAACAAACATTGCAGATAGGATAGAGCAGATTTTTGAAACAGCTCCTTTTGTTAAAATCGTGGAGCACCATACTCT  560
RCHL         TATAAAACAAACATTGCAGATAGGATAGAGCAGATTTTTGAAACAGCTCACCTTTTGTTAAAATCGTGGAGCACCATACTCT 560
CVS          TATAAGACAAACATTGCAGATAGGATAGAGCAGATTTTCGAGACAGCACCTTTTTGTTAAGATCGTGGAACACCATACCCT 560
Pasteur      TATAAGACAAACATTGCAGATAGGATAGAGCAGATTTTCGAGACAGCACCTTTTTGTTAAGATCGTGGAACACCATACCCT 560
SAD          TATAAGACAAACATTGCAGACAGGATAGAGCAGATTTTTGAGACAGCCCCTTTTGTTAAAATCGTGGAACACCATACTCT  560
ViverridSA   TATAAAACAAACATTGCAGATAGGATAGAGCAGATTTTCGAGACAGCTCCTTTCGTCAAAATAGTGGAACATCATACTTT  560

              ***** ***** ** ******** *********** ***** *********** ***** ******** **********
Nishigra     CATGACAACTCACAAAATGTGTGCTAATTGGAGTACCATACCGAACTTCAGATTTTTGGCCGGAACCTACGACATGTTTT  640
RCHL         CATGACAACTCACAAAATGTGTGCCAATTGGAGTACCATACCGAACTTCAGATTTTTGGCCGGAACCTACGACATGTTTT  640
CVS          AATGACAACTCACAAGATGTGTGCTAATTGGAGTACTATACCGAACTTCAGATTTTTGGCCGGAACCTACGACATGTTTT  640
Pasteur      AATGACAACTCACAAGATGTGTGCTAATTGGAGTACTATACCGAACTTCAGATTTTTGGCCGGAACCTACGACATGTTTT  640
SAD          AATGACAACTCACAAAATGTGTGCTAATTGGAGTACTATACCAAACTTCAGATTTTTGGCCGGAACCTATGACATGTTTT  640
ViverridSA   AATGACTACTCATAAGATGTGTGCTAATTGGAGTACTATACCGAACTTCAGATTCTTGGCTGGAACCTATGACATGTTTT  640

              **** ************* ***** ********************* ** ** ***** ** ** ** ** ** ** ***
Nishigra     TCTCCCGGATTGAGCATCTATATTCAGCAATCAGAGTGGGCACAGTTGTTACTGCTTATGAAGACTGTTCAGGGCTGGTA  720
RCHL         TCTCCCGGATTGAGCATCTATATTCAGCAATCAGAGTGGGCACAGTTGTTACTGCTTATGAAGACTGTTCAGGGCTGGTA  720
CVS          TCTCACGGATTGAGCATCTGTATTCGGCAATCAGAGTGGGCACAGTCGTCACCGCTTATGAAGACTGCTCAGGACTGGTA  720
Pasteur      TCTCACGGATTGAGCATCTGTATTCGGCAATCAGAGTGGGCACAGTCGTCACCGCTTATGAAGACTGCTCAGGACTGGTA  720
SAD          TCTCCCGGATTGAGCATCTATATTCAGCAATCAGAGTGGGCACAGTTGTCACTGCTTATGAAGACTGTTCAGGACTGGTA  720
ViverridSA   TCTCCCGGATTGAGCATCTCTATTCGGCAATCAGAGTGGGCACAGTGGTCACTGCTTACGAGGATTGCTCTGGGCTAGTA  720
```

```
                ** ***** *********** ** ** ***** * ** ** * **  * ** ** ***** ** ******* **
Nishigra    TCGTTTACTGGGTTCATAAAGCAGATCAATCTCACCGCAAGAGAAGCAATATTGTACTTTTTCCATAAGAACTTTGAAGA    800
RCHL        TCGTTTACTGGGTTCATAAAGCAGATCAATCTCACTGCGAGAGAAGCAACATTGTACTTTTTCCATAAGAACTTTGAAGA    800
CVS         TCGTTTACAGGGTTCATAAAGCAGATCAATCTCACCGCAAGGGAGGCAATACTATATTTCTTCCACAAGAACTTTGAGGA    800
Pasteur     TCGTTTACAGGGTTCATAAAGCAGATCAATCTCACCGCAAGGGAAGCAATACTATATTTCTTCCACAAGAACTTTGAGGA    800
SAD         TCATTTACTGGGTTCATAAAACAAATCAATCTCACCGCTAGAGAGGCAATACTATATTTCTTCCACAAGAACTTTGAGGA    800
ViverridSA  TCATTTACAGGGTTCATAAAGCAGATAAATTTGACTGCAAAAGAAGCGATACTGTATTTCTTCCACAAAAACTTTGAAGA    800

                *********** ***** *********** ***** ***** ****************** ************** ***
Nishigra    AGAGATAAGAAGGATGTTCGAGCCAGGGCAGGAGACGGCTGTTCCTCACTCTTATTTCATTCACTTCCGTTCACTAGGCC    880
RCHL        AGAGATAAGAAGGATGTTCGAGCCAGGGCAGGAGACAGCTGTTCCTCACTCTTATTTCATTCACTTCCGTTCACTAGGCC    880
CVS         AGAGATAAGAAGAATGTTCGAGCCAGGGCAAGAGACAGCTGTTCCTCACTCTTATTTCATCCACTTCCGTTCACTAGGCT    880
Pasteur     AGAGATAAGAAGAATGTTCGAGCCAGGGCAAGAGACAGCTGTTCCTCACTCTTATTTCATCCACTTCCGTTCACTAGGCT    880
SAD         AGAGATAAGAAGAATGTTTGAGCCAGGGCAGGAGACGGCTGTTCCTCACTCTTATTTCATCCACTTCCGTTCACTAGGCT    880
ViverridSA  AGAGATAAGAAGGATGTTCGAGCCAGGGCAGGAGACGGCTGTCCCTCACTCTTATTTCATTCACTTCCGTTCACTGGGCC    880

                ********** ** ******** ** ***** ** ***** **************************** ******* ******
Nishigra    TGAGTGGGAAGTCTCCTTATTCGTCGAATGCCGTTGGTCATGTGTTCAATCTCATTCACTTTGTTGGATGCTATATGGGT    960
RCHL        TGAGTGGGAAGTCTCCTTATTCGTCGAATGCCGTTGGTCATGTGTTCAATCTCATTCACTTTGTTGGATGCTATATGGGT    960
CVS         TGAGTGGGAAGTCTCCTTATTCATCGAATGCTGTCGGTCATGTGTTCAATCTCATTCACTTTGTTGGATGCTACATGGGT    960
Pasteur     TGAGTGGGAAGTCTCCTTATTCATCGAATGCTGTCGGTCATGTGTTCAATCTCATTCACTTTGTTGGATGCTACATGGGT    960
SAD         TGAGTGGGAAATCTCCTTATTCATCAAATGCTGTTGGTCACGTGTTCACTCTCATTCACTTTGTAGGATGCTATATGGGT    960
ViverridSA  TGAGTGGGAAGTCCCCTTATTCATCAAATGCCGTGGGTCATGTGTTCAATCTCATTCACTTTGTTGGATGCTATATGGGT    960

                *** * ** ** ******** * ** ***** ******* ** ***** *********** ******* * ** **
Nishigra    CAAATCAGATCCCTAAATGCAACAGTTATTGCCGCATGTGCTCCTCATGAGATGTCTGTTCTCGGGGGCTACCTAGGGGA    1040
RCHL        CAAATCAGATCCCTAAATGCAACAGTTATTGCCGCATGTGCTCCTCATGAGATGTCTGTTCTCGGGGGCTACCTAGGGGA    1040
CVS         CAAGTCAGATCTCTAAATGCGACCGGTTATTGCTGCATGTGCCCCTCATGAGATGTCTGTTCTAGGGGGCTATTTGGGAGA  1040
Pasteur     CAAGTCAGATCTCTAAATGCGACGGTTATTGCTGCATGTGCCCCTCATGAGATGTCTGTTCTAGGGGGCTATTTGGGAGA   1040
SAD         CAAGTCAGATCCCTAAATGCAACGGTTATTGCTGCATGTGCTCCTCATGAAATGTCTGTTCTAGGGGGCTATCTGGGAGA   1040
ViverridSA  CAAGTTAGGTCCCTAAATGCAGCAGTCATTGCTGCATGTGCTCCTCATGAGATGTCTGTTCTAGGGGGCTATCTAGGAGA   1040

                *** ** ** ** * *** **** ******* ********** ***************** ** ***** ** ***
Nishigra    GGAATTCTTTGGGAGAGGAACATTCGAAAGAAGATTCTTCAGAGATGAGAAAGAACTTCAAGAATACGAGGCAGCTGAAC    1120
RCHL        GGAATTCTTTGGGAGAGGAACATTCGAAAGAAGATTCTTCAGAGATGAGAAAGAACTTCAAGAATACGAGGCAGCTGAAC    1120
CVS         GGAATTCTTCGGAAAGGGACATTTGAAAGAAGGTTCTTCAGAGACGAGAAAGAACTTCAAGAATATGAGGCGGCTGAAC    1120
Pasteur     GGAATTCTTCGGGAAAGGGACATTTGAAAGAAGGTTCTTCAGAGACGAGAAAGAACTTCAAGAATATGAGGCGGCTGAAC   1120
SAD         GGAATTCTTCGGGAAAGGGACATTTGAAAGAAGATTCTTCAGAGATGAGAAAGAACTTCAAGAATACGAGGCGGCTGAAC   1120
ViverridSA  GGAGTTTTTTGGAAAAGGGACATTTGAAAGAAGATTCTTCAGAGATGAGAAAGAACTTCAAGAGTACGAGGCTGCCGAAT   1120

                * ****** * ***    * ********** ***** *** **** ** ***** *** * **  * ****** *****
Nishigra    TGACAAAGACTGACGTGGCATTGGCAGATGATGGAACTGTCGACTCTGATGACGAAGACTACTTCTCCGGTGAAGCCAGA    1200
RCHL        TGACAAAGACTGACGTGGCATTGGCAGATGATGGAACTGTCGACTCTGACGACGAAGACCACTTTTCCCGGTGAAGCCAGA   1200
CVS         TAACAAAGTCCGACGTGGCATTGGCAGATGACGGAACCGTCAACTCTGATGACGAGGACTATTTCTCTGGTGAAACCAGA    1200
Pasteur     TAACAAAGTCCGACGTGGCACTGGCAGATGACGGAACCGTCAACTCTGATGACGAGGACTATTTCTCTGGTGAAACCAGA    1200
SAD         TGACAAAGACTGACGTAGCACTGGCAGATGATGGAACTGTCAACTCTGACGACGAGGACTACTTTTCAGGTGAAACCAGA    1200
ViverridSA  TAACAAAGACTGACACAATACTGGCAGATGATGGAACTGTCAACTCCGACGACGAGGACTACTTCTCGGGTGAAACCAGA    1200

                * ** ** ***** *** ***************** ** ******* ** ***** ** *** ********** ** **
Nishigra    GGTCCAGAAGCTGTTTATGCTCGAATCATGATGAATGGAGGTCGACTAAAAAGATCGCATATACGGAGATATGTCTCAGT    1280
RCHL        GGTCCAGAAGCTGTTTATGCTCGAATCATGATGAATGGAGGTCGACTAAAAAGATCGCATATACGGAGATATGTCTCAGT    1280
CVS         AGTCCAGAAGCTGTCTATACTCGAATCATGATGAATGGAGGTCGACTGAAGAGATCTCATATACGGAGATATGTCTCAGT    1280
Pasteur     AGTCCAGAAGCTGTTTATACTCGAATCATGATGAATGGAGGTCGACTGAAGAGATCTCATATACGGAGATATGTCTCAGT    1280
SAD         AGTCCGGAGGCTGTTTATACTCGAATCATGATGAATGGAGGTCGACTAAAGAGATCTCACATACGGAGATATGTCTCAGT    1280
ViverridSA  AGCCCTGAAGCTGTCTATACTCGAATCATGATGAACGGGGGTCGACTAAAGAGATCACACATAAGGAGATATGTGTCCGT    1280

                ****** ******** ******** ***** ** ***** *** ********** ***** * ** ** ** * *  * **
Nishigra    CAGTTCCAATCATCAAGCCCGTCCCAAACTCATTCGCCGAGTTTCTAAACAAGACGTATTCGAGCGACTCTTAAGGAGTTA   1360
RCHL        CAGTTCCAATCATCAAGCCCGTCCCAACTCATTCGCCGAGTTTCTAAACAAGACGTATTCGAGCGACTCTTAAGGAGTTA    1360
CVS         CAGTTCCAATCATCAAGCCCGTCCCAAACTCATTCGCCGAATTTTTAAACAAGACGTATTCGAATGACTCTCATAAGGAGTTG 1360
Pasteur     CAGTTCCAATCATCAAGCCCGTCCAAACTCATTCGCCGAATTTTTAAACAAGACGTATTCGAATGACTCATAAGGAGTTG   1360
SAD         CAGTTCCAATCATCAAGCCCGTCCAAACTCATTCGCCGAGTTTCTAAACAAGACATATTCGAGTGACTCATAAGAAGTTG   1360
ViverridSA  CAGTTCTAATCATCAGGCCCGTCCCAACTCCTTTGCCGAGTTTCTAAACAAGACCTATTCTACTGATTCATGAAGATTTG   1360

                *
Nishigra    AATAACGGTG    1370
RCHL        AACAACTGTG    1370
CVS         ATTGATAGGG    1370
Pasteur     ATTGATAGGG    1370
SAD         AATAACAAAA    1370
ViverridSA  AATGGCAAGA    1370
```

Fig. 2.3A Multiple sequence alignment of the 1 370 bp N gene of the South Afican viverrid rabies virus strain with other rabies virus strains using the ClustalX program, as described in Materials and Methods. GenBank accession numbers of rabies virus isolates used in this analysis: SAD, M31046; Pasteur virus, X03673; RC-HL, AB009663; Nishigahara, AB044824; CVS, X55727.

```
          *******  *****  **  ****  ****  **  ********  ****  ****  *******  *****  **  **  *****
pntpitman ATGAGCAAGATCTTTGTCAATCCTAGTGCTATTAGAGCCGGTCTGGCCGATCTTGAGATGGCTGAAGAAACTGTTGATCT   80
pntera    ATGAGCAAGATCTTTGTCAATCCTAGTGCTATTAGAGCCGGTCTGGCCGATCTTGAGATGGCTGAAGAAACTGTTGATCT   80
pntsadb19 ATGAGCAAGATCTTTGTCAATCCTAGTGCTATTAGAGCCGGTCTGGCCGATCTTGAGATGGCTGAAGAAACTGTTGATCT   80
pntpv     ATGAGCAAGATCTTTGTCAATCCTAGTGCTATTAGAGCCGGTCTGGCCGATCTTGAGATGGCTGAAGAAACTGTTGATCT   80
pntrchl   ATGAGCAAGATCTTCTCAATCCGAGTGCTATCAGAGCCGGTATGGCCGATCTTGAGATGGCTGAAGAGACTGTTGATCT    80
pntcvs    ATGAGCAAGATCTTTGTTAATCCGAGTGCAATCAGAGCCGGTCTGGCCGATCTTGAGATGGCCGAAGAGACTGTTGATCT   80
phosSAnt  ATGAGCAAAATCTTTGTCAATCCGAGTGCAATTAGAGCCGGTCTGGCAGATCTCGAGATGGCCGAAGAGACCGTCGATCT   80

          **  **  **  **  **  **********  *****  *****  **  **  **  *****  ******  **********  ***
pntpitman GATCAATAGAAATATCGAAGACAATCAGGCTCATCTCCAAGGGGAACCCATAGAAGTGGACAATCTCCCTGAGGATATGG  160
pntera    GATCAATAGAAATATCGAAGACAATCAGGCTCATCTCCAAGGGGAACCCATAGAAGTGGACAATCTCCCTGAGGATATGG  160
pntsadb19 GATCAATAGAAATATCGAAGACAATCAGGCTCATCTCCAAGGGGAACCCATAGAGGTGGACAATCTCCCTGAGGATATGG  160
pntpv     GATCAATAGAAATATCGAAGACAATCAGGCTCATCTCCAAGGGGAACCCATAGAAGTGGACAATCTCCCTGAGGATATGG  160
pntrchl   GATTAATAGAAACATAGAAGACAATCAGGCTCACCTCCAGGGAGAACCTATAGAAGTGGACAGTCTCCCTGAGGACATGA  160
pntcvs    GATCAACAGAAACATAGAAGACAATCAGGCTCATCTCCAGGGAGAACCCATAGAAGTGGACAACCTCCCTGAGGACATGA  160
phosSAnt  AATCAATAGGAATATAGAAGACAATCAAGCTCATCTCCAGGGAGAGCCCATAGAGGTGGACAGTCTCCCTGAGGACATGA  160

          ***  *****  ****  **  ***  *  *  ***  *  *  **  *  **  **  ***  *  ***  ******
pntpitman GGCGACTTCACCTGGATGATGGAAAATCGCCCAACCCTGGTGAGATGGCCAAGGTGGGAGAAGGCAAGTATCGAGAGGAC  240
pntera    GGCGACTTCACCTGGATGATGGAAAATCGCCCAACCCTGGTGAGATGGCCAAGGTGGGAGAAGGCAAGTATCGAGAGGAC  240
pntsadb19 GGCGACTTCACCTGGATGATGGAAAATCGCCCAACCATGGTGAGATAGCCAAGGTGGGAGAAGGCAAGTATCGAGAGGAC  240
pntpv     GGCGACTTCACCTGGATGATGGAAAATCGCCCAACCCTGGTGAGATGGCCAAGGTGGGAGAAGGCAAGTATCGAGAGGAC  240
pntrchl   GCCGAATTCACTTGGATTATGGAAAGTTGCCCGACCTTGGTAGGATGTCCAAAGCGGGGGAGGGCAGGCATCAAGAGGAC  240
pntcvs    AGCGACTTCACCTGGACGATGAAAAATCGTCTAACCTTGGTGAGATGGTTAGGGTGGGAGAAGGCAAGTATCGAGAGGAC  240
phosSAnt  GGCGACTTCACTTGGACGATAGAAAACCGTCTGACCTTAGCGATATGGTCAAGTCAGGAGAAAGCAAGTATCGAGAGGAT  240

          *************  *****  **  **  **  *  ************  **  *  ********************  **
pntpitman TTTCAGATGGATGAAGGAGAGGATCCTAGCTTCCTGTTCCAGTCATACCTGGAAAATGTTGGAGTCCAAATAGTCAGACA  320
pntera    TTTCAGATGGATGAAGGAGAGGATCCTAGCTTCCTGTTCCAGTCATACCTGGAAAATGTTGGAGTCCAAATAGTCAGACA  320
pntsadb19 TTTCAGATGGATGAAGGAGAGGATCCTAGCTTCCTGTTCCAGTCATACCTGGAAAATGTTGGAGTCCAAATAGTCAGACA  320
pntpv     TTTCAGATGGATGAAGGAGAGGATCCTAGCCTCCTGTTCCAGTCTGGACAATGTTGGAGTCCAAATAGTCAGACA       320
pntrchl   TTTCAGATGGATGAAGGAGAAGATCCTAGCCTCCTGTTCCAGTCGTACCTAGACAACATTGGAGTCCAAATAGTCAGACA  320
pntcvs    TTTCAGATGGATGAGGGAGAGGACCCCAACCTCCTGTTCCAATCGTACCTGGATAATGTTGGAGTCCAAATAGTCAGACA  320
phosSAnt  TTTCAGATGGATGAAGGAGAAGATCCCATCCTCCTGTTTCAATCATACCTGGACAGTGTTGGAGTCCAAATAGTCAGGCA  320

          ***  **********  *****  **  *****  **************  *  **  **  **********  *  *******  *
pntpitman AATGAGGTCAGGAGAGAGATTTCTCAAGATATGGTCACAGACCGTAGAAGAGATTATATCCTATGTCGCGGTCAACTTTC  400
pntera    AATGAGGTCAGGAGAGAGATTTCTCAAGATATGGTCACAGACCGTAGAAGAGATTATATCCTATGTCGCGGTCAACTTTC  400
pntsadb19 AATGAGGTCAGGAGAGAGATTTCTCAAGATATGGTCACAGACCGTAGAAGAGATTATATCCTATGTCGCGGTCAACTTTC  400
pntpv     AATAAGGTCAGGAGAGAGATTTCTCAAGATATGGTCACAGACCGTAGAAGAGATTATATCCTATGTCGCGGTCAACTTTC  400
pntrchl   AATGAGGTCAGGAGAGAAGATTCCTCAAAATATGGTCACAGACCGTAGAGGAAATTGTATCCTATGTCACGGTCAACTTCC  400
pntcvs    AATGAGGTCAGGAGAGAGATTCCTCAAGATATGGTCACAGACCGTAGAGGAAATTGTATCCTATGTCACGGTCAACTTTC  400
phosSAnt  AATGAGGTCAGGAGAGAGATTCCTCAAGATATGGTCACAGACTGCAGAAGAGATTATATCCTATGTCGCAGTCAACTTTC  400

          *  ****  *  **  ***  ***  **  *****  **  **  *****  *  **  *  ***  *  **  ******  ***  *  *  **
pntpitman CCAACCCTCCAGGAAAGTCTTCAGAGGATAAATCAACCCAGACTACTGGCCGAGAGCTCAAGAAGGAGACAACACCCACT  480
pntera    CCAACCCTCCAGGAAAGTCTTCAGAGGATAAATCAACCCAGACTACTGGCCGAGAGCTCAAGAAGGAGACAACACCCACT  480
pntsadb19 CCAACCCTCCAGGAAAGTCTTCAGAGGATAAATCAACCCAGACTACTGGCCGAGAGCTCAAGAAGGAGACAACACCCACT  480
pntpv     CCAACCCTCCAGGAAAGTCTTCAGAGGATAAATCAACCCAGACTACCGGCCGAGAGCTCAAGAAGGAGACAACACCCACT  480
pntrchl   CCAACCCTTCAGGAAGGTCTTCAGAGGATAAGTCAACCCGGATCACTAGTCAAGAGCCCAAGAAGGAGACAACATCCACT  480
pntcvs    CTAACCCTCCAAGAAGGTCTTCTGAGGATAAATCAACCCAGACTACTGGCAGAGAGCTCAAGAAGGAGACAACGTCTGCT  480
phosSAnt  CCAACCGTCCAGGAAAGTCCTCTGAGGACAAATCCACCCAAACCACCGGCCAAGAACCCAGGAAGGAAACAGCATCTGCT  480

          **********  ******  *  ***  ****  *  **  *  *  ******  **  ****  ******  ****  *******
pntpitman CCTTCTCAGAGAGAAAGCCAATCATCGAAAGCCAGGATGGCGGCTCAAACTGCTTCTGGCCCTCCAGCCCTTGAATGGTC  560
pntera    CCTTCTCAGAGAGAAAGCCAATCATCGAAAGCCAGGATGGCGGCTCAAACTGCTTCTGGCCCTCCAGCCCTTGAATGGTC  560
pntsadb19 CCTTCTCAGAGAGAAAGCCAATCATCGAAAGCCAGGATGGCGGCTCAAATTGCTTCTGGCCCTCCAGCCCTTGAATGGTC  560
pntpv     CCTTCTCAGAGAGAAAGCCAATCCTCGAAAGCCAGGATGGCGGCTCAAACTGCTTCTGGCCCTCCAGCCCTTGAATGGTC  560
pntrchl   CCTTCTCAGAGAGCAAGCCAGTCCTTGAAATCCAGGACGATGGCTCAAACTGCCTCTGGCCCTCCATCCCTTGAATGGTC  560
pntcvs    TTCTCTCAGAGAGAAAGCCAACCTTCGAAAGCTAGGATGGTGGCTCAAGTTCCCCTGGTCCTCCAGCCCTTGAATGGTC  560
phosSAnt  TCTTCTCAGAGAGAAAGCCAGTCCTCGAAAGCCAGAATGGCTGCTCAAGCCGCCTCTGGTCCTCCAGCCCTCGAATGGTC  560

          **  **  *****  *  *******  *****  *************  **  *******  *****  **  **  **  **  **  *
pntpitman GGCCACCAATGAAAAGGATGATCTATCAGTGGAGGCTGAGATCGCTCACCAGATTGCAGAAAGTTTCTCCAAAAAAATATA  640
pntera    GGCCACCAATGAAAAGGATGATCTATCAGTGGAGGCTGAGATCGCTCACCAGATTGCAGAAAGTTTCTCCAAAAAAATATA  640
pntsadb19 GGCTACCAATGAAGAGGATGATCTATCAGTGGAGGCTGAGATCGCTCACCAGATTGCAGAAAGTTTCTCCAAAAAAATATA  640
pntpv     GGCCACCAATGAAAAGGATGATCTATCAGTGGAGGCTGAGATCGCTCACCAGATTGCAGAAAGTTTCTCCAAAAAAATATA  640
pntrchl   TGCTACTAATGAGGAGGATGATCTATCAGTAGAGGCTGAGATCGCTCACCAGATTGCTGAAAGCTTTTTCTAAGAAGTACA  640
pntcvs    AGCCACCAATGAAGAAGATGATCTATCAGTAGAGGCTGAGATCGCTCATCAGATTGCTGAAAGCTTTTTCCAAGAAGTACA  640
phosSAnt  TGCTACCAATGAAGAGGATGATCCATCAGTAGAGGCTGAGATCGCCCACCAGATTGCAGAAAGCTTCTCCAAGAAATACA   640
```

```
              ****** ** ** ** ***** *** ******************  **** ***** ******** ******  **
pntpitman   AGTTTCCCTCTCGATCCTCAGGGATACTCTTGTATAATTTTGAGCAATTGAAAATGAACCTTGATGATATAGTTAAAGAG  720
  pntera    AGTTTCCCTCTCGATCCTCAGGGATACTCTTGTATAATTTTGAGCAATTGAAAATGAACCTTGATGATATAGTTAAAGAG  720
pntsadb19   AGTTTCCCTCTCGATCCTCAGGGATACTCTTGTATAATTTTGAGCAATTGAAAATGAACCTTGATGATATAGTTAAAGAG  720
  pntpv     AGTTTCCCTCTCGATCCTCAGGGATACTCTTGTATAATTTTGAGCAATTGAAAATGAACCTTGATGATATAGTTAAAGAG  720
 pntrchl    AGTTTCCTTCTCGGTCTTCAGGGATATTCTTGTATAATTTTGAGCAGTTGAAAATGAACCTTGATGATATAGTTAAGGAA  720
  pntcvs    AGTTTCCCTCCCGATCTTCAGGAATATTCTTGTATAATTTTGAGCAACTGAAGATGAACCTTGATGACATAGTTAAAGAG  720
 phosSAnt   AGTTTCCCTCTCGATCATCAGGAATATTCTTGTATAATTTTGAGCAGTTGAAAATGAATCTTGATGACATAGTTAGAGAG  720


              ********** ** ** ** ****** * ** * *** ** ****** **** ** ***** ** ***** *******
pntpitman   GCAAAAAATGTACCATGAGTGACCCGTTGACCCATGACGGGTCCAAACTCCCCCTAAGATGTGTATGGGATGGGTCGC  800
  pntera    GCAAAAAATGTACCAGGTGTGACCCGTTTAGCCCATGACGGGTCCAAACTCCCCCTAAGATGTGTACTGGGATGGGTCGC  800
pntsadb19   GCAAAAAATGTACCAGGTGTGACCCGTTTAGCCCATGACGGGTCCAAACTCCCCCTAAGATGTGTACTGGGATGGGTCGC  800
  pntpv     GCAAAAAATGTACCAGGTGTGACCCGTTTAGCCCGTGACGGGTCCAAACTCCCCCTAAGATGTGTACTGGGATGGGTCGC  800
 pntrchl    GCAAAAAATGTGCCAGGCGTAACCCGTTTGGCCCATGACGGATCCAAACTCCCCCTAAGATGTGTACTGGGATGGGTCGC  800
  pntcvs    GCAAAAAATGTACCGGGCGTGACCCGTCTGGCCCATGATGGATCCAAAATCCCCCTGAGATGCGTACTGGGATGGGTCGC  800
 phosSAnt   GCAAAAAATGTGCCAGGTGTGACCCGTTTGGCACATGACGGATCCAAACTCCCTCTGAGATGTGTGCTGGGGTGGGTCGC  800


              * ** ** ** ** ******* ** * ***** *  ******* *  ** ** ********** * *******
pntpitman   TTTGGCCAACTCTAAGAAATTCCAGTTGTTAGTCGAATCCGACAAGCTGAGTAAAATCATGCAAGATGACTTGAATCGCT  880
  pntera    TTTGGCCAACTCTAAGAAATTCCAGTTGTTAGTCGAATCCGACAAGCTGAGTAAAATCATGCAAGATGACTTGAATCGCT  880
pntsadb19   TTTGGCCAACTCTAAGAAATTCCAGTTGTTAGTCGAATCCGACAAGCTGAGTAAAATCATGCAAGATGACTTGAATCGCT  880
  pntpv     CTTGGCCAACTCTAAGAAATTCCAGTTGTTAGTCGAATCCAACAAGCTGAGTAAAATCATGCAAGATGACTTGAATCGCT  880
 pntrchl    TCTGGCCAATTCTAAGAAATTCCAATTGTTAGTCGAAGCCAACAAGCTAAATAAAATCATGCAAGATGATTTGAATCGCT  880
  pntcvs    TTTGGCTAATTCCAAAAAATTCCAATTACTAGTCGAGGCTGACAAGCTAAGCAAAATCATGCAAGATGATTTGAATCGCT  880
 phosSAnt   CCTAGCCAACTCTAAGAAATTCCAGTTGTTGGTCGAGTCTGACAAGCTAAGCAAGATTATGCAAGATGATCTAAATCGCT  880


              ** **** ******
pntpitman   ATACATCTTGCTAA  894
  pntera    ATACATCTTGCTAA  894
pntsadb19   ATACATCTTGCTAA  894
  pntpv     ATACATCTTGCTAA  894
 pntrchl    ATGCATCTTGCTAA  894
  pntcvs    ATACATCCTGCTAA  894
 phosSAnt   ATACATCTTGCTAA  894
```

Fig. 2.3B Multiple sequence alignment of the 894 bp P gene of the South African viverrid rabies virus strain with other rabies virus strains using the ClustalX program. GenBank accession numbers of the isolates used: Pitman-Moore, X55729; ERA, X55728; SAD-B19, M31046; Pasteur virus (PV), X03673; RC-HL, AB009663; CVS, X55727. The rabies virus strains in the alignment is indicated with the following names: Pitman-Moore, pntpitman; ERA, pntera; SAD-B19, pntsadb19; Pasteur virus, pntpv; RC-HL, pntrchl; CVS, pntcvs and the South African viverrid strain is named phosSAnt.

primers were also designed (Table 2.1) to obtain the complete nucleic acid sequence of the respective genes. The obtained nucleic acid sequence for the N gene is indicated in Fig. 2.3A and that for the P gene in Fig. 2.3B.

The N gene of the South African viverrid rabies virus strain was found to be 1 370 nucleotides in length. Comparison of the sequence to other rabies virus strains indicated 93.79% homology. The deduced amino acid sequence of the N protein encoded by the viverrid N gene indicated that the protein consists of 450 amino acid residues and has an estimated molecular weight of 55 kDa. Alignment comparison of the amino acid sequences of the viverrid virus strain indicated 96.2% identity to the N protein of the CVS rabies virus strain (Fig 2.4A). The C-terminal 74 amino acids of the CVS rabies virus strain N protein have been reported to be involved in binding to the P protein (Schoehn *et al.*, 2001). Comparison of this region, between the viverrid and CVS strains indicated a 96% identity (Fig 2.4A). The sequence spanning amino acids 298-352 has been reported to be involved in binding of the Pasteur virus strain N protein to the viral ssRNA genome (Kouznetzoff *et al.*, 1998). Comparison of this region to the corresponding region of the viverrid rabies virus included in the analysis indicated 98% identity. Only a single amino acid change in position 328 was noted (threonine in PV to an alanine in the viverrid strain).

The phosphoprotein gene of the South African viverrid rabies virus strain is 894 nucleotides in length, displays 90.4% and 90.5% identity with the CVS and Pasteur virus strains, respectively. The encoded P protein is composed of 297 amino acids, with a predicted molecular weight of 36 kDa, and displays 90% identity to both the Pasteur virus and the CVS rabies virus strains (Fig. 2.4B). The N-terminal 19 amino acids, which is the region responsible for binding to the L protein (Chenik *et al.*, 1998), was shown to be conserved in all the strains used in this comparison. This high level in sequence conservation confirms the functional importance of this domain (Chenik *et al.*, 1998).

The amino acid sequences of the N and P proteins of the viverrid rabies virus strain were also investigated in order to identify sequences characteristic of phosphorylation sites.

The consensus phosphorylation sites for protein kinase C and casein kinase II are summarized in Table 2.2. The only other kinase recognition sites on the N protein was for Tyrosine kinase at amino acids 359 – 367 (RDEKELQEY) and 398 – 404 (RSPEAVY). These sites are not conserved in the Nishighara and the RC-HL rabies virus strains and were not considered during any further analysis. Phosphorylation of these sites will be determined by various factors, such as availability of the kinase within the host cell, and the presentation of the site within the protein for recognition by the kinase. Phosphorylation of a protein by a specific kinase do not guarantee any functional significance for the protein, and will only be known once the function of the phosphorylated and unphosphorylated forms of the protein is compared.

Table 2.2. Consensus protein kinase recognition sites identified in the viverrid rabies virus N and P proteins.

Rabies virus protein	Kinase	Phosphorylation site	
		Location	sequence
Nucleoprotein	Protein Kinase C	16-18	SLK
		190-192	THK
		295-297	SGK
	Casein Kinase II	107-110	SIVD
		114-117	TDVE
		133-136	TVSE
		215-218	SRIE
		231-234	TAYE
		252-255	TAKE
		375-378	TKTD
		389-392	SDDE
		437-440	SFAE
Phosphoprotein	Protein Kinase C	162-164	SQR
		168-170	SSK
		210-212	SKK
		271-273	SKK
	Casein Kinase II	95-98	SYLD
		121-124	TAEE
		140-143	SSED
		162-165	SQRE
		189-192	TNEE

```
              ****:***: *******:**:*.************ *** .*******************.****:*************
SAD        MDADKIVFKVNNQVVSLKPEIIVDQYEYKYPAIKDLKKPCITLGKAPDLNKAYKSVLSGMSAAKLNPDDVCSYLAAAMQF   80
Viverrid   MDADKIVFKVNNQVVSLKPEVIVDQYEYKYPAIKDSRKPSITLGKAPDLNKAYKSVLSGMNAAKLDPDDVCSYLAAAMQF   80
PV         MDADKIVFKVNNQVVSLKPEIIVDQYEYKYPAIKDLKKPCITLGKAPDLNKAYKSVLSGMNAAKLDPDDVCSYLAAAMQF   80
CVS        MDADKIVFKVNNQVVSLKPEIIVDQYEYKYPAIKDLKKPCITLGKAPDLNKAYKSVLSGMNAAKLDPDDVCSYLAAAMQF   80
Nishighara MDADRIVFRSNNQVVSLRPEIIADQYEYKYPAIKDLKKPCITLGKAPDLNKAYKSVLSGMNAAKLDPDDVCSYLAAAMQF   80
RC-HL      MDADRIVFRANNQVVSLRPEIIALQYEYKYPAIKDLKKPCITLGKAPDLNKAYKSVLSGMNAAKLDPDDVCSYLAAAMQF   80

              *** ***:******.:*****.***.*.*.***.*************.*.*:****.*********************
SAD        FEGTCPEDWTSYGIVIARKGDKITPGSLVEIKRTDVEGNWALTGGMELTRDPTVPEHASLVGLLLSLYRLSKISGQNTGN  160
Viverrid   FEGICPENWTSYGILIARKGDKITPDSIVDIKRTDVEGNWALTGGMALTKDPTVSEHASLVGLLLSLYRLSKISGQNTGN  160
PV         FEGTCPEDWTSYGILIARKGDRITPNSLVEIKRTDVEGNWALTGGMELTRDPTVSEHASLVGLLLSLYRLSKISGQNTGN  160
CVS        FEGTCPEDWTSYGILIARKGDRITPNSLVEIKRTDVEGNWALTGGMELTRDPTVSEHASLVGLLLSLYRLSKISGQNTGN  160
Nishighara FEGTCPEDWTSYGILIARKGDKITPNSLVEIKRNDVEGNWALTGGMEMTRDPTVSEHASLVGLLLSLYRLSKISGQNTGN  160
RC-HL      FEGTCPEDWTSYGILIARKGDKITPNSLVEIKRNDVEGNWALTGGMELTRDPTVSEHASLVGLLLSLYRLSKISGQNTGN  160

              ************************************************************************
SAD        YKTNIADRIEQIFETAPPVKIVEHHTLMTTHKMCANWSTIPNFRFLAGTYDMFFSRIEHLYSAIRVGTVVTAYEDCSGLV  240
Viverrid   YKTNIADRIEQIFETAPPVKIVEHHTLMTTHKMCANWSTIPNFRFLAGTYDMFFSRIEHLYSAIRVGTVVTAYEDCSGLV  240
PV         YKTNIADRIEQIFETAPPVKIVEHHTLMTTHKMCANWSTIPNFRFLAGTYDMFFSRIEHLYSAIRVGTVVTAYRDCSGLV  240
CVS        YKTNIADRIEQIFETAPPVKIVEHHTLMTTHKMCANWSTIPNFRFLAGTYDMFFSRIEHLYSAIRVGTVVTAYRDCSGLV  240
Nishighara YKTNIADRIEQIFETAPPVKIVEHHTLMTTHKMCANWSTIPNFRFLAGTYDMFFSRIEHLYSAIRVGTVVTAYEDCSGLV  240
RC-HL      YKTNIADRIEQIFETAPPVKIVEHHTLMTTHKMCANWSTIPNFRFLAGTYDMFFSRIEHLYSAIRVGTVVTAYEDCSGLV  240

              ****************:** ************************************************************
SAD        SFTGFIKQINLTAREAILYFFHKNFEEEIRRMFEPGQETAVPHSYFIHFRSLGLSGKSPYSSNAVGHVFNLIHFVGCYMG  320
Viverrid   SFTGFIKQINLTAKEAILYFFHKNFEEEIRRMFEPGQETAVPHSYFIHFRSLGLSGKSPYSSNAVGHVFNLIHFVGCYMG  320
PV         SFTGFIKQINLTAREAILYFFHKNFEEEIRRMFEPGQETAVPHSYFIHFRSLGLSGKSPYSSNAVGHVFNLIHFVGCYMG  320
CVS        SFTGFIKQINLTAREAILYFFHKNFEEEIRRMFEPGQETAVPHSYFIHFRSLGLSGKSPYSSNAVGHVFNLIHFVGCYMG  320
Nishighara SFTGFIKQINLTAREAILYFFHKNFEEEIRRMFEPGQETAVPHSYFIHFRSLGLSGKSPYSSNAVGHVFNLIHFVGCYMG  320
RC-HL      SFTGFIKQINLTAREATLYFFHKNFEEEIRRMFEPGQETAVPHSYFIHFRSLGLSGKSPYSSNAVGHVFNLIHFVGCYMG  320

              *:*****:**************************:*****************:**. *******:*****:*.**:*
SAD        QVRSLNATVIAACAPHEMSVLGGYLGEEFFGKGTFERRFFRDEKELQEYEAAELTKTDVALADDGTVNSDDEDYFSGETR  400
Viverrid   QVRSLNAAVIAACAPHEMSVLGGYLGEEFFGKGTFERRFFRDEKELQEYEAAELTKTDTILADDGTVNSDDEDYFSGETR  400
PV         QVRSLNATVIAACAPHEMSVLGGYLGEEFFGKGTFERRFFRDEKELQEYEAAELTKSDVALADDGTVNSDDEDYFSGETR  400
CVS        QVRSLNATVIAACAPHEMSVLGGYLGEEFFGKGTFERRFFRDEKELQEYEAAELTKSDVALADDGTVNSDDEDYFSGETR  400
Nishighara QIRSLNATVIAACAPHEMSVLGGYLGEEFFGRGTFERRFFRDEKNLQEYEAAELTKTDVALADDGTVLSDDEDYFSGEAR  400
RC-HL      QIRSLNATVIAACAPHEMSVLGGYLGEEFFGRGTFERRFFRDEKELQEYEAAELTKTDVALADDGTVLSDDEDHFPGEAR  400

              *****:*****************************************.**
SAD        SPEAVYTRIMMNGGRLKRSHIRRYVSVSSNHQARPNSFAEFLNKTYSSDS  450
Viverrid   SPEAVYTRIMMNGGRLKRSHIRRYVSVSSNHQARPNSFAEFLNKTYSTDS  450
PV         SPEAVYTRIMMNGGRLKRSHIRRYVSVSSNHQARPNSFAEFLNKTYSNDS  450
CVS        SPEAVYTRIMMNGGRLKRSHIRRYVSVSSNHQARPNSFAEFLNKTYSNDS  450
Nishighara GPEAVYARIMMNGGRLKRSHIRRYVSVSSNHQARPNSFAEFLNKTYSSDS  450
RC-HL      GPEAVYARIMMNGGRLKRSHIRRYVSVSSNHQARPNSFAEFLNKTYSSDS  450
```

Fig. 2.4A Multiple alignment of the amino acid sequences of the N proteins of different rabies virus strains, generated by the ClustalX program, as described under Materials and Methods. The phosphorylated serine at position 389 and threonine at position 375 are indicated in the block. The black and gray bars under the alignmets indicate the kinase recognition sites for Protein Kinase C and Casein Kinase II, respectively, as predicted by PROSITE. The red bar indicate the ssRNA-binding domain of the rabies virus N protein, and the green bar

```
              ********************************************:*********.***** ****** *..: .::.: **.*****
CVS11         MSKIFVNPSAIRAGLADLEMAEFTVDLINRNIEDNQAHLQGEPIEVDNLPEDMKRLHLDDEKSSNLGEMVRVGEGKYRED        80
Pasteur       MSKIFVNPSAIRAGLADLEMAEETVDLINRNIEDNDAHLQGEPIEVDNLPEDMKRLHLDDEKSSNLGEMVRVGEGKYRED        80
ERA           MSKIFVNPSAIRAGLADLEMAEETVDLINRNIEDNQAHLQGEPIEVDNLPEIMGRLHLDDGKSPNPGEMAKVGEGKYRED        80
Pitman-Moore  MSKIFVNPSAIRAGLADLEMAEETVDLINRNIEDNQAHLQGEPIEVDNLPDDMGRLHLDDGKSPNPGEMAKVGEGKYRED        80
SAD           MSKIFVNPSAIRAGLADLEMAEETVDLINRNIEDNQAHLQGEPIEVDNLPADMGRLHLDDGKSPNHGEIAKVGEGKYRED        80
Viverrid      MSKIFVNPSAIRAGLADLEMAEETVDLINRNIEDNQAHLQGEPIEVDSLPEDMRRLHLDDRKPSDLSDMVKSGESKYRED        80

              ********* :******:.*********************.***.***:***** * .:***********:* :***:.:
CVS11         FQMDEGEDPNLLFQSYLDNVGVQIVRQMRSGERFLKIWSQTVEEIVSYVTVNFPNPPRRSSEDKSTQTTGRELKKETTSA       160
Pasteur       FQMDEGEDPNLLFQSYLDNVGVQIVRQMRSGERFLKIWSQTVEEIVSYVTVNFPNPPRRSSEDKSTQTTGRELKKETTSA       160
ERA           FQMDEGEDPSFLFQSYLENVGVQIVRQMRSGERFLKIWSQTVEEIISYVAVNFPNPPGKSSEDKSTQTTGRELKKETTPT       160
Pitman-Moore  FQMDEGEDPSFLFQSYLENVGVQIVRQMRSGERFLKIWSQTVEEIISYVAVNFPNPPGKSSEDKSTQTTGRELKKETTPT       160
SAD           FQMDEGEDPSFLFQSYLENVGVQIVRQMRSGERFLKIWSQTVEEIISYVAVNFPNPPGKSSEDKSTQTTGRELKKETTPT       160
Viverrid      FQMDEGEDPILLFQSYLDSVGVQIVRQMRSGERFLKIWSQTAEEIISYVAVNFPNRPGKSSEDKSTQTTGQEPRKETASA       160

              *****.*.***** ** *.************:** ****************:***************:***********:*
CVS11         RSQRESQPSKARMVAQVAPGPPALEWSATNEEDDLSVEAEIAHQIAESFSKKYKFPSRSSGIFLYNFEQLKMNLDDIVKE       240
Pasteur       RSQRESQPSKARMVAQVAPGPPALEWSATNEEDDLSVEAEIAHQIAESFSKKYKFPSRSSGIFLYNFEQLKMNLDDIVKE       240
ERA           RSQRESQSSKARMAAQTASGPPALEWSATNEKDDLSVEAEIAHQIAESFSKKYKFPSRSSGILLYNFEQLKMNLDDIVKE       240
Pitman-Moore  RSQRESQSSKARMAAQTASGPPALEWSATNEKDDLSVEAEIAHQIAESFSKKYKFPSRSSGILLYNFEQLKMNLDDIVKE       240
SAD           RSQRESQSSKARMAAQIASGPPALEWSATNEEDDLSVEAEIAHQIAESFSKKYKFPSRSSGILLYNFEQLKMNLDDIVKE       240
Viverrid      SSQRESQSSKARMAAQAASGPPALEWSATNEEDPSVEAEIAHQIAESFSKKYKFPSRSSGIFLYNFEQLKMNLDDIVRE       240

              ***************.:*************:******.:*****************
CVS11         AKNVPGVTRLAHDGSKIPLRCVLGWVALANSKKFQLLVEADRLSKIMQDDLNRYTSC                            297
Pasteur       AKNVPGVTRLAHDGSKIPLRCVLGWVALANSKKFQLLVEADKLSKIMQDDLNRYTSC                            297
ERA           AKNVPGVTRLAHDGSKLPLRCVLGWVALANSKKFQLLVESDKLSKIMQDDLNRYTSC                            297
Pitman-Moore  AKNVPGVTRLAHDGSKLPLRCVLGWVALANSKKFQLLVESDKLSKIMQDDLNRYTSC                            297
SAD           AKNVPGVTRLAHDGSKLPLRCVLGWVALANSKKFQLLVESDKLSKIMQDDLNRYTSC                            297
Viverrid      AKNVPGVTRLAHDGSKLPLRCVLGWVALANSKKFQLLVESDRLSKIMQDDLNRYTSC                            297
```

Fig 2.4B Multiple alignments of the amino acid sequences of the P proteins of different rabies virus strains, generated by the ClustalX program, as described in Materials and Methods. The serine residues phosphorylated by the Rabies Virus Protein Kinase at position 63 and 64 are indicated by a solid black box. The serine residues phosphorylated by Protein Kinase C at position 162, 210, 271 are also boxed black. The black bars under the alignments indicate the Protein Kinase C recognition sites as predicted by PROSITE, while the grey bars indicate the Casein Kinase II consensus sequences as predicted by PROSITE. The red bar indicate the N-terminal 19 amino acids involved in P protein binding with the L protein.

2.4 Discussion

In this chapter, the full-length nucleo- and phosphoprotein genes of a South African viverrid rabies virus isolate were cloned and sequenced. The nucleotide sequences were translated and compared to the cognate proteins of other rabies viruses. This was aimed at characterizing structurally important domains of these two proteins and to provide information regarding possible consensus phosphorylation sites.

Functional domains on the viverrid N protein, previously identified for other rabies virus strains, were compared. Amino acid residues 298 to 352 of the Pasteur virus strain is crucial for ssRNA binding and the corresponding amino acids of the viverrid strain showed only one amino acid change at position 328 (threonine in PV and alanine in viverrid. Since threonine is a neutral polar amino acid, therefore uncharged under normal physiological conditions, and alanine is a nonpolar amino acid, there would be minimal structural differences between the N proteins of these two strains and very little difference in their RNA-binding ability. The C-terminal 74 amino acid residues play an important role during complex formation between the N and the P protein. Within this domain, we identified three amino acid differences, of which the substitution of a non-polar Val-379 in CVS with a polar threonine amino acid in the viverrid strain can effect the complex formation between these two recombinant proteins. The N protein amino acid sequence of the Pasteur rabies virus strain has previously been compared with that of VSV (Indiana serotype) (Tordo *et al.*, 1986a). This study reported conserved segments of amino acids between the N proteins of these two strains, and specifically at amino acids 72 to 112 with 41.5% homology, 140 to 150 with 54.5% homology, 225 to 247 with 39% homology and 268 to 302 with 48.5% homology.

The deduced amino acid sequences for both the P and the N proteins were used to identify consensus protein kinase recognition sites that can potentially be recognized and phosphorylated by these kinases. Three of the four identified PKC recognition sites in the phosphoprotein corresponds to known phospho-acceptors (Gupta *et al.*, 2000): *viz* Ser-162, Ser-210 and Ser-271. Although PKC has been reported to phosphorylate the rabies virus phosphoprotein, the known phospho-acceptors of the nucleoprotein, Ser-389

and Thr-375, do not correspond to the identified PKC recognition sites (Gupta *et al.*, 2000). Anzai *et al.* (1997) identified these phospho-amino acids with a monoclonal antibody directed against a specific phosphatase-sensitive epitope as well as through amino acid hydrolysis. It is very likely that there are other phosphorylated amino acids besides Ser-389 and Thr-375 in the Nucleoprotein, with other kinases involved in phosphorylating the rabies virus N protein.

With the use of the PROSITE database and software, nine Casein kinase II consensus sequences in the N protein amino acid sequence were identified. Two of these corresponded to the location of the known phospho-acceptors, amino acid residues 375-378 (TKTD) and 389-392 (SDDE). Gupta *et al.* (2000) reported that, in addition to several isomers of PKC, the rabies virus P protein are also phosphorylated by a unique heparin-sensitive non-CKII protein kinase. This enzyme appears to be selectively packaged within the matured virions, and were designated RVPK (rabies virus protein kinase). The authors were unable to identify this protein kinase by comparing its biochemical properties with those of other known cellular protein kinases in protein data banks. This work by Gupta *et al.* (2000) open the door for further studies on the phosphorylation of rabies virus proteins and its role during virus transcription and replication.

CHAPTER 3

Expression of the nucleo- and phosphoproteins in an eukaryotic and prokaryotic expression system

3.1 Introduction

Due to their essential roles in virus replication and transcription, the N and P proteins of rabies viruses have been studied in some detail during recent years. These studies have been facilitated by the ability to express the respective genes in a variety of different eukaryotic, mammalian and prokaryotic expression systems. An important finding emanating from these studies was the fact that both the N and P proteins are phosphorylated by cellular kinases (Anzai *et al.*, 1997; Gupta *et al.*, 2000). Although the phosphorylated amino acid residues have been identified (Anzai *et al.*, 1998; Chenik *et al.*, 2000), knowledge regarding the functional significance of phosphorylation of these proteins is still limited.

The rabies virus N and P proteins have been cloned and expressed by means of recombinant baculoviruses (Fu *et al.*, 1991, 1994; Iseni *et al.*, 1998; Prehaud *et al.*, 1990, 1992; Reid-Sanden *et al.*, 1990), as vaccinia virus recombinants (Chenik *et al.*, 1994) and in *E. coli* (Goto *et al.*, 1995; Gupta *et al.*, 1997, 2000; Takamatsu *et al.*, 1998; Gigant *et al.*, 2000). A general problem encountered in expressing the N protein is the high level of insolubility of the expressed protein, irrespective of the expression system used. In the case of the N protein of VSV, the protein has been reported to rapidly aggregate into insoluble high-molecular weight complexes (Sprague *et al.*, 1983). Using *E. coli* as an expression host, Das and Banerjee (1993) have reported the expressed N protein of VSV to be insoluble, but it could be solubilized in a 1 M NaCl solution. Several groups (Fu *et al.*, 1994; Chenik *et al.*, 1994; Gupta and Banerjee, 1997; and Prehaud *et al.*, 1992) have reported that co-expression of the rabies virus N and P proteins led to complex formation which maintains the N protein in a soluble form.

Since the primary aim of this study was to investigate the functional significance of rabies virus N and P protein phosphorylation, these proteins would have to be expressed to high levels, as soluble proteins, and preferentially in phosphorylated and unphosphorylated forms. Two different expression systems were subsequently chosen

based on their ability or inability to perform post-translational modifications such as phosphorylation.

The first system selected for expression of the rabies virus N and P genes was the eukaryotic baculovirus expression system. The baculovirus expression vector system can express high levels of heterologous genes, and in most cases the recombinant proteins are processed, modified and targeted to their appropriate cellular locations where they are functionally similar to the authentic protein (Galarza *et al.*, 1992; Theron *et al.*, 1996). The successful expression of proteins relies on cloning the heterologous gene under the transcriptional control of the promoters of two dispensable baculovirus gene products, p10 and polyhedrin. These proteins are expressed to very high levels in baculovirus-infected cells and their replacement by almost any heterologous gene result in expression of the recombinant genes to similarly high levels. The construction of recombinant baculoviruses have been greatly facilitated by the introduction of the BAC-to-BAC™ baculovirus expression system (Life Technologies, BRL Gibco) in which foreign genetic material is introduced into the viral genome within *E. coli* cells (Luckow *et al.*, 1993). Recombinant Bacmids are obtained following transposition of a mini-Tn7 element from the donor plasmid to the mini-*att*Tn7 attachment site on the Bacmid DNA. The Tn7 transposition functions are provided *in trans* by a helper plasmid. The mini-Tn7 contains an expression cassette consisting of a gentamycin gene, baculovirus-specific promoters, an extensive multiple cloning site, and an SV40 poly(A) signal. Subsequently, only recombinant Bacmid DNA is introduced into insect cell culture.

The expression of large amounts of protein from a cloned gene introduced into *E. coli*, has been of great value in the purification, localization and functional analysis of proteins. Additional advantages of expressing genes in *E. coli* are that it utilizes general laboratory protocols, and is much cheaper than eukaryotic tissue culture systems. Although prokaryotic cells do not possess the appropriate machinery required to perform post-transcriptional modification of expressed heterologous proteins, they do, however, facilitate studies regarding the significance or function of such modifications present in the appropriately modified proteins. We have elected to express the rabies virus N and P

genes as fusion proteins in *E. coli* using the pGEX plasmid expression vectors (Smith and Johnson, 1989). Each of these plasmids contain the IPTG-inducible *tac* promoter which directs the high level synthesis of heterologous proteins as fusions with the C-terminus of Sj26, a 27.5 kDa glutathione *S*-transferase (GST) protein encoded by the parasitic helminth *Schistosoma japonicum*. In the absence of IPTG, the plasmid-encoded *lacI*^{q} allele is efficient in repressing transcription from the *tac* promoter, thus preventing the accumulation of foreign proteins which may lead to growth retardation and cell death. Furthermore, the recombinant fusion proteins can easily be purified from crude bacterial lysates by glutathione affinity chromatography under non-denaturing conditions.

Therefore, the aims of this part of the study were to express the full-length N and P genes of a viverrid rabies virus strain, singularly and in combination, in insect cells by means of recombinant baculoviruses, and secondly to express the full-length N and P genes in *E. coli*, as recombinant GST fusion proteins.

3.2 Materials and Methods

3.2.1 Construction of recombinant baculoviruses

3.2.1.1 Construction of the recombinant Bacmid donor vectors (pFastBac1 and pFastBac™DUAL)

- *Cloning of the full-length rabies virus N and P genes into the pFastBac1 donor vector*

To enable cloning of the N gene into pFastBac1, both the recombinant NpGEM and the pFastBac1 plasmid DNA were restricted with BamH1 and EcoR1. The 1.4 kb full-length N gene and restricted pFastBac1 vector DNA were purified from an agarose gel, ligated overnight and then transformed into competent *E. coli* JM109 cells. Similarly, the 0.9 kb P gene was recovered from the recombinant PpGEM plasmid following EcoR1 digestion and then cloned into a EcoR1-digested, dephosphorylated pFastBac1 vector. The vector DNA was dephosphorylated by treatment with 0.5 U Alkaline Phosphatase (Roche) for 15 min at 37°C according to the manufacturer's instructions. Recombinant transformants were selected from LB-agar plates supplemented with ampicillin (100 µg/ml) after overnight incubation at 37°C, and further characterized by restriction digestion following

plasmid DNA extraction. These recombinant plasmids were designated NpFastBac and PpFastBac, respectively.

- *Cloning of the full-length N and P genes into the pFastBac™DUAL donor plasmid*

For co-expression of the full-length N and P genes, the donor plasmid pFastBac™DUAL was utilized. The full-length N gene was recloned into pFastBac™DUAL using the procedures described above. To allow for directional cloning of the P gene into pFastBac™DUAL, the full-length P gene was recovered from PpGEM by EcoR1 digestion and then recloned into an EcoR1-linearised, dephosphorylated pBluescript SKII (-) phagemid vector. Following orientation determination of the cloned P gene by BamH1 digestion, the P gene was recovered by Kpn1 and Sma1 digestion and finally cloned into the identically cut recombinant NpFastBac™DUAL plasmid. The resultant recombinant plasmid was designated N+PpFastBac™DUAL.

3.2.1.2 Transformation of competent *E. coli* DH10BAC™ cells

Competent *E. coli* DH10BAC™ cells, containing the Bacmid genome as well as a helper plasmid, were prepared as previously described (Section 2.2.8) and then transformed with the recombinant donor plasmids as follows. Competent DH10BAC™ cells (100 µl) were mixed with 100 ng recombinant donor plasmid in glass transformation tubes and incubated on ice for 30 min. The cells were heat-shocked for 45 s at 42°C, and chilled on ice for 2 min. Following addition of 900 µl of SOC medium (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM glucose; final pH 7.0), the transformation mixtures were incubated for 4 h at 37°C with agitation. The transformed cells were selected by plating the cells onto LB-agar plates containing kanamycin (50 µg/ml), gentamycin (7 µg/ml), tetracyclin (10 µg/ml), IPTG (100 µg/ml) and X-gal (40 µg/ml). The plates were incubated at 37°C for 24 h after which white colonies were selected and restreaked onto fresh plates in order to verify the phenotype. Recombinant transformants were subsequently selected and inoculated into 5 ml of LB-broth containing antibiotics (50 µg/ml kanamycin, 7 µg/ml gentamycin, 10 µg/ml tetracyclin) and incubated at 37°C for 16 h with shaking.

3.2.2 Analysis of recombinant Bacmid DNA

The cells from 1.5 ml of the overnight cultures were harvested by centrifugation at 15 000 rpm for 1 min in microfuge tubes. The bacterial cell pellets were suspended in 300 µl of Solution I (15 mM Tris-HCl (pH 8.0), 10 mM EDTA, 100 µg/ml RNase A), and after addition 300 µl of Solution II (0.2 N NaOH, 1% (w/v) SDS), the tubes were incubated at room temperature for 5 min. This was followed by the addition of 300 µl of 3 M potassium acetate (pH 5.5) and incubation on ice for 10 min. The precipitate consisting of protein and *E. coli* genomic DNA was removed by centrifugation at 15 000 rpm for 10 min. The Bacmid DNA was precipitated from the supernatant by the addition of 800 µl isopropanol on ice for 10 min. The precipitated Bacmid DNA was collected by centrifugation at 15 000 rpm for 15 min, washed with 70% ethanol and air-dried for 10 min at room temperature. The Bacmid DNA was suspended in 40 µl 1 × TE and stored at –20°C until needed.

- *Analysis of recombinant Bacmid DNA*

A PCR reaction was performed in order to verify transposition of the full-length N and P genes into the Bacmid DNA. The recombinant Bacmid DNA was diluted 10-fold after which 1 µl of the DNA was used as template in the PCR reactions. The PCR was carried out as described in Section 2.2.5, using 10 pMol of JJ1 and PHF for the N and P genes respectively in combination with 25 pMol M13 reverse primer. The presence of both the N and P genes in Bacmid DNA transposed with the P+NpFastBacDual donor plasmid was investigated by using, a multiplex PCR, in which three primers, the M13 reverse primer, the N_3 and the PHF primers were used in one PCR reaction. Conditions for the multiplex PCR were as described in Section 2.2.5.

3.2.3 Transfection of *Spodoptera frugiperda* cells with recombinant Bacmid DNA

Spodoptera frugiperda (Sf9) cells were used for the transfection and propagation of the wild-type and recombinant baculoviruses using Grace's insect medium (Highveld Biological) containing 10% fetal calf serum (FCS). Cell concentrations of viable cells were determined using a haemacytometer and staining with trypan blue.

The *Sf*9 cells were transfected with the purified recombinant Bacmid DNA using Cellfectin™ reagent (Life Technologies), according to the instructions of the manufacturer. *Sf*9 cells were seeded in 35 mm diameter wells (1.2×10^6 cells/well) and following attachment of the cells for 1 h at room temperature, were rinsed twice with Grace's medium, without FCS. A further 1 ml of FCS-free medium was then added to the cells. The recombinant Bacmid DNA (5 μg) was diluted with 100 μl Grace's medium without antibiotic and FCS. In a separate tube, 6 μl of the Cellfectin™ reagent was diluted to 100 μl with Grace's medium without antibiotics and FCS. Just prior to transfection, the two solutions were mixed and added dropwise to the monolayers of *Sf*9 cells. After an incubation period of at least 5 h at 28°C under humid conditions, the transfection mixtures were removed and replaced by 2 ml Grace's medium supplemented with 10% FCS and antibiotics. The cells were then incubated at 28°C for 4 days, after which the supernatants were recovered and used in titration assays to investigate the presence of recombinant baculoviruses. Cells transfected with Bacmid DNA or mock-infected cells were included as controls.

3.2.4 Virus titration and plaque purification

In order to obtain purified single plaques or to determine the virus titre, a method based on the procedures described by Brown and Faulkner (1977), Possee and Howard (1987) and Kitts *et al.* (1990) was used. Serial dilutions (1.0×10^{-1} to 1.0×10^{-9}) of the transfection supernatants were prepared in Grace's medium. *Sf*9 cells were seeded in 35 mm diameter wells at a density of 1.5×10^6 cells/well and after adsorption for 1 h at room temperature, the medium was replaced with 1 ml of the serial dilutions. The virus particles were left to infect the cells for 2 h after which the virus dilutions were removed and the cells gently overlaid with sterile 3% (w/v) low melting temperature agarose at 37°C, diluted 1:1 in Grace's medium. The tissue culture dishes were incubated at 28°C for 4 days in a humid environment. The cells were then stained with 2 ml Neutral Red (100 μg/ml in Grace's medium) for 5 h at 28°C. The liquid overlay was removed and the dishes were incubated overnight before screening the plaques. Putative recombinant translucent plaques were plucked as an agarose plug with a pipette tip, transferred to 1 ml

Grace's medium in Eppendorf tubes, vortexed vigorously and stored at 4°C (*ca.* 10^4 pfu/plug). Following one round of plaque purification, high titre viral stocks were subsequently prepared from single plaques by infecting monolayer cultures at a MOI of 0.01 to 0.1 and harvesting the virus-containing supernatants at 96 h post-infection. The recombinant baculoviruses were stored at 4°C after filter-sterilization of viral stocks and virus titres of the prepared virus stocks was determined as described above.

3.2.5 Expression of the N and P proteins by recombinant baculoviruses

Monolayers of *Sf*9 cells in 75 cm^3 tissue culture flasks (1.0×10^7 cells/flask) were mock-infected or infected with wild-type Bacmid or recombinant Bacmid baculovirus recombinants at a MOI of 10 and incubated at 28°C. The cells were harvested from the bottom of the flasks at 72 h post-infection and collected by low speed centrifugation (2 000 rpm for 10 min) and washed in phosphate-buffered saline (1 × PBS). The cell pellets were resuspended in 0.01 M STE (1 mM Tris (pH 7.4), 0.1mM EDTA, 10 mM NaCl) and incubated on ice for 10 min. Subcellular fractions were then prepared according to the procedures described by Huismans *et al.* (1987). The cell suspensions were homogenized by 15 strokes with a Dounce homogenizer, the nuclei were removed by centrifugation at 1 000 rpm for 1 min. The supernatants (cytoplasmic or S10 fractions) were recovered and an equal volume 2 × protein solvent buffer (PSB) (50 mM Tris (pH 6.8), 100 mM DTT, 2% (w/v) SDS, 0.1% (w/v) bromophenol blue, 10% (w/v) glycerol) was added to the respective samples. The samples were heated for 5 min in boiling water and sonicated for 1 min prior to analysis by 12% SDS-PAGE.

3.2.6 Cloning of the full-length rabies virus N and P genes into bacterial pGEX expression vectors

- *Construction of recombinant pGEX expression vectors*

To enable cloning of the rabies virus N gene into pGEX-1, both the recombinant NpGEM and pGEX-1 plasmids were restricted with BamH1 and EcoR1. The 1.4 kb full-length N gene and restricted pGEX-1 vector DNA were purified from the agarose gel, ligated overnight and then transformed into competent *E. coli* JM109 cells. Similarly, the 0.9 kb P-gene was recovered from the recombinant PpGEM plasmid following EcoR1 digestion

and then cloned into a EcoR1-digested, dephosphorylated pGEX-1 vector. Recombinant transformants were selected from LB-agar plates supplemented with ampicillin (100 µg/ml) after overnight incubation at 37°C, and further characterized by restriction digestion following plasmid DNA extraction. The orientation of the cloned P gene was determined by BamH1 restriction mapping. The recombinant plasmids were designated NpGEX and PpGEX, respectively.

3.2.7 Expression of N and P proteins as fusion proteins in *E. coli*

Individual colonies of *E. coli* JM109 transformed with recombinant pGEX-1 and parental pGEX-1 plasmids were inoculated into 1 ml LB-broth supplemented with ampicillin and cultured overnight at 37°C for 16 h. The overnight cultures were diluted 1:10 in 10 ml of fresh medium and incubated for 2 h at 37°C. To each culture, IPTG was added to a final concentration of 1 mM in order to induce expression of the recombinant fusion proteins (Smith and Johnson, 1989). After a further 6 h of culturing, the cells were harvested by centrifugation at 4 000 rpm for 5 min at 4°C in a Sorvall® RC-5B Refrigerated Superspeed centrifuge (Du Pont Instruments). A aliquot (10 µl) of the cell lysates were added to an equal volume of 2 × PSB, boiled for 5 min and after sonication analyzed by SDS-PAGE.

3.2.8 Purification of soluble N and P fusion proteins

- *Solubilization of induced proteins*

To increase the solubility of the expressed recombinant fusion proteins, the pellet of each 10 ml-induced culture was resuspended in 1 ml ice-cold Mix 1 (20% (w/v) sucrose, 100 mM Tris-HCl (pH8), 10 mM EDTA) and incubated on ice for 1 h, after which 10 mM DTT, 1% (v/v) Tween-20, 1% (v/v) Triton X-100 and 1 × PBS (final concentrations) was added. The suspensions were either incubated on ice for 90 min or overnight at 4°C. The suspensions were further homogenized by using two pulses of 20 s each of a probe Ultrasonic Homogenizer (Cole-Palmer Instruments). Following centrifugation at 4 000 rpm for 3 min, the supernatants containing the soluble proteins were recovered and stored at –20°C, until later use.

- *Glutathione agarose affinity chromatography*

The recovered supernatants were purified by affinity chromatography with glutathione agarose (Sigma-Aldrich) as described by Smith and Johnson (1988). The glutathione agarose (0.087 g) was pre-swollen in 7 ml MT-PBS buffer (150 mM NaCl, 16 mM Na_2HPO_4, 4 mM NaH_2PO_4; pH 7.3), washed twice with MT-PBS and stored at 4°C as a 50% (v/v) solution. The fusion proteins were purified by mixing 200 µl of the prepared glutathione agarose with 250 µl of the P-GST or 400 µl of the N-GST cleared lysates, respectively. To facilitate proper mixing, the reactions were diluted to 800 µl in MT-PBS, and incubated at room temperature for 40 min on an orbital shaker. After adsorption, the beads were collected by brief centrifugation at 2 000 rpm for 2 min and washed 3 times with 1 ml MT-PBS each. In the case of N-GST protein purification, this process was repeated twice prior to elution of the fusion protein. The fusion proteins were eluted from the agarose beads by competition with free glutathione in two 2 min washes with 1 bead volume of MT-PBS, followed by 1 wash with 1 bead volume of 5 mM reduced glutathione, in 50 mM Tris-HCl (pH 8.0), at a final pH of 7.5. The proteins were analyzed by SDS-PAGE and Western blot analysis.

3.2.9 Detection of the expressed proteins

- *SDS-polyacrylamide gel electrophoresis (SDS-PAGE)*

Protein samples were analyzed by electrophoresis in the discontinuous gel system as described by Laemmli (1970). The 12% separating gel (1.5 M Tris-HCl (pH 8.8), 0.1% (w/v) SDS, 12 % (w/v) acrylamide) and 5% stacking gel (1.0 M Tris-HCl (pH6.8)l, 0.1% (w/v) SDS, 5% (w/v) acrylamide) were polymerized by the addition of 0.001% TEMED and 0.1% ammonium persulfate. Electrophoresis was performed in a Hoefer Mighty Small™ electrophoresis unit at a constant voltage of 10 V/cm gel in 1 × TGS buffer (25 mM Tris, 250 mM glycine (pH 8.3), 0.1% (w/v) SDS). After electrophoresis, the gels were stained for 10 min with a Coomassie Brilliant Blue solution (prepared in 50% (v/v) methanol and 10% (v/v) acetic acid) and then counter-stained for 1 h in a solution containing 25% (v/v) methanol and 10% (v/v) glacial acetic acid.

- *Western blot analysis* [1]

After SDS-PAGE, the gel and a Hybond-C nitrocellulose membrane (Life Technologies, Inc.), cut to the same size as the gel, were equilibrated for 15-30 min in transfer buffer (0.2 M Tris, 1 M Glycine). The proteins were transferred to the membrane at 28 V, 120 mA for 90 min using the TE Series Transfor Electrophoresis unit (Hoefer Scientific Instruments). Following transfer, the gel was recovered and stained with Coomassie Brilliant Blue to determine the efficiency of the transfer process. The membrane was rinsed in 1 × TBS (20 mM Tris, 137 mM NaCl, 1 M HCl; pH 7.6) for 5 min and a non-specific binding sites were blocked by immersing the membrane in blocking solution (1 × TBS containing 1% (w/v) fat-free milkpowder) for 30 min at room temperature on an orbital shaker. The primary antibody was added to the membrane and incubated with gentle agitation for 1 h at room temperature. The unbound primary antibodies were removed by first rinsing the membrane in washing buffer (0.05% (v/v) Tween-20 in 1 × TBS), and then by washing the membrane once for 10 min and twice for 5 min each in washing buffer. The secondary antibody was added to the membrane and incubated at room temperature for 1 h with gentle agitation. The membrane was washed as described above. Calorimetric detection of immunoreactive proteins was performed by immersing the membrane in ice-cold peroxidase enzyme substrate (60 mg 4-chloro-1-naftol in 20 ml ice-cold methanol, mixed with 100 ml 1 × TBS containing 60 µl H_2O_2). Once the bands become visible, the membrane was rinsed with water and air-dried.

Primary antibodies used in the analysis included a rabies virus-specific serum (diluted 1:100 in TBS) as well as monoclonal and polyclonal anti-GST antibodies (Santa Cruz Biotechnology, Inc.). The anti-GST antibodies were both diluted 1:1000 in 1 × TBS prior to use. For detection of the bound primary antibodies, horseradish peroxidase-conjugated Protein A (Sigma-Aldrich) was used as the secondary antibody and diluted 1:1000 in 1 × TBS.

---

[1] The Rabies Unit at the Onderstepoort Veterinary Research Institute, kindly donated the rabies virus serum used in this analysis.

3.3 Results

3.3.1 Expression and co-expression of the rabies virus N and P proteins in insect cells using the BAC-to-BAC™ baculovirus expression system

- *Construction of donor plasmids containing full-length N and P genes*

Full-length copies of the N and P genes were obtained by restriction enzyme digestion of recombinant pGEM®-T Easy vectors (Chapter 2). The N gene was recovered from NpGEM by digestion with both EcoR1 and BamH1 to allow for directional cloning into similarly digested pFastBac1 and pFastBac™DUAL vectors. The P gene was recovered with EcoR1 restriction enzyme digestion from PpGEM for cloning into linearised and dephosphorylated pFastBac1 vector. The resultant recombinant plasmids were selected and characterized by agarose and restriction enzyme analysis. The results are shown in Figures 3.1 and 3.2, respectively. Digestion of the recombinant NpFastBac and NpFastBac™DUAL plasmids with BamH1 and EcoR1 resulted in excision of a 1.4 kb fragment, corresponding in size to the full-length N gene (Fig. 3.1, lanes 6 and 8, respectively). EcoR1 digestion of recombinant PpFastBac1 plasmids excised a DNA fragment of *ca* 900 bp, corresponding with the expected size of the full-length P-gene (Fig 3.2, lane 5). The recombinant PpFastBac plasmids were also digested with BamH1, which recognizes one site in the vector and cuts asymmetrically in the P gene 300 nt from the 5' end, in order to verify the transcriptional orientation of the cloned insert (not shown).

For construction of the N+PpFastBacDUAL vector the P gene was first cloned into pBluescript SKII (-) to facilitate directional cloning of the P gene. The full-length P gene was recovered from PpGEM by restriction with Sma1 and Kpn1 and cloned into pBluescript SKII (-). The P gene was subsequently recloned into the Sma1 and Kpn1 sites of the recombinant NpFastBacDUAL plasmid to generate the P+NpFastBac™DUAL donor plasmid. Recombinant P+NpFastBac™DUAL plasmids were screened by digestion with EcoR1 which resulted in digestion products of *ca* 0.9 kb, 1.7 kb and 5.0 kb (Fig 3.2). The sizes of these fragments corresponded with the expected sizes of the P
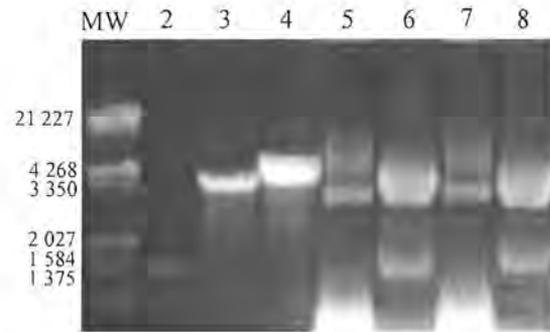
Fig. 3.1 Agarose gel electrophoretic analysis of the recombinant NpFastBac and NpFastBacDUAL donor plasmids. Lane MW, Molecular weight marker; Lane 2, purified N gene amplicon; Lane 3, pFastBac1 plasmid restricted with BamH1 and EcoR1; Lane 4, pFastBacDUAL plasmid digested with BamH1 and EcoR1; Lane 5, uncut recombinant NpFastBac; Lane 6, recombinant NpFastBac digested with BamH1 and EcoR1; Lane 7, uncut recombinant NpFastBacDUAL; Lane 8, recombinant NpFastBacDUAL digested with BamH1 and EcoR1. The sizes of molecular weight marker, phage Lambda restricted with EcoR1 and HindIII, are indicated to the left of the figure in bp.
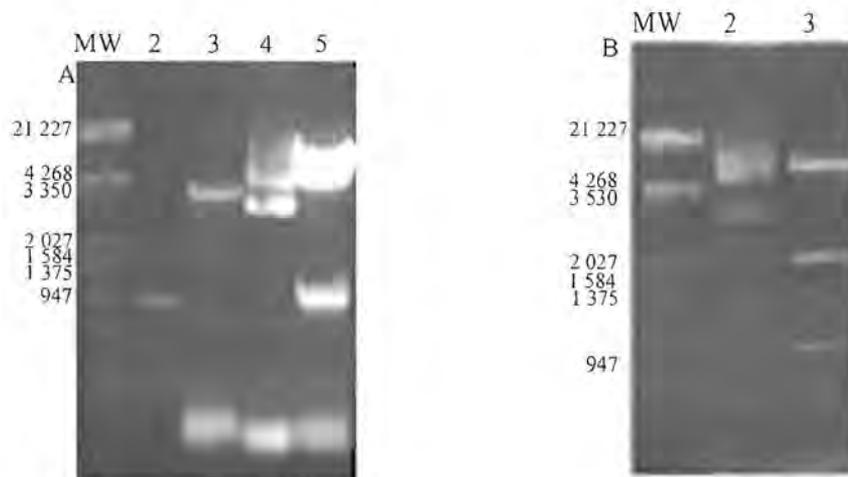
Fig. 3.2 Agarose gel electrophoretic analysis of (A) recombinant pFastBac and (B) P+NpFastBacDUAL donor plamids. (A) Lane MW, Molecular weight marker; Lane 2, purified P gene amplicon; Lane 3, pFastBac1 digested with EcoR1; Lane 4, uncut recombinant PpFastBac; Lane 5, recombinant PpFastBac digested with EcoR1. (B) Lanes: MW, Molecular weight marker; 2, uncut N+PpFastBacDUAL; 3, recombinant N+PpFastBacDUAL digested with EcoR1. The sizes of molecular weight marker, phage Lambda digested with EcoR1 and HindIII, are indicated on the left of the figure.

gene, the N gene together with 300 bp of flanking sequences, and the remainder of the pFastBac™DUAL vector.

- *Construction of recombinant baculoviruses*

Recombinant donor plasmids NpFastBac, PpFastBac and P+NpFastBacDual were used for transposition to the Bacmid DNA in *E. coli* DH10BAC™ cells. During the site-specific transposition of the recombinant donor plasmids, the mini-Tn7 cassette is inserted into the mini-*att*Tn7 attachment site on the Bacmid, thereby disrupting expression of the lacZα peptide. Consequently, white colonies containing the recombinant Bacmid could be distinguished from blue colonies (harboring the non-recombinant Bacmid), when plated in the presence of X-gal and IPTG. The high molecular weight recombinant Bacmid DNA was isolated from *E. coli* DH10BAC™ cells, and the presence of the rabies virus N and P genes in the Bacmid DNA was investigated by a PCR approach. The primers used in this assay were the gene-specific primers, JJ1 and PHF, that anneals to the 5' end of the N and P genes respectively (Table 2.1), and the universal M13 reverse primer that anneals on the mini-*att*Tn7 site within the *lacZα* gene of the Bacmid DNA. By making use of recombinant Bacmid DNA transposed with the NpFastBac donor plasmid as template for the PCR with primers JJ1 and M13, a single discreet band of the expected size (2 kb) was obtained. The presence of the P gene in the recombinant Bacmid DNA was indicated by an amplicon of 1.5 kb following PCR amplification. A multiplex PCR protocol was used to verify the presence of both the N and the P genes in the N+PpFastBacDual transposed Bacmid DNA. The N_3 primer that anneals to the 3' end of the N gene and the PHF primer that anneals to the 5' end of the P gene were used in combination with the M13 reverse primer in one PCR reaction, resulting in two amplicons of 3.2 kb and 1.5 kb, respectively. The controls (no template and Bacmid transposed with pFastBac) were amplified using the M13 forward and reverse primers, and resulted in no amplification and a 2.3 kb amplicon respectively.

The recombinant Bacmid DNA was subsequently used to transfect *S. frugiperda* cells using Cellfectin™ reagent. Progeny viruses resulting from each transfection were obtained by plaque assay of the recovered supernatant fluids. Following plaque

purification, single plaques were selected and used in the preparation of recombinant viral stocks of which the titres were determined by plaque assays. These recombinant baculoviruses were designated BacN, BacP and BacN+P, respectively.

- *Analysis of proteins synthesized in infected S. frugiperda cells*

To investigate whether the rabies virus N and/or P proteins were expressed in *Sf*9 cells by the baculovirus recombinants, monolayers of *Sf*9 cells were mock-infected and infected at a MOI of 10 with the wild-type and respective recombinant baculoviruses. The cells were harvested at 96 h post-infection and subcellular cytoplasmic and particulate fractions were prepared, as described under Materials and Methods (Section 3.2.5). The fractions were then subjected to SDS-PAGE. Analysis of the Coomassie-stained gels indicated that no unique recombinant proteins were visible following comparison of the recombinant baculovirus-infected cells to the control mock- or wild-type infected cells (results not shown). The presence of rabies virus N and P proteins could however be indicated by Western blot using a rabies virus-specific serum as the primary antibody (Fig. 3.4). The results indicated that the serum reacted with a protein of *ca.* 55 kDa in size, which was present in the particulate fraction of cells infected with the BacN baculovirus recombinant (Fig.3.4, lane 2). No proteins that reacted with the rabies virus antiserum could be identified in the cytoplasmic fractions (Fig. 3.4, lane 6). The antiserum also reacted strongly with a 36 kDa protein present in the cytoplasmic and particulate fractions prepared from BacP-infected cells (Fig. 3.4, lanes 3 and 7). In the case of the recombinant baculovirus BacN+P, the virus-specific serum reacted with both the 36 kDa and 55 kDa proteins present in the particulate fraction prepared from the infected cells (Fig. 3.4, lane 4). However, only the 30 kDa protein reacted with the serum in the corresponding cytoplasmic fraction (Fig. 3.4, lane 8).

It could be concluded that the rabies virus N and P proteins were indeed expressed by the baculovirus recombinants, albeit in very low amounts as the proteins could not be observed in the Coomassie-stained gels. Furthermore, the N protein appears to be insoluble as it could only be observed in the particulate fraction prepared from the
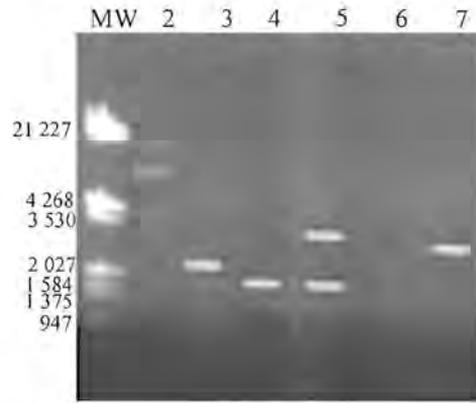
Fig. 3.3 Agarose gel electrophoresis of PCR amplicons, in order to verify the presence of rabies virus N and P genes in the recombinant Bacmid DNA. Lanes: MW, Molecular weight marker; 2, purified Bacmid DNA; 3, N gene in reombinant Bacmid DNA amplified with primers JJ1 and M13; 4, P gene in recombinant Bacmid DNA amplified with primers PHF and M13; 5, Multiplex PCR of the N and P genes in the recombinant Bacmid DNA transposed with the P+NpFastBac DUAL vector; 6, no template DNA control; 7, Bacmid DNA with transposed pFastBac1 plasmid amplified with the pUC forward and reverse primers. The sizes of molecular weight marker, phage Lambda restricted with EcoR1 and Hind III, are indicated to the left of the figure in bp.

Fig. 3.4 Western blot analysis of recombinant baculovirus-expressed rabies virus N and P proteins using rabies virus-specific serum as primary antibody. Lane 1, Mock-infected $Sf9$ cells; Lanes 2 and 6, BacN-infected cells; Lanes 3 and 7, BacP-infected cells; Lanes 4 and 8, BacN+P-infected cells; Lane 5, Wild-type baculovirus-infected cells. P10 refers to particulate fraction prepared from the infected cells, while S10 refers to the corresponding cytoplasmic fraction. Sizes of the protein molecular weight marker is indicated to the left of the figure in kDa. The * indicates the recombinant N protein, and # indicates the recombinant P protein expressed by the recombinant baculoviruses.

recombinant baculovirus infected cells. The assays used to analyze the co-expression of the N protein with the P protein were not sensitive enough to indicate whether complex formation did occur.

3.3.2 Expression of the rabies virus N and P proteins in *E. coli*

• *Cloning of the full-length N and P genes into the pGEX-1 expression vectors*

In order to express the N and the P genes in *E. coli*, the respective genes were cloned into the pGEX-1 expression vector (Smith and Johnson, 1988), thereby positioning the respective genes in the correct reading frame for translation of N-GST and P-GST fusion proteins. The full-length N gene was recovered from the NpGEM clone (Chapter 2) by restriction with EcoR1 and BamH1. The gene was agarose gel-purified and then directionally cloned into the EcoR1 and BamH1 sites of the pGEX-1 vector. The full-length P gene was recovered from the PpGEM clone (Chapter 2) by restriction with EcoR1, agarose gel-purified and then cloned into the dephosphorylated EcoR1 site of the pGEX-1 vector. Transformants were selected and cultured overnight at 37°C after which the extracted plasmids were subjected to restriction enzyme analysis. Restriction of the recombinant NpGEX plasmids with BamH1 and EcoR1 (Fig. 3.5A, lane 5) and restriction of recombinant PpGEX plasmids with EcoR1 (Fig. 3.5B, lane 3) indicated that the N and P gene-specific inserts were of full-length. The correct transcriptional orientation of the cloned P gene was verified by further restriction mapping with BamH1 (not shown).

• *Expression of recombinant GST fusion proteins*

For bacterial expression, overnight cultures of *E. coli* JM109 cells containing either the parental or recombinant pGEX-1 vectors were diluted and grown in the presence of IPTG. After incubation at 37°C, the cells were harvested and protein samples were prepared for analysis by SDS-PAGE as described in Section 3.2.9. Analysis of the Coomassie-stained gel (Fig. 3.6) indicated that *E. coli* cells transformed with pGEX-1 only produced the 27.5 kDa GST protein. Induced cultures of *E. coli* containing the recombinant NpGEX clone, indicated the synthesis of a polypeptide of *ca.* 82 kDa, while *E. coli* containing the recombinant PpGEX clone indicated the synthesis of a 57 kDa
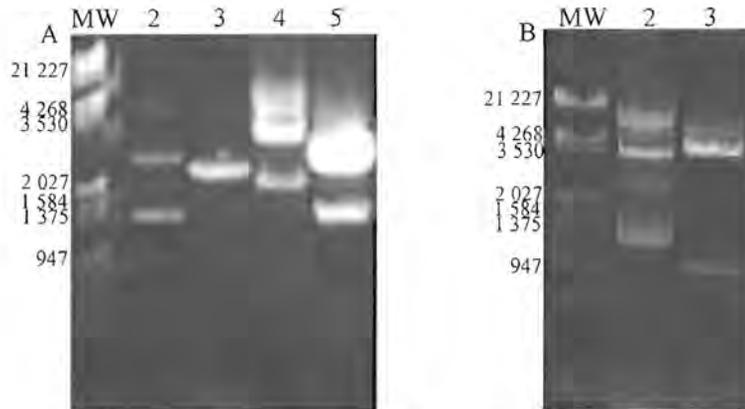
Fig. 3.5 Agarose gel electrophoretic analysis of the recombinant (A) NpGEX and (B) PpGEX plasmids. (A) Lanes: MW, Molecular weight marker; 2, uncut pGEX-1; 3, pGEX-1 digested with BamH1 and EcoR1; 4, uncut recombinant NpGEX; 5, recombinant NpGEX digested with BamH1 and EcoR1. (B) Lanes: MW, Molecular weight marker; 2, recombinant PpGEX (uncut); 3, recombinant PpGEX digested with EcoR1. Sizes of the molecular weight marker phage, Lambda digested with EcoR1 and Hind III are indicated on the left of the figure in bp.
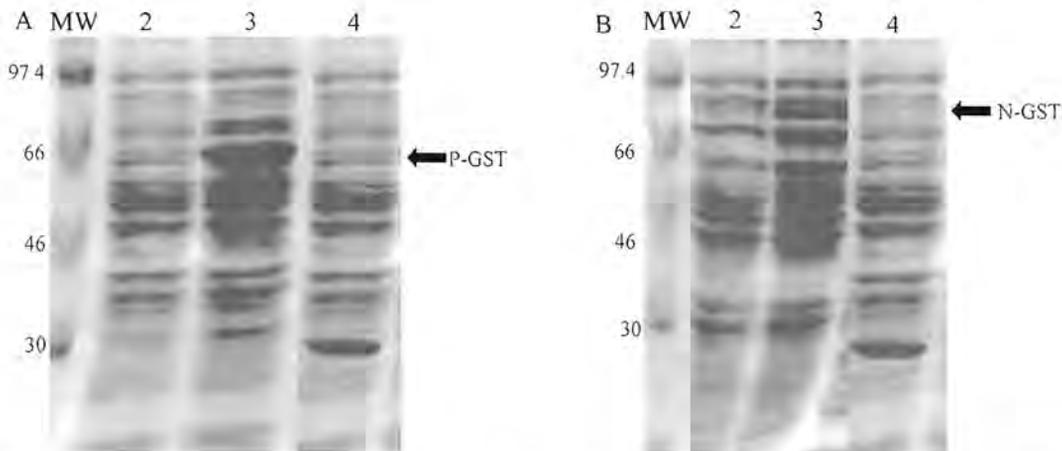
Fig. 3.6 Coomassie-stained SDS-PAGE gel of the *E. coli*-expressed rabies virus (A) P-GST and (B) N-GST fusion proteins. (A) Lanes: MW, Rainbow molecular weight marker (in kDa); 2, IPTG-induced *E. coli* culture; 3, IPTG-induced *E. coli* culture transformed with PpGEX; 4, IPTG-induced *E. coli* culture transformed with parental pGEX; (B) Lanes: MW, Rainbow molecular weight marker (in kDa); 2, IPTG-induced *E. coli* culture; 3, IPTG-induced *E. coli* culture transformed with NpGEX; 4, IPTG-induced *E. coli* culture transformed with parental pGEX. The position of the P and N fusion proteins are indicated by the arrows.

protein. In both cases, the sizes of these unique proteins corresponded with the predicted sizes of the N-GST and P-GST proteins. A duplicate, but unstained SDS-PAGE gel was subjected to Western blotting using a polyclonal anti-GST antibody. The antibody reacted with the overproduced 82 and 57 kDa polypeptides as well as the control 27.5 kDa GST protein (Fig. 3.7A). To verify that the respective fusion proteins included the rabies virus N and P proteins, Western blot analysis was performed using a rabies virus-specific serum (Fig. 3.7B). The antiserum reacted strongly with only the overproduced 82 kDa and 60 kDa proteins, indicating the specific expression of the rabies virus N and P proteins as GST fusion proteins.

Western blot analysis using the anti-GST and rabies virus serum, also indicated that the 82 kDa N-GST protein as well as proteins migrating slightly faster than the N-GST fusion protein, reacted with the antisera (Fig. 3.7A and B, lanes 5). Since no cross-reaction of the antisera with host cell proteins were observed, it was concluded that these smaller proteins represented degraded derivatives of the 82 kDa N-GST protein. Similar degraded derivatives of the 57 kDa P-GST fusion protein was also occasionally observed. Such instability of expression products appears to be a common phenomenon in *E. coli* (Marston, 1988).

To determine whether the expressed viverrid rabies virus N-GST and P-GST fusion were present in a soluble form, the GST-fusion proteins were purified from the bacterial extracts by glutathione affinity chromatography and the yields and purity were assessed by SDS-PAGE (Fig. 3.8). Although high concentrations of purified control GST and P-GST fusion proteins could be obtained, the concentration of the purified N-GST fusion protein was much lower. This may have been due to the majority of N-GST protein being expressed in an insoluble form and could thus not be purified under the conditions used. Nevertheless, by expressing the N gene as a GST fusion protein in *E. coli*, a source of soluble N protein was established which could be used in subsequent investigations.
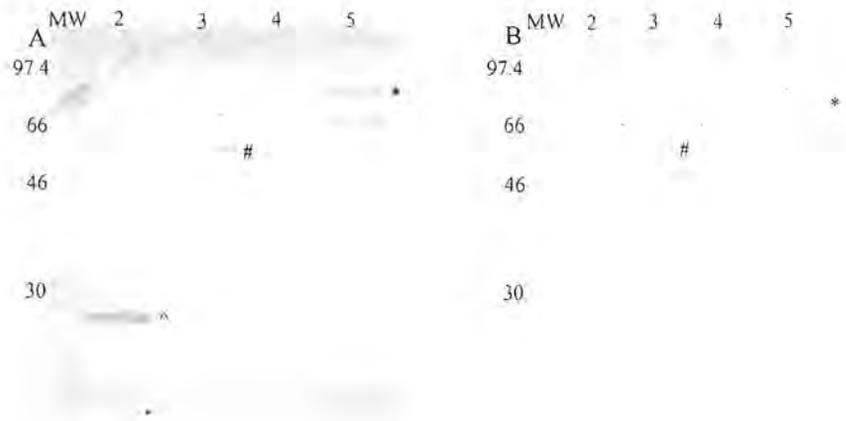
Fig. 3.7 Western blot analysis of recombinant rabies virus N and P fusion proteins using (A) anti-GST antibody and (B) anti-rabies virus serum as primary antibody. Lanes: MW, Rainbow molecular weight marker (in kDa); 2, IPTG-induced *E. coli* culture transformed with pGEX-1; 3, IPTG-induced *E. coli* culture transformed with PpGEX; 4, IPTG-induced *E.coli* culture; 5, IPTG-induced *E. coli* culture transformed with NpGEX. The * indicate the N-GST, the # indicate the P-GST and ^ indicate the GST.
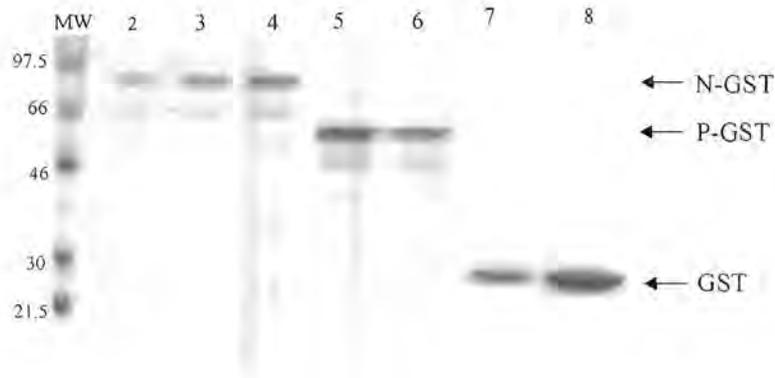
Fig. 3.8 Coomassie-stained SDS-PAGE gel of purified GST fusion proteins following glutathione affinity chromatography. Lane MW, Rainbow molecular weight marker (in kDa); Lanes 2, 3, and 4, purified N-GST protein; Lanes 5 & 6, purified P-GST protein; Lanes 7 & 8, purified GST protein.

3.4 Discussion

We have been able to express the N and P proteins of a South African rabies virus strain in both an eukaryotic and a prokaryotic expression system. Expression of the rabies virus proteins in a BAC-to-BAC™ baculovirus expression system enabled us to express recombinant rabies virus N and P proteins individually or in combination. The baculovirus expression system has been used successfully to express a wide variety of genes. In the case of rabies virus, both the N and P proteins have been expressed in the baculovirus system (Fu *et al.*, 1991, 1994; Iseni *et al.*, 1998; Prehaud *et al.*, 1990, 1992; Reid-Sanden *et al.*, 1990). The synthesis of N and P proteins in cell culture has made significant contributions to the understanding of the function of the individual proteins. The co-expression of these proteins have also been used for studying complex formation between the N and the P proteins (Fu *et al.*, 1994; Chenik *et al.*, 1994), the maturation of the N protein (Kawai *et al.*, 1999) and specificity of RNA encapsidation (Yang *et al.*, 1998).

In this chapter, the N and P proteins of a viverrid rabies virus isolate were expressed in the BAC-to-BAC™ baculovirus expression system. This system offers the advantage of characterizing the viral DNA prior to its introduction into insect cells and thus only modified recombinant baculoviral DNA is introduced into insect cell cultures. Recombinant baculoviruses expressing the N or P genes were subsequently constructed. *S. frugiperda* cells were infected with the respective recombinant baculoviruses after which subcellular extracts were prepared and subjected to SDS-PAGE analysis. It was found that the proteins were expressed to very low levels, as they could not be detected in Coomassie-stained gels. However, proteins corresponding in size to the N and P proteins reacted strongly with an anti-rabies virus-specific serum in Western blot analysis. From the obtained results (Fig. 3.4), it could be concluded that the P protein was expressed in a soluble form, but the expressed N protein was insoluble as it could only be detected in the particulate fraction of infected cells. It has been previously reported that co-expression of the rabies virus N and P proteins maintains the N protein in a soluble form (Yang *et al.*, 1998). To obtain soluble N protein, a baculovirus recombinant capable of co-expressing the viverrid rabies virus N and P proteins was also constructed. The use of a single

recombinant baculovirus for the simultaneous expression of these genes offers certain advantages over co-infection of insect cells with the previously constructed N and P baculovirus recombinants. By co-expressing both proteins in the same cell, the proteins can interact *in vivo* to form N-P protein complexes whereas by co-infection of the cells, not all the cells may be infected by both individual recombinant baculoviruses. It is therefore to be expected that these interactions will be inefficient and the yield of N-P protein complexes may be low. Under the centrifugation conditions used during this study the proteins pelleted and complex formation could not be observed. Optimized centrifugation conditions, sucrose gradient centrifugation, or co-immunoprecipitation would have given more definite answers regarding complex formation between these dual expressed proteins as well as the solubility of the N protein as part of the N-P complex.

The results obtained indicated that although both proteins were being synthesized in the recombinant baculovirus-infected cells, complex formation between the N and P proteins could not be observed. Consequently, the N protein was insoluble and could not be recovered for the cytoplasmic subcellular fraction. The rabies virus genes were placed under the control of the polyhedrin and p10 promoters, since both these promoters are very late promoters these two proteins will be expressed at the same time (O'Reilly *et al.*, 1992).

A further complication that was encountered during co-expression of the rabies virus N and P genes, was the apparent instability of the constructed baculovirus recombinant, in spite of several plaque purifications. First round infection of the insect cells with the recombinant baculovirus did result in expression of both the N and P proteins. However, use of the virus-containing supernatant to re-infect *Sf*9 monolayers resulted in only the P protein being expressed. Such instability with regards to heterologous protein expression using this system has also been noted during the expression of Bluetongue virus and African horsesickness virus gene products (Maree, 2001).

To overcome the problems associated with the BAC-to-BAC™ baculovirus expression system, the use of *E. coli* as an alternative expression host was investigated. The rabies virus N (Goto *et al.*, 1995) as well as VSV N and P proteins (Das and Banerjee, 1993; Gupta and Banerjee, 1997) has previously been expressed in *E. coli*. Although the solubility of the rabies virus N protein when expressed in *E. coli* was not investigated by Goto *et al.* (1995), Das and Banerjee (1993) reported that the bacterial-expressed N protein of VSV was insoluble. The protein could, however, be solubilized in 1 M NaCl. In a subsequent report, Gupta and Banerjee (1997) co-expressed the VSV N and P proteins in *E. coli* and found that the resulting N-P protein complexes were soluble. In order to obtain soluble rabies virus N and P protein, the use of a pGEX expression system was investigated since it has been reported that the expressed heterologous fusion proteins are predominantly soluble (Wang and Zang, 1997; Berg *et al.*, 1998). The N and P genes of a typical viverrid rabies virus were cloned into pGEX-1 after which crude cell extracts of the selected recombinants were analyzed by SDS-PAGE. The respective proteins were synthesized as GST fusion proteins and found to be expressed to high levels. Their identity was subsequently confirmed by Western blot analysis using both anti-GST and rabies virus-specific antisera (Fig 3.7). The rabies virus P-GST fusion protein was found to be soluble and could easily be purified to high levels from bacterial extracts by glutathione affinity chromatography. In contrast, the bacterial-expressed N-GST fusion protein remained largely insoluble. Different approaches were therefore investigated in order to obtain soluble N-GST fusion proteins. These included varying the growth parameters such as lowering the growth temperature to between 20-30°C, inducing protein expression for a shorter period of time and also at a higher cell density for a shorter period of time. Treatment of the bacterial cell lysates with varying concentrations of N-laurylsarcosine was also investigated (Frangioni and Neel, 1993). None of these approaches appeared to be adequate as only a small amount of the soluble N-GST fusion protein could be recovered. However, a higher level of soluble protein could be obtained by treatment of the bacterial cell extracts in a solution containing DTT, Tween-20 and Triton-X100 (Theron *et al.*, 1996). This resulted in soluble N-GST fusion protein, which could subsequently be purified using glutathione affinity chromatography.

In conclusion, the results obtained in this part of the investigation indicated that the viverrid rabies virus N and P genes could be expressed in insect cells infected with the constructed recombinant baculoviruses, and confirmed the integrity of the cloned rabies virus genes. However, the insolubility of the expressed N proteins hampered further use of the recombinant protein. The synthesis of the N and P proteins as GST fusion proteins in *E. coli* provided a source of soluble proteins which were subsequently used for ssRNA- and protein-protein binding studies of which the results will be detailed in the following chapter.

CHAPTER 4

Role of rabies virus nucleoprotein phosphorylation during interaction with single stranded RNA and the phosphoprotein

4.1   Introduction

Rabies virions contain two kinds of phosphorylated proteins, namely the nucleoprotein and phosphoprotein. The nucleoprotein is a 55 kDa RNA-binding protein and represents *ca.* 30% of all virion proteins (Kawai, 1977). The N protein binds to the single-stranded viral genomic RNA *via* the phosphate-sugar backbone of the RNA and exposes the nucleotide bases to the RNA polymerase during transcription and replication (Emerson, 1987; Iseni *et al.*, 2000; Keene *et al.*, 1981; Klumpp *et al.*, 1997). The 36 kDa phosphoprotein is present in the virions as well as in infected cells. Although being present abundantly in the cell, the P protein constitutes less than 6% of the virion (Kawai, 1977). In addition to being a non-catalytic subunit of the viral polymerase, the P protein forms complexes with nascent N protein to prevent it from self-aggregating (Emerson and Schubert, 1987; Bell *et al.*, 1984). It has been suggested that this N-P protein association ensures that N protein is available for encapsidation of the viral genomic RNA and is also responsible for conferring specificity to the RNA encapsidation, thereby preventing binding to non-specific RNAs (Yang *et al.*, 1998).

In the case of VSV, only the P protein is phosphorylated (Sokol and Clark, 1973) and its role during viral replication and transcription has been studied extensively. The VSV P protein has been reported to comprise at least two different phosphorylated components, a hypophosphorylated P1 and a hyperphosphorylated P2 form (Barik and Banerjee, 1991). The interconversion between these two forms of the P protein is a result of phosphorylation and dephosphorylation of phosphate groups which may regulate its functions (Takacs *et al.*, 1992; Barik and Banerjee, 1992b; Chen *et al.*, 1997). The VSV P protein is phosphorylated by cellular casein kinase II (CK II) and a viral L protein-associated kinase (Chen *et al.*, 1997) at two sites located in the N-terminal half and C-terminus of the protein (Barik and Banerjee, 1992a; Chattopadhyay and Banerjee, 1987b). Phosphorylation in both these domains is essential for transcription, but not for replication of the viral genome. Similar to the VSV P protein, the rabies virus P protein also appears to exist in heterogeneous forms (Takamatsu *et al.*, 1998). By making use of an anti-P antibody, a 40 kDa protein (p40) was detected in the virion, but not in the infected cell. The p40 is more phosphorylated than the 37 kDa component of the P

protein (Chenik *et al.*, 1995; Tuffereau *et al.*, 1985). Although the p40 protein is continuously produced in the cell, most of these proteins are immediately dephosphorylated, resulting in p40 becoming almost absent in the cell (Takamatsu *et al.*, 1998). The kinases responsible for rabies virus P protein phosphorylation, namely protein kinase C (PKC) and a novel rabies virus protein kinase (RVPK), as well as the phosphorylation sites, have been identified (Gupta *et al.*, 2000). The RVPK appears to be packaged with the virions and phosphorylation of unphosphorylated *E. coli*-expressed P proteins by this kinase resulted in an alteration of the protein mobility in SDS-PAGE. The protein migrates as a 40 kDa protein, in contrast to phosphorylation with PKC which does not change the mobility of the 37 kDa P protein during electrophoresis (Gupta *et al.*, 2000). Functions of the rabies virus P protein are also thought to be regulated by phosphorylation and dephosphorylation of the protein.

In contrast to the rabies virus P protein, phosphorylation of the N protein has not been well characterized. Although both serine and threonine have been identified as phospho-amino acid residues (Anzai *et al.*, 1997; Yang *et al.*, 1999), the kinase(s) responsible for N protein phosphorylation is not known. In addition, information regarding the role of N protein phosphorylation in regulating its functions is limited. Such investigations have been largely restricted due to an inability to obtain sufficient amounts of soluble N protein for functional studies. Recently, however, using a reverse-genetics system to express a rabies virus minigenome, Yang *et al.* (1999) have however shown that unphosphorylated N protein encapsidated more leader RNA than phosphorylated N protein. In addition, the lack of N protein phosphorylation resulted in a decrease of both viral transcription and replication.

The objectives of this part of the study were to investigate the influence of N protein phosphorylation on its ssRNA-binding activity and on its ability to interact with the P protein.

4.2 Materials and Methods

4.2.1 Expression of the rabies virus P protein in *E. coli* using the pPROEX™HT prokaryotic expression system

- *Cloning of the P gene into pPROEX™HT expression vector*

The rabies virus P gene was cloned into the prokaryotic pPROEX™HTc vector by making use of the procedures described in Section 2.2. The full-length P gene was recovered from the recombinant PpGEM vector by EcoRI excision and ligated into the EcoRI-linearised, dephosphorylated pPROEX™ vector. Following transformation of competent *E. coli* JM109 cells and plating of the cells onto LB agar plates supplemented with ampicillin (100 µg/ml), recombinant transformants were selected for further analysis. Plasmid DNA was isolated by the alkaline lysis method (Section 2.2.9) and analyzed by agarose gel electrophoresis and EcoRI restriction enzyme digestion. The translational orientation of the cloned inserts was determined by BamHI restriction mapping. A clone containing the full-length P gene, in the correct translational orientation, was selected and designated PpPROEX.

- *Expression of the P fusion protein*

Overnight *E. coli* JM109 cells transformed with the parental PpPROEX™ HTc and recombinant PpPROEX™ HTc vectors were diluted 1:10 in 10 ml of fresh medium and grown for 2 h at 37°C. Expression of the recombinant P-His fusion protein was induced by the addition of 1 mM IPTG after which incubation was continued for a further 5-6 h. The cells were harvested at 4 000 rpm for 5 min at 4°C in a Sorvall® RC-5B Refrigerated Superspeed centrifuge (Du Pont Instruments) using a SS-34 rotor. The cells were resuspended in 4 volumes of lysis buffer (50 mM Tris (pH 8.5 at 4°C), 5 mM 2-mercaptoethanol), sonicated and debris pelleted by centrifugation at 2 000 rpm for 1 min. The crude cell lysates were either stored at -20°C until needed or analyzed by SDS-PAGE and immunoblotting (Section 3.2.9).

4.2.2 Radiolabeling of the rabies virus fusion proteins synthesized in *E. coli*

Bacterial proteins were labeled *in vivo* with [^{35}S]methionine as described by Barik and Banerjee (1991). Briefly, single colonies of *E. coli* cells transformed with either PpGEX (Chapter 3), PpPROEX or parental plasmids were grown in 10 ml low-phosphate glucose-peptone medium (1.2% (w/v) Tris-base, 0.75% (w/v) peptone, 0.3% (w/v) NaCl, 0.01 mM MgSO$_4$.7H$_2$O, 0.001 mM CaCl$_2$, 0.3% (w/v) glucose). The pH of the medium was adjusted to 7.1 before addition of Ca^{2+} and Mg^{2+}. Simultaneous with the induction of the respective cultures with IPTG, 30 μCi of [^{35}S]methionine (1 mCi/ml, 1000 Ci/mmol; Amersham) was added per ml of culture. Following harvest, the cells were washed once with 1 × PBS and the proteins were analyzed by SDS-PAGE and autoradiography.

4.2.3 *In vitro* phosphorylation of the *E. coli*-expressed fusion proteins

A method for phosphorylating the bacterial-expressed proteins was developed based on the procedures described by Ogita *et al.* (1991), Gupta *et al.* (2000) and Kikkawa *et al.* (1982). The cells from IPTG-induced cultures of *E. coli* transformed with NpGEX, PpGEX, PpPROEX and parental plasmids were harvested by centrifugation and lysed as previously described (Section 4.2.1). The phosphorylation reaction mixtures (250 μl) contained 0.5 to 1 μg of the different fusion proteins, 5 μCi of [γ^{32}P]ATP in a kinase buffer (50 mM Tris (pH8), 1 mM DTT, 2 mM MgCl$_2$, 50 μM ATP) and 0.2 U of protein kinase C enzyme (PKC; Roche). The PKC activity was started with the addition of 200 μM CaCl$_2$, 8 μg 3-sn-Phosphatidyl-L-serine (Fluka) per ml, and 0.8 μg of 1,2-Dioleoyl-sn-glycerol (Fluka) per ml. The Phosphatidylserine and diacylglycerol, stored separately as chloroform solutions at -20°C, were mixed first in chloroform, dried, suspended in 20 mM Tris (pH 7.5) and sonicated. The reaction mixture was incubated at 30°C for 1 h and the reactions were terminated by the addition of ETDA (12 mM final concentration). The reaction mixtures or immunoprecipitates were analyzed by SDS-PAGE and autoradiography.

4.2.4 Immune-precipitation of GST fusion proteins

The bacterial-expressed rabies virus N- and P-GST fusion proteins were immunoprecipitated using a monoclonal anti-GST antibody (Santa Cruz Biotechnology,

Inc.). The samples were diluted in 500 µl in 1 × PBS. After addition of 1 µl of the anti-GST antibody, the reactions were incubated at room temperature for 90 min with gentle shaking. Then 40 µl of a 10% (w/v) suspension of insoluble Protein A (Sigma-Aldrich) was added, and incubation continued for 2 h. The immune-complexes were pelleted by centrifugation at 2 000 rpm for 1 min, and washed twice by resuspension in 40 µl 1 × PBS. The precipitated proteins were resuspended in 15 µl 1 × PBS and an equal volume of 2 × PSB was added. After boiling for 5 min, the samples were subjected to SDS-PAGE. For radiolabeled proteins, the gels were dried and autoradiographed. In the case of unlabeled proteins, the gels were further analyzed by Western blot using a rabies virus-specific antiserum as previously described (Section 3.2.9).

4.2.5   RNA-binding analysis

The ssRNA-binding ability of purified, *in vitro* phosphorylated and unphosphorylated N-GST and P-GST fusion proteins were investigated by means of a poly(U)- and poly(A)-Sepharose binding assay as previously described by Theron *et al.* (1994). Three hundred µl of the N-GST or 100 µl of the P-GST fusion proteins were each incubated with 10 mg poly(U)- or Poly(A)-Sepharose in 300 µl STE-TX buffer at NaCl concentrations between 0.01 M and 0.4 M NaCl, as indicated in Figs. 4.5 and 4.6. After allowing the proteins to bind to the respective Sepharose substrates for 30 min at room temperature with gentle agitation, the Sepharose was collected by centrifugation at 2 000 rpm for 1 min and washed twice with 300 µl of the corresponding binding buffer. The supernatant fractions were pooled, and the proteins precipitated with 4 volumes of acetone at − 70°C for 1 h and then pelleted by centrifugation at 18 000 rpm for 30 min at 4°C in a Beckmann SW50.1 rotor. The bound protein (Sepharose pellets) and unbound proteins (precipitated supernatants) were resuspended in 35 µl 0.01 M STE-TX and analyzed by SDS-PAGE. As a control, the affinity of the GST protein for poly(U)- and poly(A)-Sepharose was also assayed in 0.015 M STE-TX buffer. The RNA-binding was quantified by scanning the Coomassie-stained protein bands representing the bound and unbound fractions. The bound protein fraction was expressed as a fraction of the total protein recovered, bound and unbound.

4.2.6 Protein-protein binding analysis

[^{35}S]methionine-labeled cell lysates containing *in vitro* phosphorylated or unphosphorylated N-GST and P-His fusion proteins were used in a modified version of the GST pull-down assay described by Ansari *et al.* (1997). Crude N-GST lysates (400 µl) were incubated with 200 µl glutathione agarose for 30 min at room temperature with gentle agitation. Following brief centrifugation, the pellets were washed twice with MT-PBS to remove unbound N-GST fusion proteins. The pellet was gently resuspended in 300 µl of the binding buffer (50 mM Tris (pH7.4), 150 mM NaCl, 5mM EDTA) after which 300 µl of a P-His cell lysate was added. The mixture was incubated for 2 h at 4°C. Following incubation, the agarose-bound complexes were collected by low-speed centrifugation at 2 000 rpm for 1 min, washed twice with 300 µl binding buffer and then resuspended in 40 µl 1 × PSB. The recovered agarose-immobilized protein-complexes were subjected to SDS-PAGE, followed by autoradiography of the dried gel. As a control, the affinity of the P-His protein for glutathione agarose was also assayed. Quantification of protein complexes was performed by scanning of the radiolabeled bands representing the bound N and P proteins. The peaks were determined and expressed as the ratio of P bound to N.

4.3 Results

4.3.1 Cloning and expression of the rabies virus P protein in *E. coli* with the pPROEX™ HT prokaryotic expression system

In order to investigate the role of protein phosphorylation on rabies virus N and P protein interaction, the P protein was expressed as a His fusion protein in *E. coli* by making use of the pPROEX™ HT prokaryotic expression system. The use of a different affinity tag to GST was aimed at preventing the non-specific binding of the P-His fusion protein to glutathione agarose, thereby ensuring that the agarose-immobilized protein complex were the result of N and P protein interaction. The pPROEX™ HT vector contains a *trc* promoter and *lac*I^q gene that enables inducible expression of the cloned P gene with IPTG, thus resulting in a fusion protein with a histidine affinity tag at the amino terminus.

To enable expression of the rabies virus P protein in *E. coli*, the P gene was recovered from PpGEM by EcoRI excision and recloned into an EcoRI-digested and dephosphorylated PpPROEX™ HTc vector. Following plasmid extraction, the DNA was characterized by restriction enzyme digestion (Fig. 4.1). EcoRI restriction of selected plasmid DNAs resulted in the recovery of an *ca.* 900 bp DNA fragment, corresponding in size to the full-length P gene. Following restriction mapping with BamHI, a clone containing the P gene-specific insert in the correct translational orientation was selected for further use and designated as PpPROEX. For bacterial expression, overnight cultures of *E. coli* JM109 cells containing either parental or recombinant pPROEX vectors were diluted and grown in the presence of IPTG, as described under Materials and Methods (Section 4.2.1). Protein samples prepared from the cultures were subsequently analyzed by SDS-PAGE. The results indicated the presence of a unique protein band of 36 kDa in the extracts prepared from *E. coli* cells containing PpPROEX (Fig. 4.2A). No similar band was observed in the control cellular extracts. The size of the unique protein corresponds with the predicted size of the rabies virus P protein. To confirm the rabies virus origin of the induced protein, a Western blot was performed by making use of rabies virus serum as the primary antibody. The antibody reacted strongly with both the 36 kDa P-His fusion protein as well as the control 82 kDa N-GST fusion protein (Fig. 4.2B).

4.3.2 *In vitro* phosphorylation of the recombinant *E. coli*-expressed rabies virus N and P fusion proteins

In order to investigate the role of phosphorylation in the interaction of rabies virus N protein with ssRNA and the P protein, a source of phosphorylated bacterial-expressed rabies virus proteins was required. Since bacterial kinases are generally specific for their natural substrates (Barik and Banerjee, 1991), various exogenous kinase sources were investigated for their ability to phosphorylate the bacterial-expressed rabies virus fusion proteins. The use of uninfected *S. frugiperda* and BHK-21 cytoplasmic extracts for this purpose was investigated. However, incubation of the bacterial-expressed rabies virus fusion proteins with these extracts in the presence of [γ-^{32}P]ATP did not result in labeling of the respective fusion proteins (results not shown).
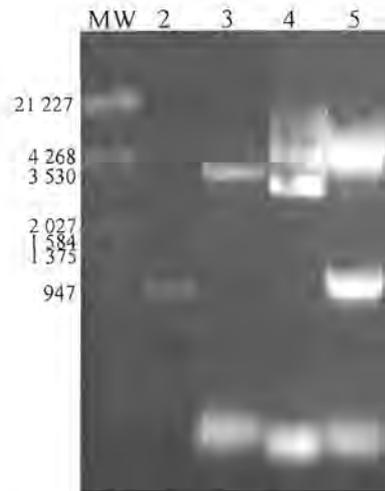
Fig. 4.1 Ethidium bromide stained 1% agarose gel, showing the construction of the recombinant PpPROEX Hc expression vector. Lane MW, Molecular weight marker; Lane 2, purifed P gene amplicon; Lane 3, EcoR1-linearized pPROEX Hc; Lane 4, uncut recombinant PpPROEX; Lane 5, EcoR1 digested PpPROEX. The sizes of molecular weight marker, phage Lambda DNA digested with EcoR1 and HindIII, are indicated to the left of the figure (in bp).
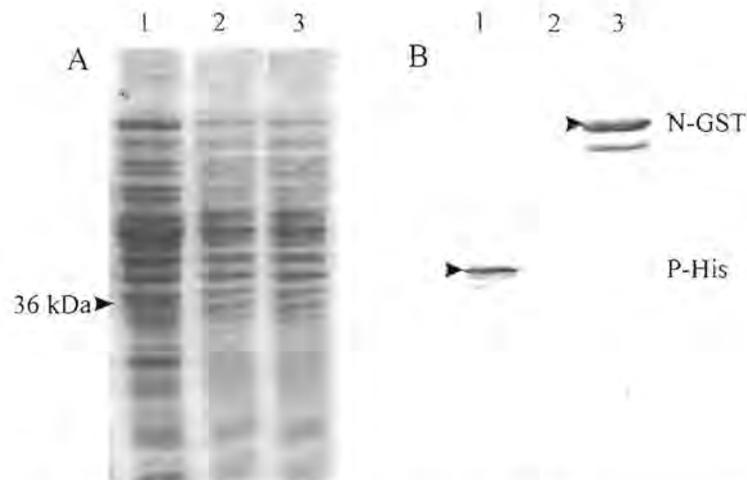
Fig. 4.2 Expression of the P protein in *E. coli* using the prokaryotic pPROEX expression system. (A) SDS-PAGE, and (B) Western blot with anti-rabies virus serum. (A) Lanes: 1, IPTG-induced cell cultures of *E. coli* transformed with PpPROEX Hc; 2, IPTG-induced *E. coli* transformed with pPROEX Hc; 3, unduced *E.coli*. (B) Lanes: 1, IPTG-induced *E. coli* cultures transformed with PpPROEX H; 2, IPTG-induced *E. coli* culture; 3, IPTG-induced cell-cultures of *E. coli* transformed with recombinant NpGEX. The arrows indicate the position of the 36 kDa P-His and the 80 kDa fusion protein.

Since both the rabies virus N and P proteins were found to contain the consensus sites for potential phosphorylation by protein kinase C (Chapter 2), an assay was developed whereby these proteins could be phosphorylated *in vitro* by making use of a commercially available phospholipid-dependent Protein Kinase C. Each phosphorylation reaction consisted of *E. coli* lysates containing either the N-GST, P-GST or P-His fusion proteins and [γ-^{32}P]ATP in the presence or absence of PKC enzyme together with its required activators. As controls, induced *E. coli* JM109 cells transformed with the parental pGEX-1 or pPROEX vector were included in the assay. Reaction mixtures containing N-GST, P-GST and the control GST protein were first immunoprecipitated with an anti-GST antibody prior to analysis by SDS-PAGE and autoradiography. Analysis of the autoradiographs indicated that neither the GST (Fig. 4.4A, lane 3) nor the histidine (Fig 4.4C, lane 3) affinity tags were phosphorylated when the respective extracts were incubated with [γ-^{32}P]ATP and the kinase enzyme. Furthermore, no incorporation of the radiolabel by the respective fusion proteins were observed in reaction mixtures from which the kinase enzyme was omitted (Fig. 4.4, lanes 2). However, the respective rabies virus fusion proteins were strongly labeled when incubated in the presence of both the PKC enzyme and [γ-^{32}P]ATP (Fig. 4.4, lanes 1). From these results it could be concluded that the rabies virus N and P proteins lacks autophosphorylation activity and, more importantly, that both the proteins could be *in vitro* phosphorylated by protein kinase C and that the kinase exhibits a specificity for both the N and P proteins.

4.3.3 Effect of phosphorylation on the ssRNA-binding activity of the rabies virus N protein

In the previous section it was shown that the bacterial-expressed rabies virus N and P proteins were unphosphorylated, but could be phosphorylated *in vitro* by protein kinase C. These fusion proteins could therefore be used to investigate the influence of phosphorylation on its ssRNA-binding ability. The ability of the unphosphorylated *E. coli*-expressed N-GST and P-GST fusion proteins to bind to ssRNA was therefore investigated before and after *in vitro* phosphorylation of the respective proteins by means

of poly(U)- and poly(A)-Sepharose affinity assays. In these assays, the purified fusion proteins (as described in Section 3.2.8) were allowed to interact with the respective Sepharose substrates in the presence of increasing concentrations of NaCl and the step-wise increase in salt concentration was used to measure changes in the degree of protein-RNA interaction. The purified GST protein served as a control in these assays. The control GST protein did not bind to either poly(U)- or poly(A)-Sepharose (results not shown). As expected, the P-GST fusion protein also did not bind to the respective substrates, irrespective of it being phosphorylated or not (results not shown). This corresponds with the lack of RNA-binding activity of the rabies virus P protein described in the literature (Yang *et al.*, 1999). The results obtained from the affinity binding assays of unphosphorylated and phosphorylated N-GST protein to poly(A)- and poly(U)-Sepharose are indicated in Fig. 4.5. The results indicated that *in vitro* phosphorylated and unphosphorylated N protein could bind to both poly(U)- and poly(A)-Sepharose. At low salt concentrations (0.025 to 0.1 M STE-TX) more of the phosphorylated N protein remained bound to poly(A)-Sepharose compared to poly(U)-Sepharose, but at higher salt concentrations (0.2 to 0.3 M STE-TX) more of the phosphorylated N protein remained bound to the poly(U)-Sepharose. Similarly, at salt concentrations ranging from 0.025 M to 0.2 M STE-TX, more of the unphosphorylated N protein remained bound to poly(A)-Sepharose compared to poly(U)-Sepharose. At salt concentrations between 0.25 and 0.4 M STE-TX, no binding of the unphosphorylated N protein to poly(A)-Sepharose was observed, although the unphosphorylated N protein still bound to the poly(U)-Sepharose.

In the case of the poly(U)-Sepharose affinity assays (Fig 4.6A), slightly less of the unphosphorylated N bound to the poly(U)-Sepharose at low salt concentrations (0.01 to 0.1 M) compared to the phosphorylated N protein. The average difference in protein bound was less than 1.2 %. At higher salt concentrations of 0.15 M to 0.3 M STE-TX, there was a clear difference in the capacity of phosphorylated and unphosphorylated N protein to bind to the poly(U)-Sepharose. At these salt concentrations, more of the phosphorylated N protein remained bound to the poly(U)-Sepharose compared to the unphosphorylated N protein. The most drastic difference was observed at 0.15 M STE-TX with almost 6-fold more of the phosphorylated N protein being bound compared to

the unphosphorylated N protein. Similarly, throughout a wide range of salt concentrations, more of the phosphorylated N protein remained bound to poly(A)-Sepharose compared to the unphosphorylated N protein (Fig. 4.6B). At a salt concentration of 0.15 M STE-TX, very little difference in the capacity of phosphorylated and unphosphorylated N proteins to bind to poly(A)-Sepharose was observed. However, at salt concentrations of 0.25 M STE-TX and higher, none of the unphosphorylated N protein remained bound. At the same salt concentrations, the phosphorylated N protein remained bound to the poly(A)-Sepharose, albeit at reduced capacity, and no binding was observed at the highest salt concentration investigated (0.4 M STE-TX).

4.3.4   Protein-protein binding assay

The ability of phosphorylated and unphosphorylated rabies virus N and P proteins to associate with each other were also investigated. The bacterial-expressed N-GST and P-His fusion proteins were metabolically labeled with [^{35}S]methionine during induction of proteins synthesis (Fig. 4.3). The [^{35}S]methionine-labeled GST or N-GST fusion proteins were bound to glutathione agarose, as described in Section 3.2.8, and then incubated with similarly labeled P-His bacterial lysates. After incubation, the unbound material was removed by several washing steps. Finally, the material bound to the glutathione agarose was analyzed by SDS-PAGE and autoradiography. The interaction between the N-GST and P-His proteins were determined by measuring the amount of N-GST and P-His in each interaction and expressing the results as the ratio of P bound to N (Fig. 4.7). The ratio of unphosphorylated P bound to unphosphorylated N is 2.07, while the ratio of phosphorylated P bound to unphosphorylated N, phosphorylated N bound to phosphorylated P is 0.48 and 0.45 respectively. In the case of phosphorylated N bound to unphosphorylated P the ratio is 1. The results from the control GST assay indicated that the P-His protein did not bind to the GST protein (results not shown).
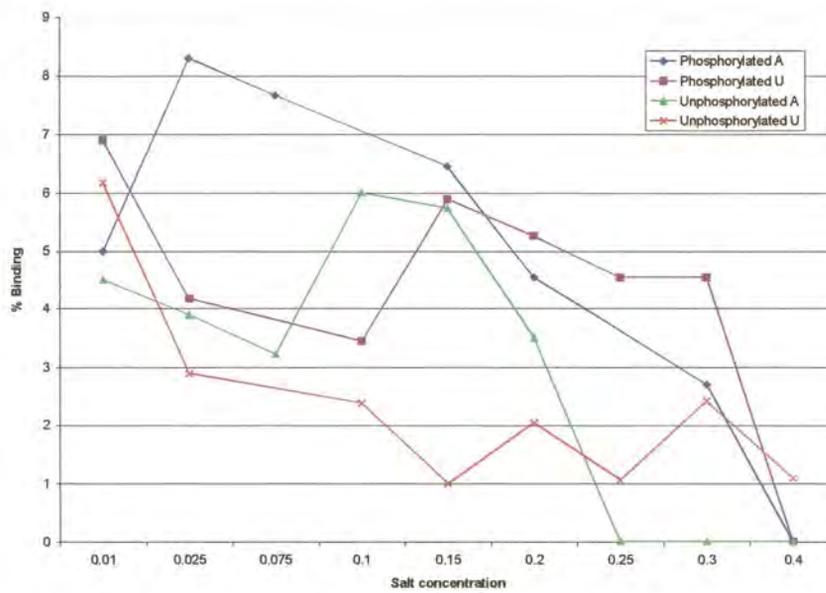
Fig. 4.5 Affinity of unphosphorylated and phosphorylated N-GST for poly(A) and poly(U) ssRNA. The N-GST was expressed in *E. coli*, purified, with or without phosphorylation and assayed for poly(A)- and poly(U)-Sepharose-binding ability over a range of different ionic strengths, as described in Materials and Methods. Phosphorylated N-GST binding to poly(A)-Sepharose is indicated in blue, phosphorylated N-GST binding to poly(U)-Sepharose is indicated in pink, unphosphorylated N-GST binding to poly(A)-Sepharose is indicated in green and unphosphorylated N-GST binding to poly(U)-Sepharose is indicated in red.
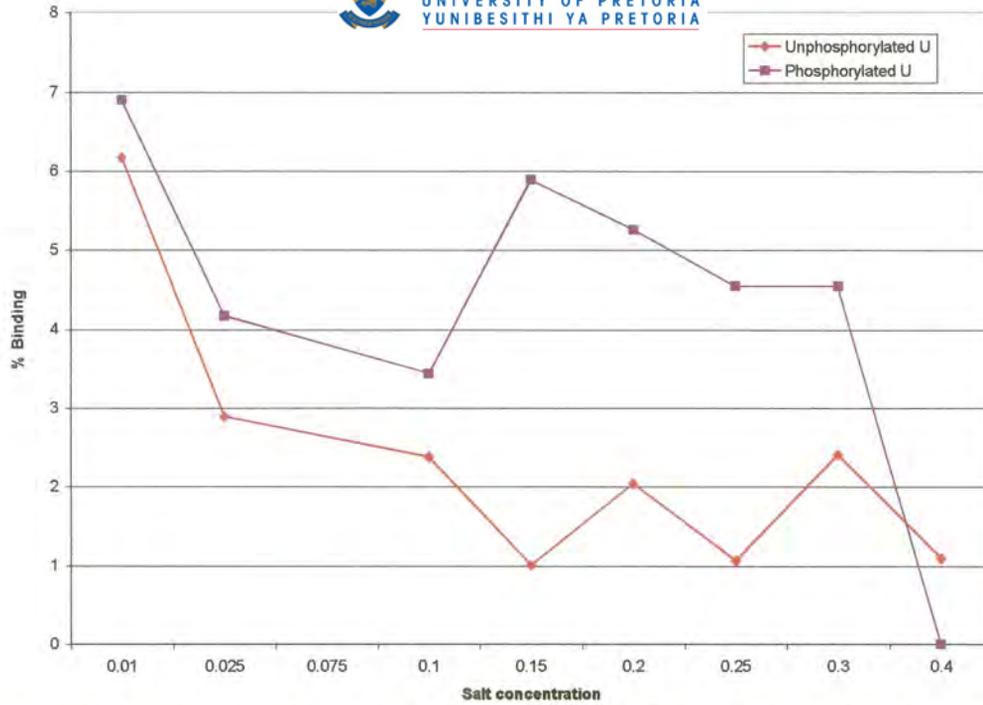
Fig. 4.6(A) Affinity of phosphorylated and unphosphorylated N-GST for poly(U) ssRNA. The N-GST was expressed in *E. coli*, purified, phosphorylated or not and assayed for poly(U)-Sepharose-binding ability over a range of different ionic strengths, as described in Materials and Methods. Phosphorylated N-GST is indicated in red and unphosphorylated N-GST is indicated in pink.
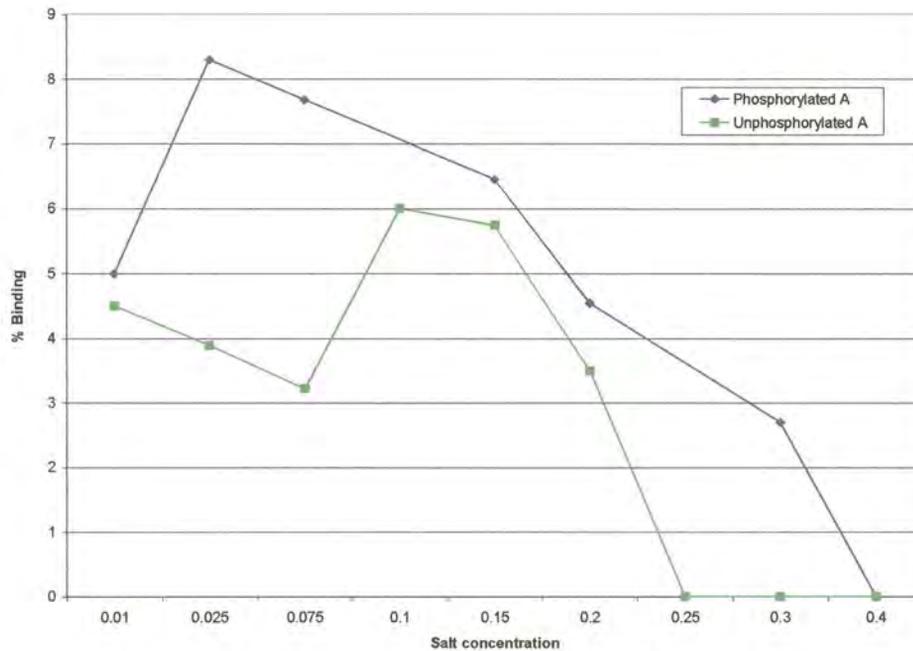
Fig. 4.6(B) Affinity of phosphorylated and unphosphorylated N-GST for poly(A) ssRNA. The N-GST was expressed in *E. coli*, purified, phosphorylated or not and assayed for poly(A)-Sepharose-binding ability over a range of different ionic strengths, as described in Materials and Methods. Phosphorylated N-GST is indicated in blue and unphosphorylated N-GST is indicated in green.
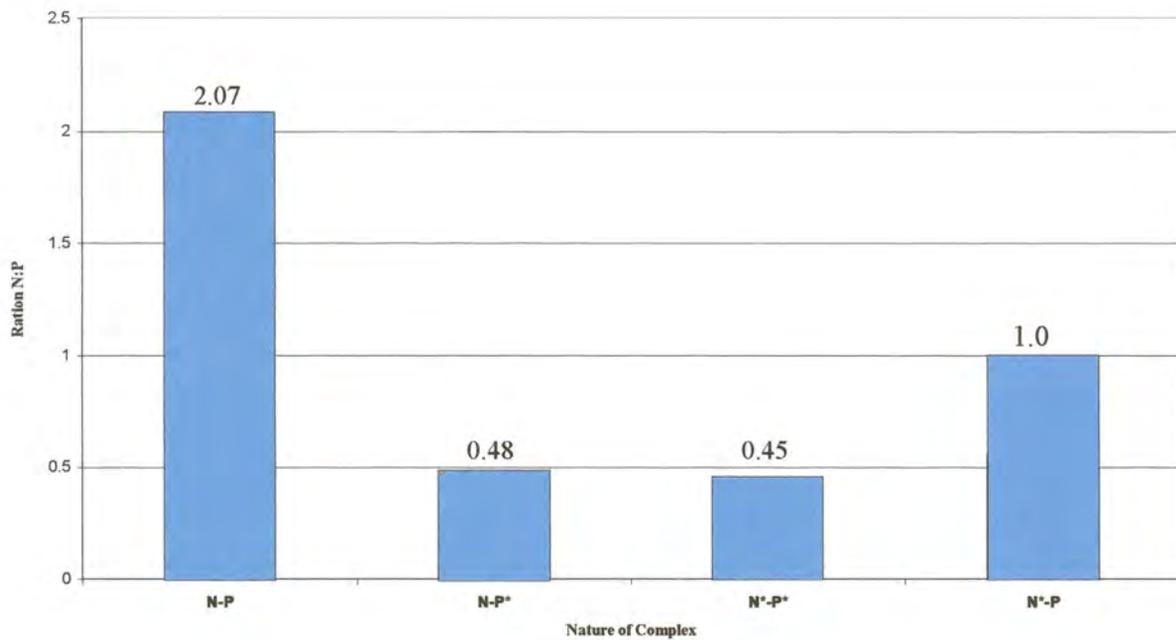
Fig. 4.7 Graphical presentation of the complexes formed between rabies virus N and P protein, with various combinations of phosphorylated and unphosphorylated N and P, respectively. The N:P ratio is given in the bar chart, with actual value given on top of the bar. The * indicate the protein that is phosphorylated.

4.4 Discussion

The importance of phosphorylation of the rabies virus N protein for its interaction with ssRNA and with the rabies virus P protein was studied *in vitro* using bacterial-expressed fusion proteins. For this purpose phosphorylated versions of the unphosphorylated bacterial-expressed rabies virus N and P proteins were required. Initially, the use of eukaryotic cell extracts as sources of kinases for phosphorylation of the N and P fusion proteins were investigated. These attempts at phosphorylating the bacterial-expressed proteins were, however, unsuccessful. Similarly, Gupta *et al.* (1995) also could not phosphorylate *E. coli*-expressed P protein *in vitro* by using a BHK cell extract and Prehaud *et al.* (1992) found that expression of the P protein in insect cells did not result in phosphorylation of the protein. It would thus appear that phosphorylation of the rabies virus proteins requires a specific protein kinase, which may be absent from the cell extracts used above.

Analysis of the rabies virus N and P amino acid sequences indicated the presence of consensus recognition sites for protein kinase C (PKC) and casein kinase II (CKII) (Chapter 2). Recently, Gupta *et al.* (2000) demonstrated that the rabies virus P protein could be phosphorylated by PKC and a second unique kinase, rabies virus protein kinase (RVPK), which is selectively packaged in the mature virion. The respective kinases were shown to phosphorylate the P protein at the C- and N terminal, respectively. In the case of P protein phosphorylation by PKC, four isomers (α, β, γ, ζ) of the enzyme were capable of phosphorylating the protein, but the γ isomer was found to be preferred to the other isomers (Gupta *et al.*, 2000). In contrast to the rabies virus P protein, the kinase(s) responsible for N protein phosphorylation has not been characterized. By making use of a commercially available preparation of PKC, bacterial-expressed rabies virus N and P fusion proteins could both be specifically phosphorylated. This also constitutes the first evidence for phosphorylation of the rabies virus N protein by PKC.

The N protein of rabies virus is a RNA-binding protein that encapsidates both positive and negative sense RNA, thereby protecting it from nuclease digestion and keeping it available for transcription and replication (Emerson, 1987; Iseni *et al.*, 2000; Keene *et al.*,

1981; Klumpp *et al.*, 1997). The rabies virus N protein binds to an A-rich encapsidation site in the leader RNA with the following sequence: 5'-AAGAAAAAACA-3' (Yang *et al.*, 1998). To investigate whether the N protein-RNA interaction may be the result of the A-rich nature of the RNA substrate, the bacterial-expressed *in vitro* phosphorylated and unphosphorylated N-GST fusion proteins were subjected to RNA-binding affinity assays using two different Sepharose substrates, poly(U)- and poly(A)-Sepharose. The percentages ssRNA-binding obtained in this study are very low. The binding experiments done by other groups (Yang *et al.*, 1998, 1999; Kawai *et al.*, 1999) did not quantify the binding results and only concluded whether binding occurred or not. The conditions under which the binding experiments during this study was done, is not optimal, but the fact that P-GST do not bind to the ssRNA template, while binding do occur with the N-GST indicate that the results reflect the binding characteristics of the rabies virus N protein. The results obtained indicated that the unphosphorylated N protein had an affinity for both poly(U)- and poly(A)-Sepharose. In the case of *in vitro* phosphorylated N-GST protein, the protein bound preferentially to poly(A)-Sepharose at salt concentrations lower than 0.15 M NaCl, and to poly(U)-Sepharose at salt concentrations higher than 0.15 M NaCl. At physiological salt concentrations (0.15 M), the phosphorylated N-GST protein bound equally well to both the ssRNA substrates. From these results it can be concluded that the N protein binds nonspecifically to ssRNA. These results are in agreement with those reported by Yang *et al.* (1999). The rabies virus N protein was reported to bind non-specifically to viral and non-viral cellular RNA in the absence of P protein. Simultaneous addition of the N and P proteins to leader RNA resulted in the specific binding of viral RNA. It would thus appear that the specificity in N protein-virus RNA binding is conferred by the P protein.

It has previously been reported that unphosphorylated N protein encapsidates more leader RNA than the phosphorylated N protein, but that N protein phosphorylation increases viral transcription and replication (Yang *et al.*, 1999). Protein phosphorylation was prevented by mutating Ser-389 to an alanine amino acid, this amino acid forms part of a Casein Kinase II recognition site (Fig. 2.4A). Newly synthesized N proteins became associated with the P protein prior to it being phosphorylated and is only then able to

encapsidate viral RNA (Kawai *et al.*, 1999). The encapsidation-associated conformational change is a prerequisite step for N protein to be phosphorylated at Ser-389 (Kawai *et al.*, 1999). To investigate the role of Protein Kinase C specific phosphorylation on the ssRNA-binding activity of the rabies virus N protein, *in vitro* phosphorylated and unphosphorylated bacterial-expressed N-GST were used in similar RNA-binding assays as above. It was found that throughout a wide range of different ionic strengths the unphosphorylated N-GST protein preparations were less effective in binding ssRNA, compared to the phosphorylated N-GST preparations. The control GST and P-GST protein preparations did not display an affinity for the ssRNA, but it can be that the 27.5 kDa fusion part of the N-GST protein changes the conformation to such an extent that the unphosphorylated protein is less effective during its binding to ssRNA. From these results it could be concluded that specific phosphorylation by Protein Kinase C changes the conformation of the rabies virus N protein to such a degree that it enhances its ability to bind to ssRNA.

The role of phosphorylation on the ability of rabies virus N and P proteins to associate with each other during complex formation was also investigated. *In vivo* [^{35}S]methionine-labeled N-GST and P-His fusion proteins were phosphorylated *in vitro* or left unphosphorylated and then mixed in different combinations to allow complexes to form. In the case of both the N and P proteins being unphosphorylated, N and P proteins interact to form complexes with a 2:1 stoichiometry. This is similar to the N:P ratios found in rabies virus virions (Fu *et al.*, 1994; Wunner, 1991) and also in VSV infected cells (Davis *et al.*, 1986; Howard *et al.*, 1987). Phosphorylation of the P protein complexed with either phosphorylated or unphosphorylated N protein resulted in a stoichiometry of 1:2 of N to P protein. In the event of unphosphorylated P protein complexed to the phosphorylated N protein, a N:P ratio of 1:1 was found. The association of newly synthesized unphosphorylated N protein with the P protein is required for phosphorylation of the N protein at Ser-389 (located within a CKII recognition site) where after encapsidation of viral RNA occurs (Kawai *et al.*, 1999). It would thus seem that both the N and P proteins must be unphosphorylated in order to correctly interact with each other. Phosphorylation of either N or P protein may alter the

protein conformation due to a charge difference thereby preventing the proteins to form complexes with a 2:1 stoichiometry.

CHAPTER 5

Concluding Remarks

The nucleoprotein of rabies virus, the prototype virus of the *Lyssavirus* genus, plays an important role during virus transcription and replication. It is capable of binding to the viral RNA genome *via* the phosphate-sugar backbone thereby exposing the nucleotide bases to the RNA polymerase and simultaneously protecting it from nuclease digestion (Emerson, 1987; Iseni *et al.*, 2000; Klumpp *et al.*, 1997). It also forms complexes with the phosphoprotein which maintains the N protein in a soluble form. In contrast to vesicular stomatitis virus (VSV), the prototype model for the *Rhabdoviridae* virus family, the rabies virus N protein is phosphorylated. Although much research regarding the kinases responsible for P protein phosphorylation as well as the identification and localization of the phosphorylated amino acid residues have been performed (Gupta *et al.*, 2000), no information regarding the kinase(s) responsible for phosphorylation of the N protein is available. Neither has the functional role that phosphorylation may play in N protein function been extensively investigated.

Towards investigating the functional significance of N protein phosphorylation, cDNA copies of the N and P genes, respectively, of a South African viverrid rabies virus isolate was cloned and sequenced. The nucleotide sequences displayed high degrees of homology when compared to the cognate sequences of other rabies virus strains. Important functional domains such as the ssRNA-binding domain and P protein-binding site on the N protein as well as the domain responsible for P protein binding to the L protein was found to be highly conserved amongst all the proteins included in the analysis. Analysis of the N and P amino acid sequences furthermore indicated the presence of several potential kinase recognition sites for protein kinase C and casein kinase II. Previously identified phosphorylated amino acid residues in the rabies virus P protein (Ser-162, Ser-210 and Ser-271; Gupta *et al.*, 2000), were found to be conserved in the P protein of the South African rabies virus strain and was located within protein kinase C recognition sites. The rabies virus N protein has been reported to be phosphorylated at Ser-389 and Thr-375 (Yang *et al.*, 1998; Kawai *et al.*, 1999) and these sites were also conserved in the South African rabies virus strain.

The rabies virus N and P proteins were subsequently expressed in a eukaryotic expression system. Expression of the rabies virus N and P proteins in the eukaryotic BAC-to-BAC™ baculovirus expression system resulted in very low levels of expression and the expressed proteins could only be detected following Western blot analysis using a rabies virus-specific serum. Whereas the recombinant baculovirus-expressed P protein was found to be soluble, the recombinant baculovirus-expressed N protein was insoluble and could only be detected in the particulate fraction prepared from the infected cells. Previous reports have indicated that co-expression of the rabies virus N and P proteins maintain the N protein in a soluble form (Yang *et al.*, 1999). A baculovirus capable of co-expressing the rabies virus N and P proteins was subsequently constructed. Although both proteins were synthesized in the recombinant baculovirus-infected cells, soluble N-P protein complexes could not be observed. This may have been due to the low expression levels obtained.

To overcome the low levels of protein synthesis obtained in the baculovirus expression system, the rabies virus N and P proteins were expressed in a prokaryotic pGEX expression system as GST fusion proteins. The respective proteins were successfully expressed as soluble fusion proteins that could be purified from bacterial lysates by glutathione affinity chromatography. The availability of purified unphosphorylated soluble N-GST allowed for further investigations regarding the functional significance of N protein phosphorylation. To obtain phosphorylated versions of the rabies virus N and P proteins, an *in vitro* phosphorylation assay was developed by making use of a commercially available phospholipid-dependent protein kinase C. It was found that protein kinase C phosphorylated both the rabies virus N and P proteins specifically. The *in vitro* phosphorylated proteins were subsequently used to investigate the influence of phosphorylation on the ssRNA-binding ability of the N protein. Poly(U)- and poly(A)-Sepharose binding assays, performed in the presence of increasing salt concentrations, indicated that throughout a wide range of salt concentrations, phosphorylated N-GST protein displayed a higher affinity for ssRNA than unphosphorylated N-GST. As expected, the P protein did not bind to the ssRNA. Yang *et al.* (1999) reported that phosphorylation of the N protein resulted in encapsidation of less RNA. In their

investigtion, the unphosphorylated N protein was obtained following mutation of Ser-389 to alanine, thereby preventing phosphorylation of the protein by casein kinase II. In this investigation, a commercially available protein kinase C enzyme was used to phosphorylate the N protein *in vitro*. It is known what the effect of phosphorylation by casein kinase II, if any, on the ssRNA-binding ability of the N protein is. This aspect of N protein phosphorylation should be further investigated.

Since the encapsidation initiation site on the virus genome has been mapped to an A-rich region in the leader RNA, it was investigated whether specificity in N protein-RNA interaction may be the result of the A-rich nature of the RNA substrate. The bacterial-expressed *in vitro* phosphorylated N-GST fusion protein, however, did not display a higher affinity for the A-rich ssRNA substrate as both poly(U)- and poly(A)-Sepharose was bound equally well at physiological salt concentration of 0.15 M. These findings are in agreement with those of Iseni *et al.* (1998) and Yang *et al.* (1999) who reported that the rabies virus N protein binds non-specifically to RNA in the absence of P protein. Previous reports have indicated that when preformed N-P protein complexes were used in the binding studies, the ability of the N protein to bind to the viral RNA was reduced. However, when the N and P proteins were simultaneously added to leader RNA, the N protein specifically bound to the viral RNA. It would thus appear that the specificity in N protein-virus RNA binding is conferred by binding of the P protein to the N protein.

The influence of rabies virus N and P protein phosphorylation on complex formation between these two proteins was furthermore investigated. Different combinations of *in vitro* phosphorylated N-GST and P-His fusion proteins were allowed to interact and the ratio of N:P was determined. A 2:1 stoichiometry of N to P in *in vitro* protein complexes could be obtained when both the N and P fusion proteins were unphosphorylated. This corresponds to the 2:1 ratio of N to P observed in purified rabies virions (Fu *et al.*, 1994; Wunner, 1991). However, phosphorylation of the P protein resulted in *in vitro* protein complexes having a N to P ratio of 1:2. However, phosphorylated N protein bound to unphosphorylated P protein to form complexes with a ratio of 1:1. These results suggest that the N and P proteins interact with each other inside the virus before phosphorylation

to form N:P complexes of the correct stoichiometry. Phosphorylation of either the N and/or the P protein resulted in *in vitro* complexes with different stoichiometries compared to that found in the rabies virions. This is consistent with the model proposed by Kawai *et al.* (1999) in which newly synthesized unphosphorylated N protein associates with the P protein before being phosphorylated at Ser-389 (located within a CKII recognition site).

Historically rabies virus research has focused on the epidemiology of the virus and characterization of the different proteins to determine their role in the virulence and antigenicity of the virus. In the last decade the focus has shifted towards understanding the role of the viral proteins during the process of transcription and replication. The identification of the kinases involved in phosphorylation of the phosphoprotein (Gupta *et al.*, 2000) and reverse genetics approaches (Yang *et al.*, 1999), will enable researchers to elucidate the mechanisms involved in controlling the replication and transcription processes. Future research should center on the role of the P protein in conferring specificity to the encapsidation of viral RNA, a similar approach to the one used for investigating the protein-protein interactions could be adopted, but by allowing the different complexes to bind to different virus and non-virus RNA. The role of N protein phosphorylation by both casein kinase II and protein kinase C can be investigated by making cocktails of these two kinases. This will indicate the effect of differentially phosphorylated rabies virus N protein and N-P complexes on ssRNA binding. By investigating the functional significance of phosphorylation of the N and P proteins during RNA-binding and protein complex formation, new aspects of the control process may be revealed.

References

Abraham, G. and Banerjee, A.K. 1976. Sequential transcription of the genes of vesicular stomatitis virus. *Proc. Natl. Acad. Sci. USA.* **73**, 1504-1508.

Anzai, J., Takamatsu, F., Takeuchi, K., Kohno, T., Morimoto, K., Goto, H., Minamoto, N. and Kawai, A. 1997. Identification of a phosphatase-sensitive epitope of rabies virus nucleoprotein which is recognized by a monoclonal antibody 5-2-26. *Microbiol.Immunol.* **41**, 229-240.

Ball, L.A., White, C.N. 1976. Order of transcription of genes of vesicular stomatitis virus. *Proc. Natl. Acad. Sci. U S A.* **73**, 442-6.

Baltimore, D., Huang, A.S. and Stampfer, M. 1970. Ribonucleic acid synthesis of vesicular stomatitis virus. II An RNA polymerase in the virion. *Proc. Natl. Acac. Sci. USA.* **66**, 572-576.

Banerjee, A.K. 1987. Transcription and replication of Rhabdoviruses. *Microbiol.Rev.* **51**, 66-87.

Banerjee, A.K., Abraham, G. and Colonno, R.J. 1977. Vesicular stomatitis virus: mode of transcription. *J. Gen. Virol.* **34**, 1-8.

Banerjee, A.K. and Barik, S. 1992. Gene expression of vesicular stomatitis virus genome RNA. *Virology*, **188**: 417-428.

Barge, A., Gagnon, J., Chaffotte, A., Timmins, P., Langowski, J., Ruigrok, R.W.H. and Gaudin, Y. 1996 Rod-like shape of vesicular stomatitis virus matrix protein. *Virology*, **219**, 465-470.

Barik, S. and Banerjee, A.K. 1991. Cloning and expression of the vesicular stomatitis virus phosphoprotein gene in *Escherichia coli*: analysis of phosphorylation status versus transcriptional activity. *J. Virol.* **65**, 1719-1726.

Barik, S. and Banerjee, A.K. 1992a. Phosphorylation by cellular casein kinase II is essential for transcriptional activity of vesicular stomatitis virus phosphoprotein P. *Proc. Natl. Acad. Sci. USA.* **89**, 6570-6574.

Barik, S. and Banerjee, A.K. 1992b. Sequential phosphorylation of the phosphoprotein of vesicular stomatitis virus by cellular and viral protein kinases is essential for transcription activation. *J. Virol.* **66**, 1109-1118.

Barr, J.N., Whelan, S.P.J. and Wertz, G.W. 1997a. *cis*-Acting signals involved in termination of vesicular stomatitis virus mRNA synthesis include the conserved AUAC and the U7 signal for polyadenylation. *J. Virol.* **71**, 8718-8725.

Barr, J.N., Whelan, S.P.J. and Wertz, G.W. 1997b. Role of the intergenic dinucleotide in vesicular stomatitis virus RNA transcription. *J.Virol.* **71**, 1794-1801.

Bell, J.C., Brown, E.J., Takayesu, D. and Prevec, L. 1984. Protein kinase activity associated with immunoprecipitate of the vesicular stomatitis virus phosphoprotein NS. *Virology.* **132**, 220-238.

Berg, M., Ehrenborg, C., Blomberg, J., Pipkorn, R. and Berg, A. 1998. Two domains of the borna disease virus p40 protein are required for interation with the p23 protein. *J.Gen.Virol.* 79, 2957-2963.

Birnboim, H.C. and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*, 7, 1513-1523.

Blumberg, B.M., Giorgi, C. and Kolakofsky, D. 1983. N protein of vesicular stomatitis virus selectively encapsidates leader RNA in vitro. *Cell.* 32, 559-567.

Blumberg, B.M., Giorgi,C., Rose, K. and Kolakofsky, D. 1984. Preparation and analysis of the nucleocapsid proteins of vesicular stomatitis virus and Sendai virus, and analysis of the sendai virus leader-NP gene region. *J. Gen. Virol.* 65, 769-779.

Blumberg, B.M. and Leppert, M. and Kolakofsky, D. 1981. Interaction of VSV leader RNA and Nucleocapsid protein may control VSV genome replication. *Cell.* 23, 837-845.

Bourhy, H, Kissi, B. and Tordo, N. 1993. Molecular diversity of the *Lyssavirus* genus. *Virology*, 194, 70-81.

Bourhy, H., Tordo, N., Lafon, M. and Sureau, P. 1989. Complete cloning and molecular organization of a rabies-related virus, mokola virus. *J.Gen.Virol.*, 70, 2063-2074.
Brown, N. and Faulkner, P. 1977. A plaque assay for nuclear polyhedrosis virus using a solid overlay. *J. Gen. Virol.*, 36, 361-364.

Chattopadyay, D. and Banerjee, A.K. 1987b. Phosphorylation within a specific domain of the phosphoprotein of vesicular stomatitis virus regulates transcription *in vitro*. *Cell*, 49, 407-414.

Chattopadhyay, D. and Banerjee, A.K. 1988. NH_2-terminal acidic region of the phosphoprotein of vesicular stomatitis virus can be functionally replaced by tubulin. *Proc. Natl. Acad. Sci. USA*. 85, 7977-7981.

Chen, J., Tapas, D. and Banerjee, A.K. 1997. Phosphorylated states of vesicular stomatitis virus P protein *in vitro* and *in vivo*. *Virology.* 228, 200-212.

Chenik, M., Chebli, K. and Blondel, D. 1995. Translation initiation at alternate in-frame AUG codons in the rabies virus phosphoprotein mRNA is mediated by ribosomal leaky scanning mechanism. *J.Virol.* 69, 707-712.

Chenik, M., Chebli, K., Gaudin, Y. and Blondel, D. 1994. *In vivo* interaction of rabies virus phosphoprotein (P) and nucleoprotein (N): existence of two N-binding sites on P protein. *J. Gen. Virol.* 75, 2889-2896.

Chenik, M., Schnell, M., Conzelmann, K.K. and Blondel, D. 1998. Mapping the interacting

domains between the rabies virus polymerase and phosphoprotein. *J. Virol.* **72**, 1925-1930.

Conzelmann, K., Cox, J.H., Schneider, L.G. and Thiel, H. 1990. Molecular cloning and complete nucleotide sequence of the attenuated rabies virus SAD B19. *Virology.* **175**, 485-499.

Coslett, D.G., Holloway, B.P. and Obijeski, J.F. 1980. The structural proteins of rabies virus and evidence for their synthesis form separate monocistronic RNA species. *J. Gen. Virol.* **49**, 161-180.

Coulon, P., Lafay, F. and Flamand, A. 1993. Rabies virus antigenicity: an overview. *Onderstepoort J. Vet. Res.* **60**, 271-275.

Das, T. and Banerjee, A.K. 1993. Expression of the vesicular stomatitis virus nucleocapsid protein gene in *Escherichia coli*: analysis of its biological activity *in vitro*. *Virology.* **193**, 340-347.

Davis, N.L., Arnheiter, H. and Wertz, G.W. 1986. Vesicular stomatitis virus N and NS proteins form multiple complexes. *J. Virol.* **59**, 751-754.

De, B.P. and Banerjee, A.K. 1985. Requirements and functions of vesicular stomatitis virus L and NS proteins in the transcription process *in vitro*. *Biochem. Biophys, Res. Commun.* **126**, 40-49.

De, B.P., Das, T. and Banerjee, A.K. 1997. Role of cellular kinases in the gene expression of nonsegmented negative strand RNA viruses. *Biol.Chem.* **378**, 489-493.

Emerson, S.U. 1982. Reconstitution studies detect a single polymerase entry site on the vesicular stomatitis virus genome. *Cell,* **31**, 635-642.

Emerson, S.U. 1987. *The rhabdoviruses.* New York: Plenum.

Emerson, S.U. and Schubert, M. 1987. Location of the binding domains for the RNA polymerase L and the ribonucleocapsid template within different halves of the NS phosphoprotein of vesicular stomatitis virus. *Proc. Natl. Acad. Sci. USA.* **84**, 5655-5659.

Emerson, S.U. and Yu, Y.H. 1975. Both NS and L proteins are required for in vitro RNA synthesis by vesicular stomatitis virus. *J. Virol.* **15**, 1348-1356.

Erlich, H.A., Gelfand, D. and Sninsky, J.J. 1991. Recent advances in the polymerase chain reaction. *Science,* **252**, 1643-1651.

Finke, S. Cox, J.H. and Conselmann, K. 2000. Differential transcription attenuation of rabies virus genes by intergenic regions: Generation of recombinant viruses overexpressing the polymerase gene. *J. Virol.* **74**, 7261-7269.

Flamand, A., and Delagneau, J.F. 1978. Transcriptional mapping of rabies virus *in vivo.* *J. Virol.* **28**, 518-523.

Frangioni, J.V. and Neel, B.G. 1993. Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal. Biochem.* **210**, 179-187.

Fraenkel-Conrat, H., Kimball, P.C., Levy, J.A. 1988. *Virology*, 2nd Ed., New Jersey, Prentice Hall, Inc.

Follett, E.A.C., Pringle, C.R., Wunner, W.H. and Skehel, J.J. 1974. Virus replication in enucleated cells: vesicular stomatitis virus and influenza virus. *J. Virol.* **13**, 394-399.

Fu, Z.F., Dietzschold, B., Schumacher, C.L., Wunner, W.H., Ertl, H.C.J. and Koprowski, H. 1991. Rabies virus nucleoprotein expressed in and purified from insect cells is efficacious as a vaccine. *Proc. Natl. Acad. Sci. USA.* **88**, 2001-2005.

Fu, Z.F., Zheng, Y., Wunner, W.H., Koprowski, H. and Dietzschold, B. 1994. Both the N- and the C-terminal domains of the nominal phosphoprotein of rabies virus are involved in binding to the nucleoprotein. *Virology.* **200**, 590-597.

Galarza, J.M., Sowa, A., Hill, V.M., Skorko, R. and Summers D.F. 1992. Influenza A virus NP protein expressed in insect cells by a recombinant baculovirus is associated with a protein kinase activity and possesses single-stranded RNA binding activity. *Virus Res.* **24**, 91-106.

Gao, Y. and Lenard, J. 1995. Multimerization and transcriptional activation of the phosphoprotein (P) of vesicular stomatitis virus by casein kinase-II. *EMBO J.*, **14**, 1240-1247.

Gigant, B., Iseni, F., Gaudin, Y., Knossow, M. and Blondel, D. 2000. Neither phosphorylation nor the amino-terminal part of rabies virus phosphorylation is required for its oligomerization. *J.Gen.Virol.* **81**, 1757-1761.

Gill, D.S., Chattopadhyay, D., and Banerjee, A.K. 1986. Identification of a domain within the phosphoprotein of vesicular stomatitis virus that is essential for transcription *in vito*. *Proc. Natl. Acad. Sci. USA.* **83**, 8873-8877.

Gill, D.S. and Banerjee, A.K. 1985. Vesicular stomatitis virus NS proteins: structural similarity without extensive sequence homology. *J. Virol.* **55**, 60-66.

Giorgi, C., Blumberg, B. and Kolakofsky, D. 1983. Sequence determination of the (+) leader RNA regions of the vesicular stomatitis virus Chandipura, Cocal, and Piry serotype genomes. *J. Virol.* **46**, 125-130.

Gupta, A.K. and Banerjee, A.K. 1997. Expression and purification of vesicular stomatitis virus N-P complex from *Escherichia coli*: role in genome RNA transcription and replication *in vitro*. *J. Virol.* **71**, 4264-4271.

Gupta, A.K., Blondel, D., Choudhary, S. and Banerjee, A.K. 2000. The phosphoprotein of rabies virus is phosphorylated by a unique cellular protein kinase and specific isomers of Protein Kinase C.

J. Virol. **74**, 91-98.

Gupta, A.K., Das, T. and Banerjee, A.K. 1995. Casein kinase II is the P protein phosphorylating cellular kinase associated with the ribonucleoprotein complex of purified vesicular stomatitis virus. *J. Gen. Virol.* **76**, 365-372.

Gupta, A.K., Drazba, J.A., Banerjee, A.K. 1998. Specific interaction of heterogeneous nuclear ribonucleoprotein particle U with the leader RNA sequence of vesicular stomatitis virus. *J. Virol.* **72**, 8532-40

Goto, H., Minamoto, N., Ito, H., Luo, T.R., Sugiyama, M., Kinjo, T. and Kawai, A. 1995. Expression of the nucleoprotein of rabies virus in *Escherichia coli* and mapping of antigenic sites. *Arch.Virol.* **140**, 1061-1074.

Hercyk, N., Horikami, S.M. and Moyer, S.A. 1988. The vesicular stomatitis virus L protein possesses the m RNA methyltransferase activities. *Virology*, **163**, 222-225.

Holloway, B.P., and Obijeski, J.F. 1980. Rabies virus-induced RNA synthesis in BHK-21 cells. *J.Gen. Virol.* **49**, 181-195.

Hudson, L.D., Condra, C., and Lazzarini, R.A. 1986. Cloning and expression of a viral phosphoprotein: Structure suggests vesicular stomatitis virus NS may function by mimicking an RNA template. *J. Gen. Virol.* **67**, 1571-1579.

Huismans, H., van Dijk, A.A. and Bauskin, A.R. 1987. In vitro phosphorylation and purification of a nonstructural protein of bluetongue virus with affinity for single-stranded RNA. *J.Virol.* **61**, 3589-3595.

Hummeler, K., Koprowski, H. and Wiktor, T.J. 1967. Structure and development of rabies virus in tissue culture. *J. Virol.* **1**, 152-170.

Hunt, D.M., Smith, E.G. and Buckley, D.W. 1984. Aberrant polyadenylation by a vesicular stomatitis mutant is due to an altered L protein. *J. Virol.* **52**, 515-521.

Hwang, L.N., Englun, N. and Pattnaik, A.K. 1998. Polyadenylation of vesicular stomatitis virus mRNA dictates efficient transcription termination at the intercistronic gene junction. *J. Virol.* **72**, 1805-1813

Isaac, C.L. and Keene, J.D. 1982. RNA polymerase-associated interactions near template promoter sequences of defective interfering particles of vesicular stomatitis virus. *J. Virol.* **43**, 241-249.

Iseni, F., Baudin, F., Blondel, D. and Ruigrok, R.W.H. 2000. Structure of the RNA inside the vesicular stomatitis virus nucleocapsid. *RNA.* **6**, 270-281.

Iseni, F., Barge, A., Baudin, F., Blondel, D. and Ruigrok, R.W.H. 1998. Characterization of rabies

virus nucleocapsids and recombinant nucleocapsid-like structures. *J.Gen.Virol.* **79**, 2909-2919.

Iverson, L.E. and Rose, J.K. 1981. Localized attenuation and discontinuous synthesis during vesicular stomatitis virus transcription. *Cell.* **23**, 477-484.

Kawai, A. 1977. Transcriptase activity associated with rabies virion. *J. Virol.*, **24**, 826-835.

Kawai, A., Toriumi, H., Tochikura, T.S., Takahashi, T., Honda, Y. and Morimoto, K. 1999. Nucleocapsid formation and/or subsequent conformational change of rabies virus nucleoprotein (N) is a prerequisite step for acquiring the phosphatase-sensitive epitope of monoclonal antibody 5-2-26. *Virol.* **263**, 395-407.

Keene, J.D., Schubert, M., Lazzarini, R.A. and Rosenberg, M. 1978. Nucleotide sequence homology at the 3' termini of RNA from vesicular stomatitis virus and its defective interfering particles. *Proc. Natl. Acad. Sci. USA.* **75**, 3225-3229.

Keen, J.D., Thornton, B.J. and Emerson, S.U. 1981. Sequence specific contacts between the RNA polymerase of vesicular stomatitis virus and the leader RNA gene. *Proc. Natl. Acad. Sci. USA.* **78**, 6191-6195.

Kikkawa, U., Takai, Y., Minakuchi, R., Inohara, S. and Nishizuka, Y. 1982. Calcium-activated, phospholipid-dependent protein kinase from rat brain. *J. Boil. Chem.* **257**, 13341-13348.

King, A.A. and Crick, J. 1988. *Rabies.* Boston :Kluwer Academic.

King, A., Meredith, C. and Thompson, G. 1993. Canid and viverrid rabies viruses in South Africa. *Onderstepoort J. Vet. Res.*, **60**, 295-299.

Kitts, P.A., Ayres, M.D. and Possee, R.D. 1990. Linearization of baculovirus DNA enhances the recovery of recombinant virus expression vectors. *Nucleic Acid Res.*, **18**, 5667-5672

Klumpp, K., Ruigrok, R.W.H. and Baudin, F. 1997. Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. EMBO J. **16**, 1248-1257.

Krebs, J.W., Wilson, M.L. and Childs, J.E. 1995. Rabies - epidemiology, prevention and future research. *J.Mammalogy*, **76**, 681-694.

Kurilla, M.G., Cabradilla, C.D., Holloway, B.P. and Keene, J.D. 1984. Nucleotide sequence and host La protein interactions of rabies virus leader RNA. *J.Virol.* **50**, 773-778.

Kouznetzoff, A., Buckle, M. and Tordo, N. 1998. Identification of a region of the rabies virus N protein involved in direct binding to the viral RNA. *J.Gen.Virol.* **79**, 1005-1013.

Kolakofsky, D., Vidal, S. and Curran, J. 1991. *The Paramyxoviruses.* New York/London: Plenum.

Kozak, M. 1986. The scanning model of translation: an update. *J.Cell Biol.* **108**, 229-241.

Kurilla, M. and Keene, J.D. 1983. The leader RNA of vesicular stomatitis virus is bound by a cellular protein reactive with anti-La lupus antibodies. *Cell.* **34**, 837-845.

Kurilla, M.G., Piwnica-Worms, H. and Keene, J. 1982. Rapid and transient localization of the leader RNA of vesicular stomatitis virus in the nuclei of infected cells. *Proc. Natl. Acad. Sci. USA.* **79**, 5240-5244.

Laemmli, U.K. 1970. Cleaving of the structural proteins during the assembly of the head of bacteriophage T4.*Nature*, **227**, 680-685.

Lamb, R.A. and Kolakofsky, D. 1996. *Virology.* Philadelphia: Lippincott-Raven Publishers.

Larson, J.K. and Wunner, W.H. 1990. Nucleotide and deduced amino acid sequences of the nominal nonstructural phosphoprotein of the ERA, PM and CVS-11 strains of rabies virus. *Nucleic Acids Res.* **18**, 7172.

Leppert, M., Rittenhouse, L., Perrault, J., Summers, D.F. and Kolakofsky, D. 1979. Plus and minus strand leader RNAs in negative strand virus-infected cells. *Cell.* **18**, 735-747.

Li, T. and Pattnaik, A.K. 1999. Overlapping signals for transcription and replication at the 3' terminus of the vesicular stomatitis virus genome. *J. Virol.* **73**, 444-452.

Luckow, V.A., Lee, S.C., Barry, G.F. and Olins, P.O. 1993. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in *Escherichia coli. J. Virol.*, **67**, 4566.

Marcus, P.I. and Sekellick, M.J. 1975. Cell killing by viruses. II. Cell killing by VSV: a requirement for virion derived transcription. *Virology*, **63**, 176-190.

Marcus, P.I., Sekellick, M.J., Johnson, L.D. and Lazzarini, R.A. 1977. Cell kiling by viruses, V. Transcribing defective interfering particles of VSV function as cell-killing particles. *Virology*, **82**, 242-246.

Maree, F.F. 2001. Multimeric protein structures of African horsesickness virus and their use as antigen delivery systems. Ph.D. thesis, University of Pretoria, Pretoria.

Marston, F.A.O. 1988. The purification of eukaryotic polypeptides synthesized in *Escherichia coli. Biochem. J.* **240**, 1-12.

Matsumoto, S.1974. Morphology of rabies virion and cytopathology of virus infected cells. *Symp. Ser. Immunobiol. Handb.* **21**, 25-34.

McGowan, J.J., and Wagner, 1981. Inhibition of cellular DNA synthesis by vesicular stomatitis

virus. *J. Virol.* **38**, 356-367.

McGowan, J.J., Emerson, S.U. and Wagner, R.R. 1982. The plus-strand leader RNA of VSV inhibits DNA-dependent transcription of adenovirus and SV40 genes in a soluble whole cell extract. *Cell*, **28**, 325-333.

Mebatsion, T., Weiland, F. and Conzelmann, K.K. 1999. Matrix protein of rabies virus is responsible for the assembly and budding of bullet-shaped particles and interacts with the transmembrane spike glycoprotein G. *J. Virol.* **73**, 242-250.

Morimoto, K., Akamine, T., Takamatsu, F. and Kawai, A. 1998. Studies on rabies virus RNA polymerase: 1. cDNA cloning of the catalytic subunit (L protein) of avirulent HEP-Flury strain and its expression in animal cells. *Microbiol.Immunol.* **42**, 485-496.

Nadin-Davis, S.A., Huang, W. and Wandeler, A.I. 1997. Polymorphism of rabies viruses within the phosphoprotein and matrix protein genes. *Arch.Virol.* **142**, 979-992.

Naito, S. and Ishihama, A. 1976. Function and structure of RNA polymerase from vesicular stomatitis virus. *J. Biol. Chem.* **251**, 4307-4314

Nel, L., Jacobs, J., Jaftha, J., Meredith, C. 1997. Natural spillover of a distinctly Canidae-associated biotype of rabies virus into an expanded wildlife host range in southern Africa. *Virus Genes.* **15**, 79-82.

Ogita, K., Ono, Y., Kikkawa, U. and Nishizuka, Y. 1991. Expression, separation, and assay of protein kinase C subspecies. *Methods in Enzymology.*, **200**, 228-234.

Olivier, M. 1997. Rabies associated with wildlife in the family *Viverridae* in South Afirca: Molecular analysis of the virus nucleoprotein and glycoprotein encoding genes. M.Sc(Agric) thesis, University of Pretoria, Pretoria.

Pattnaik, A.K., Hwang, L., Li, T., Englund, N., Mathur, M., Das, T. and Banerjee. A.K. 1997. Phosphorylation within the amino-terminal acidic domain I of the phosphoprotein of vesicular stomatitis virus is required for transcription but not for replication. *J. Virol.* **71**, 8167-8175.

Patton, J.T., Davis, N.L. and Wertz, G.W. 1984. N protein alone satisfies the requirement for protein synthesis during RNA replication of vesicular stomatitis virus. *J.Virol.* **49**, 303-309.

Peluso, R.W. and Moyer, S.A. 1988. Viral proteins required for the *in vitro* replication of vesicular stomatitis virus defective interfering particle genome RNA. *Virology.* **162**, 369-376

Possee, R.D. and Howard, S.C. 1987. Analysis of the polyhedrin gene promoter of the *Autographa californica* nuclear polyhedrosis virus. *Nucleic Acids Res.* **15**, 10233-48.

Préhaud, C., Harris, R.D., Fulop, V., Koh, C., Wong, J., Flamand, A. and Bishop, D.H.L. 1990.

Expression, characterization, and purification of a phosphorylated rabies nucleoprotein synthesized in insect cells by Baculovirus vectors. *Virology*, **178**, 486-497.

Préhaud, C., Nel, K. and Bishop, D.K. 1992. Baculovirus expressed rabies virus M1 protein is not phosphorylated: it forms multiple complexes with expressed rabies N protein. *Virology*, **189**, 766-770.

Reid-Sanden, F.L., Sumner, J.W., Smith, J.S., Fekadu, M., Shaddock, J.H. and Bellini, W.J. 1990. Rabies diagnostic reagents prepared from a rabies N gene recombinant expressed in Baculovirus. *J. Clin. Micorbiol.*, **28**, 858-863.

Rose, J.K. 1980. Complete intergenic and flanking gene sequences from the genome of vesicular stomatitis virus. *Cell*, **19**, 415-421.

Sacramento, D., Bourhy, H., and Tordo, N. 1991. PCR technique as an alternative method for diagnosis and molecular epidemiology of rabies virus. *Mol. Cell. Probes.* **5**, 229-240.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Molecular cloning: A laboratory manual. Second edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Schuber, M., Harmison, G.G., Richardson, C.D. and Meier, E. 1985. Expression of a cDNA encoding a functional 241-kilodalton vesicular stomatitis virus RNA polymerase. *Proc. Natl. Acad. Sci. USA.* **82**, 7984-7988.

Schubert, M., Keene, J.D., Herman, R.C. and Lazzarini, R.A. 1980. Site of the vesicular stomatitis virus genome specifying polyadenylation and the end of the L gene mRNA. *J. Virol.* **34**, 550-559.

Schnell, M.J., Buonocorek L., Whitt, M.A. and Rose, J.K. 1996. The minimal conserved transcription stop/start signal promotes stable expression of a foreign gene in vesicular stoatitis virus. *J.Virol.* **70**, 2318-2323.

Schoehn, G., Iseni, F., Mavrakis, M., Blondel, D. and Ruigrok, R.W.H. 2001. Structure of recombinant rabies virus nucleoprotein-RNA complex and identification of the phosphoprotein binding site. *J. Virol.* **75**, 490-498.

Simonsen, C.C., Batt-Humphries, S. and Summers, D.F. 1979. RNA synthesis of vesicular stomatitis viurs-infected cells: in vivo regulation of replication. *J. Virol.* **31**, 124-132.

Smith, D.B. and Johnson, K.S. 1989. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. *Gene*, **67**, 31-39.

Sokol, F. and Clark, H.F. 1973. Phosphoproteins, structural components of Rhabdoviruses. *Virology.* **52**, 246-263.

Spadafora, D., Canter, D.M., Jackson, R.L. and Perrault, J. 1996. Constitutive phosphorylation of

the vesicular stomatitis virus protein modulates polymerase complex formation but is not essential for transcription or replication. *J. Virol.* **70**, 4538-4548.

Sprague, J., Condra, J.H., Arnheiter, H. and Lazzarini, R.A. 1983. Expression of a recombinant DNA gene coding for the vesicular stomatitis virus nucleocapsid protein. *J.Virol.* **45**, 773-781.

Stillman, E.A. and Whitt, M.A. 1997. Mutational analysis of the intergenic dinucleotide and the transcriptional start sequence of vesicular stomatitis virus (VSV) define sequences required for efficient termination and initiation of VSV transcripts. *J.Virol.* **71**, 2127-2137.

Swanepoel, R., Barnard, B.J.H., Meredith, C.D., Bishop, G.C., Brückner, G.K., Foggin, C.M. and Hübschle, O.J.B. 1993. Rabies in southern Africa. *Onderstepoort J. Vet. Res.* **60**, 325-346.

Takacs, A.M. and Banerjee, A.K. 1995. Efficient interaction of the vesicular stomatitis virus P protein with the L protein or the N protein in cells expressing the recombinant proteins. *Virology.* **208**, 821-826.

Takacs, A.M., Barik, S., Das, T. and Banerjee, A.K. 1992. Phosphorylation of specific serine residues within the acidic domain of the phosphoprotein of vesicular stomatitis virus regulates transcription *in vitro*. *J. Virol.* **66**, 5842-5848.

Takacs, A.M., Das, T. and Banerjee, A.K. 1993. Mapping of interacting domains between the nucleocapsid protein and the phosphoprotein of vesicular stomatitis virus by using a two-hybrid system. *Proc. Natl. Acad. Sci. USA.* **90**, 10375-10379.

Takamatsu, F., Asakawa, N., Morimoto, K., Takeuchi, K., Eriguchi, Y., Toriumi, H. and Kawai, A. 1998. Studies on the rabies virus RNA polymerase: 2. Possible relationships between the two forms of the non-catalytic subunit (P protein). *Microbiol.Immunol.* **42**, 761-771.

Theron, J., Uitenweerde, J.M., Huismans, H. and Nel, L.H. 1994. Comparison of the expression and phosphorylation of the non-structural protein NS2 of three different orbiviruses: evidence for the involvement of an ubiquitous cellular kinase. *J.Gen.Virol.* **75**, 3401-3411.

Theron, J., Huismans, H. and Nel, L.H. 1996. Site-specific mutations in the NS2 protein of epizootic haemorrhagic disease virus markedly affect the formation of cytoplasmic inclusion bodies. *Arch. Virol.*, **141**, 1143-1151.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, **25**, 4876-4882.

Tordo, N. and Kouknetzoff, A. 1993. The rabies virus genome: an overview. *Onderstepoort J. Vet. Res.* **60**, 263-269.

Tordo, N. and Poch, O. 1988. Structure of rabies virus. In: *Rabies.* J.B. Campbell, and K.M.

Charlton, Eds, Kluwer Academic, Boston, pp 25-45.

Tordo, N., Poch, O., Ermine, A. and Keith, G. 1986a. Primary structure of leader RNA and nucleoprotein genes of the rabies genome: segmented homology with VSV. *Nucleic Acids Res.* **14**, 2671-2683.

Tordo, N., Poch, O., Ermine, A., Keith, G. and Rougeon, F. 1986b. Walking along the rabies genome: is the large G-L intergenic region a remnant gene? *Proc. Natl. Acad. Sci. USA.* **83**, 3914-3918.

Tuffereau, C., Fischer, S. and Flamand, A. 1985. Phosphorylation of the N and M_1 proteins of rabies virus. *J.Gen.Virol.* **66**, 2285-2289.

Vidal, S. and Kolakofsky, D. 1989. Modified model for the switch from Sendai virus transcription to replication. *J.Virol.* **63**, 1951-1958.

Von Teichman, B.F., Thompson, G.R., Meredith, C.D. and Nel, L.H. 1995. Molecular epidemiology of rabies in South Africa: evidence for two distinct virus groups. *J. Gen. Virol.* **76**, 73-82.

Wang, Y. and Zhang, X. 1999. The nucleocapsid protein of coronavirus mouse hepatitis virus interacts with the cellular heterogeneous nuclear ribonucleoprotein A1 *in vitro* and *in vivo*. *Virology.* **265**, 96-109.

Weck, P.K. and Wagner, R.R. 1978. Inhibition of RNA synthesis in mouse myeloma cells infected with vesicular stomatitis virus. *J. Virol.* **25**, 770-780.

Weck, P.K. and Wagner, R.R. 1979. Transcription of VSV is required to shut off cellular RNA synthesis. *J. Virol.* **30**, 410-413.

Wertz, G.W., Davies, N.L. and Patton, J. 1987. *The rhabdoviruses*. New York: Plenum.

Wertz, G.W. and Youngner, R.R. 1972. Transcription of VSV is required to shut off cellular RNA synthesis. *J. Virol.* **25**, 770-780.

Whelan, S.P.J. and Wertz, G.W. 1999a. Regulation of RNA synthesis by the genomic termini of vesicular stomatitis virus: identification of distinct sequences essential for transcription but not replication. *J.Virol.* **73**, 297-306.

Wiktor, T.J. and Koprowski, H. 1974. Rhabdovirus replication in encucleated host cells. *J. Virol.* **14**, 300-306.

Wilusz, J., Kurilla, M. and Keene, J.D. 1983. A host protein (La) binds to a unique species of minus-sense leader RNA during replication of vesicular stomatitis virus. *Proc. Natl. Acad. Sci. USA.* **80**, 5827-5831.

Wunner, W.H. 1991. *The natural history of rabies*. 2nd ed. Boston: CRC press.

Yan, J.X., Packer, N.H., Gooley, A.A. and Williams, K.L. 1998. Protein phosphorylation: technologies for the identification of phosphoamio acids. *J.Chromatography A.* **808**, 23-41.

Yang, J., Hooper, D.C., Wunner, W.H., Koprowski, H. Dietzschold, B. and Fu, Z.F. 1998. The specificity of rabies virus RNA encapsidation by nucleoprotein. *Virol.* **242**, 107-117.

Yang, J., Koprowski, H., Dietzschold, B. and Fu, Z.F. 1999. Phosphorylation of rabies virus nucleoprotein regulates viral RNA transcription and replication by modulating leader RNA encapsidation. *J.Virol.* **73**, 1661-1664.