Chapter 1:
Literature Review

1.1 General introduction to the disease

Foot-and-mouth disease (FMD) virus is a highly contagious pathogen responsible for one of the most economically devastating diseases affecting cloven-hoofed animals (Bachrach, 1978). FMD threatens the economies of southern African countries, like Botswana, Namibia, Zimbabwe and South Africa that are dependant on agricultural export of animal origin. The disease is characterised by a non-lethal infection culminating in temporary oral and pedal vesicles, which can result in a significant reduction in the production of meat or dairy products (Shahan, 1962).

Due to the infectious nature of the virus, trade embargoes are enforced against regions where FMD is endemic (Bachrach, 1968; Bachrach, 1978). Countries that are free of FMDV maintain rigid quarantine and import restrictions to prevent the introduction of the disease, while control programs to eradicate FMD include the slaughter of infected or exposed animals and vaccination using chemically inactivated vaccines.

The virus is classified in the family *Picornaviridae*, which contains nine genera *viz.* enterovirus, rhinovirus, cardiovirus, hepatovirus, parechovirus, erbovirus, kobuvirus, teschovirus and aphthovirus (van Regenmortel *et al.*, 2000). Historically FMDV has been the only member of the aphthovirus genus, but recently Equine Rhinitis A virus, formerly known as Equine Rhinovirus 1, was included within this group and FMDV appointed as the type species of the aphthoviruses.

There are seven recognised FMDV serotypes of which six occur on the African continent. The South African Territories (SAT) types 1, 2 and 3 are three serotypes unique to sub-Saharan Africa and are predominantly responsible for outbreaks of the disease in the region (Thomson, 1994). In addition of these, serotypes A, O and C are also occasionally found to be associated with outbreaks in Africa.

Under experimental conditions the viruses show little or no cross-protection against infection (Brooksby, 1982). This would account for reports that a single dose of monovalent vaccine fail to protect against challenge with heterologous virus (Cartwright *et al.*, 1982). The level of serological intratypic variation differs for each of the seven FMDV

serotypes. A number of identifiable subtypes exist for both serotypes A and O while intratypic variation within the C and Asia-1 serotypes appear to be markedly less. Despite the high degree of variation within the SAT types, no serological distinct subtypes have been identified for these serotypes (Esterhuysen *et al.*, 1988). Traditionally subtyping has been based on results generated by cross-neutralization assays and compliment fixation tests and shows good correlation with cross-protection in vivo (Thomas, 1986).

Molecular characterisation of field isolates have aided significantly in broadening our understanding of the epidemiology of the disease and enabled researchers to more effectively address the problems associated with the disease. A characteristic feature of FMDV epidemiology is the frequent emergence of new strains with altered antigenic properties and is due to the genetic instability of the virus (Domingo *et al.*, 1996). Although the European types of FMDV have been studied intensively, little is known about the genetic makeup of the SAT type viruses.

1.2 Epidemiology of the disease

1.2.1 Geographical distribution and recent history of FMDV

FMD has a global distribution occurring in South America, Europe, Asia, the Middle East, the Far East and Africa. The geographical distribution of the seven virus types is heterogeneous. Types A, O and C is widely distributed and has been responsible for outbreaks of FMD in South America, Europe, Africa, the Middle East and Far East. In contrast, the geographical distribution of the Asia 1 serotype is relatively restricted, since it is confined to the Far and Middle East and the Balkan states (Daborn, 1982; Rweyemanu *et al.*, 1982). All three SAT types are largely restricted to sub-Saharan Africa although there have been incursion of SAT 1 and 2 into North Africa and the Middle East respectively (Pereira, 1981).

FMD is classified as a List A disease defining it as a transmissible disease of one or more animal species, that has the potential for very serious and rapid spread, irrespective of national borders, that it is of serious socio-economic or public health consequence and that it is of major importance in the international trade of animals and animal products. Countries affected by the disease is required to report all cases of FMDV to the Office

International des Epizooties (OIE) as often as necessary to comply with Articles 1.1.3.2 and 1.1.3.3. of the International Animal Health Code. The FMD status of individual countries are continually monitored by both the World Reference Laboratory (Pirbright) UK, and the OIE. A graphical representation of the current global distribution of FMDV as published by the EIO is given in Figure 1.2.1.

In Europe and North America the effects of FMD have historically been great and vast amounts of money have been spent in attempts to eradicate or control the incidence of the disease. Northern America is considered to be free of FMD while in northern Europe incidence of the disease has decreased to such an extent that the disease is now considered to be exotic to all member states of the European Economic Community (Kitching, 1998). Despite recent reports of the eradication of the disease in western Europe and the declaration of Paraguay and Argentina as "Free from FMD with vaccination" areas, little progress have been made in the global control of FMDV. The impact of these and other significant advances, have been offset by the reappearance of the FMD in regions such as Malaysia and the introduction of the disease into areas formally free of FMDV such as Taiwan. Political and socio-economic instability in the Independent States of the former Soviet Union and frequent outbreak in the countries of southern Europe constantly pose a threat of large outbreaks in Europe (Kitching, 1998).

The continual global threat posed by FMD is highlighted by recent developments in Europe. Since the cessation of vaccination in the European union in 1991, there have been outbreaks of FMD in Italy in 1993 and in Greece in 1994 and 1996. Sporadic outbreaks of FMD were also recorded in Bulgaria, Russia, Albania and the former Yugoslavia at different times over the same period. Of particular concern to Europe is the situation in Turkey. Until 1989 a buffer zone existed on the northern border of Turkey, in which susceptible animals were routinely vaccinated. The zone was moved into Anatolia and vaccination discontinued in Turkish Thrace. However, due to the endemic status of FMD in Anatolia, outbreaks soon spread into Thrace posing a renewed threat to the Balkan states north of Turkey. Due to the nomadic culture of the people of the Middle East, it has proven difficult for countries to control the disease in isolation. In addition to this, wars and a general lack of infrastructure within the region have hampered efforts to control FMD.

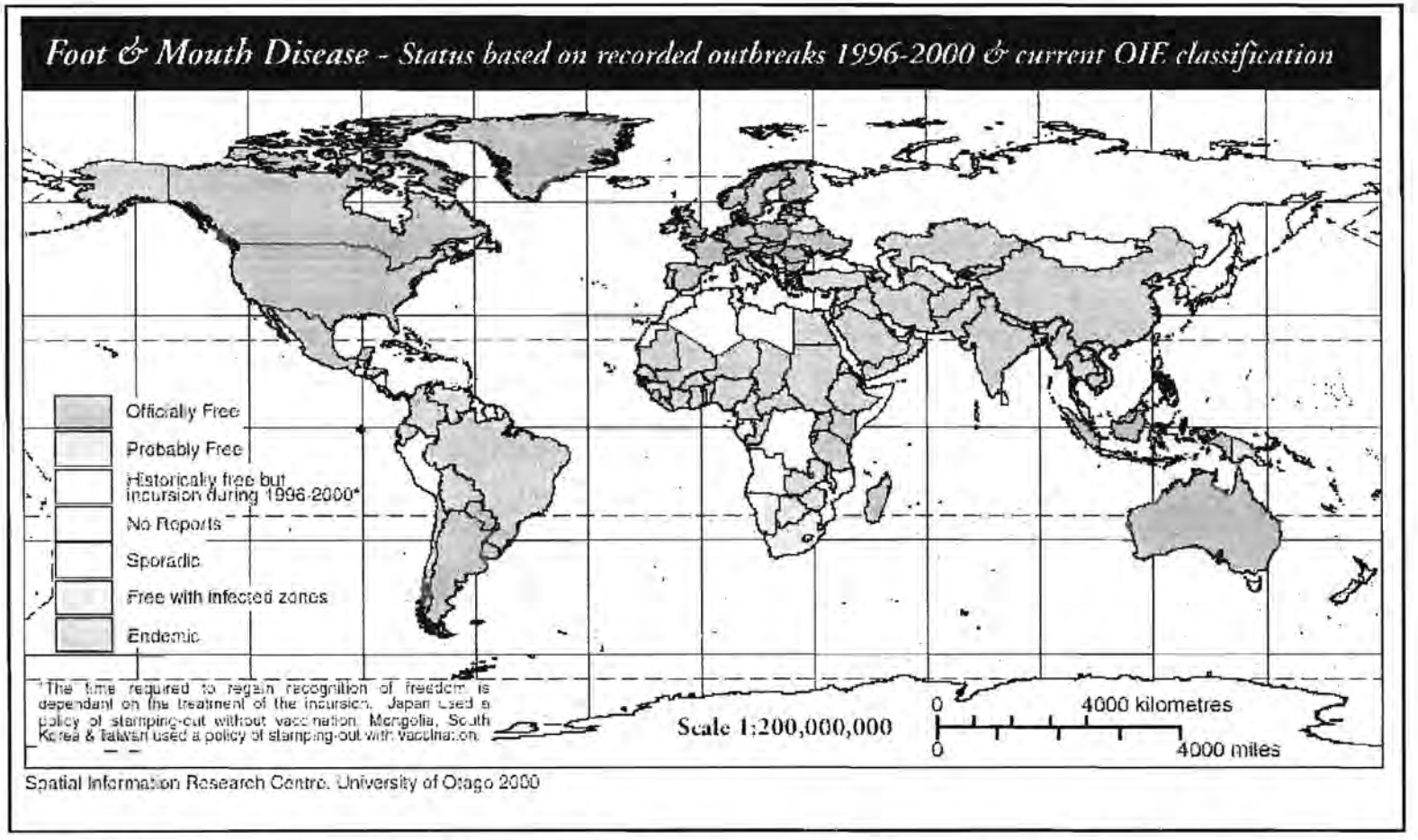

Figure 1.2.1. The global distribution of Foot-and-mouth disease virus. Spatial Information Research Centre. University of Otago, 2000. http://www.fao.org/ag/AGA/AGAH/EUFMD/findmaps/Global.htm.

FMD serotypes A, O, C and Asia-1 are widespread through India and the Far East. Although FMD is endemic to most of the region, Malaysia had been free of FMD up until 1990. In 1991 an isolated outbreak occurred in quarantine station on Malaysian border with Thailand. This was however followed by annual outbreaks in the surrounding areas between 1993 and 1996. Since then sporadic outbreaks of the disease have occurred in Malaysia with the most recent involving an isolated outbreak of FMD in the southern province of Kar.

A significant event in the recent history of FMD was the report of an outbreak in Taiwan during 1997. Until then, Taiwan had been free of FMDV. Although early outbreaks were brought under control by the mass slaughter of pigs, FMD is still prevalent in the region with cases of FMD occurring as recent as February 2000. Genetic characterisation of the viruses responsible for outbreaks in the region revealed that they significantly differ from historical strains of FMD isolated in the Far East and reconfirmed the importance of genetic evolution in the epidemiology of the disease.

In contrast to the situation in Europe and Asia, the incidence of FMDV in South America has greatly reduced. This is mainly due to marked improvement of in the efficiency of FMD control programs in the region. Several countries in the region have remained free of FMD for consecutive years. The last outbreak of FMDV in Uruguay occurred in 1990 and in 1993 the country was awarded the then unique status of "free of FMD with vaccination". In 1994 Uruguay ceased prophylactic vaccination and in 1995 declared itself free of FMD. Argentina and Paraguay have report similar successes in controlling FMD and was declared "free of FMD with vaccination" in 1997. Regional co-operation to control FMD has undoubtedly contributed to the general decrease of FMD incidence in South America (Kitching, 1998).

FMD is endemic to the most part of sub-Saharan Africa. It is however difficult to accurately determine the distribution of the disease on the subcontinent due to inconsistent reporting of the disease in several countries (Thomson, 1994). SAT type 1, 2 and 3 have been responsible for the overwhelming majority of outbreaks in cattle in sub-Saharan Africa, although sporadic outbreaks of FMD have been caused by serotypes A, O and C. Of the SAT types, the highest incidence in domestic animals is caused by SAT 2 (48%).

Outbreaks of SAT 1 make up approximately 36% of cases in cattle, while SAT 3 are only responsible for the remaining 16% of outbreaks. SAT 3 has a restricted distribution and has never been diagnosed in domestic animals outside of southern Africa (Thomson, 1994).

Extensive phylogenetic characterisation of SAT1 viruses isolated from African buffalo has revealed that these viruses evolve independently in at least three distinct geographical locations within southern Africa and can broadly be grouped in one of three topotypes (Bastos *et al.*, 2001). Most southern African countries have a single SAT 1 topotype within their borders, which is shared by at least two neighboring countries. Similar observations of strict geographical clustering of strains have been also reported for the SAT 2 and 3 type viruses respectively (Bastos 1999, personal communication).

The implementation of improved control strategies in southern Africa has significantly decreased the incidence of the disease in domestic livestock. In South Africa, zoning and control of FMDV has lead to the restriction of FMD outbreaks to wildlife populations in the northeastern extremes of the country. However, after several years of the absence of the disease in domestic livestock two independent outbreaks occurred during the summer of 2000 in Mpumalanga and KwaZulu-Natal respectively. Eradication of the FMD from sub-Saharan Africa is unlikely due to the persistent infection of African buffalo (*Syncerus caffer*) in the region. Buffalo act as a potential source of infection of other susceptible wildlife and domestic species. Thus, the focus is on control of FMD in sub-Saharan Africa rather than the eradication.

1.2.2 Persistent infection in wildlife

African buffalo is the only species of wildlife known to be persistently infected by FMDV (Hedger, 1976). The vast majority of infections are however clinically inapparent. Infection in a single animal can be sustained for at least five years with the continuous circulations of virus in the pharyngeal region of the animal. In isolated herds, SAT 1and 2 isolates can stably be maintained for 24 years by continual circulation within the population (Condy *et al.*, 1985). In cattle as in buffalo, virus levels in the pharyngeal region begin to decline soon after initial infection (Burrows, 1966). Six months post infection virus levels become nominal, but virus may persist in 20% of the herd for up to

two years (Hedger, 1968). During this period, animals are not protected against superinfections (McVicar & Sutmöller, 1976; Van Bekkum *et al.*, 1959).

Although several other wildlife species periodically become infected, there is little or no evidence that any other species of wildlife are capable of sustaining FMD infections over extended periods of time. In the Kruger National Park (KNP) in South Africa, regular epidemics of FMD occur in Impala (*Aepyceros melampus*), a species of gazelle found throughout southern Africa. Outbreaks among Impala are frequently characterised by the onset of severe symptoms. Serological surveillance however, showed that the virus is not maintained in the population during interepidemic periods. Despite the fact that researchers have shown that under experimental conditions, it is in fact possible to transmit FMDV from acutely infected Impala to cattle, very little evidence exits to support speculation that outbreaks in cattle in areas surrounding the park originate from Impala (Meeser, 1962; Thomson *et al.*, 1984).

Infection rates in buffalo are exceptionally high (Hedger, 1972; Falconer and Child, 1975). African buffalo breed seasonally with most calves being born in midsummer. This causes the synchronised infection of juveniles resulting in the simultaneous onset of disease in more than 10% of a breeding herd (Thomson, 1994). This in turn leads to the significant increase in the environmental contamination and provides a source of infection for other species in the immediate vicinity. African buffalo occur in or adjacent to all endemic FMD areas in the subcontinent. In addition to this, high FMD seropositivity rates suggested that that buffalo act as a major source of infection of livestock (Condy *et al.*, 1969). Recent studies have unequivocally confirmed that African buffalo is the predominant source of infection for outbreaks in domestic and wildlife species (Dawe *et al.*, 1994; Bastos *et al.*, 2000).

1.2.3 Spread of infection

Although indirect transfer of virus occasionally involves contaminated objects or material, most often spread is affected by direct contact between infected animals actively excreting virus and susceptible hosts (Sellers, 1971; Rweyemanu *et al.*, 1982). FMD virus can routinely be detected in all physiological fluids of viraemic animals and excretion can occur as much as four days before the onset of clinical symptoms (Hyslop, 1970; Burrows

et al., 1971). With the exception of the oesophago-pharyngeal secretions, virus is not secreted in any of the major animal secretions or excretions for more than two weeks post-infection. (Graves *et al.*, 1971).

Of all secretions and excretions in acutely infected animals, saliva contains the highest concentration of infectious units. This is the major contributing factor in the cross-contamination of the immediate environment and subsequent spread of the disease in susceptible populations. Due to the large respiratory volume of cattle, these animals are especially vulnerable to infection *via* the respiratory tract (Sellers, 1971). Cattle can be infected by inhaling doses of the virus as low as 25 infections units (Donaldson *et al.*, 1987). The infectivity of aerosols is heavily dependent on the size of the droplets and is associated with larger droplets, which are most likely to be trapped in the upper respiratory tract. Strong circumstantial evidence exists that long-distance transmission of FMD is possible under favourable environmental conditions (Donaldson, 1979; Gloster *et al.*, 1982). However, reports of airborne spread of infection from primary outbreaks are restricted to outbreaks in northern Europe and no convincing evidence for its involvement of FMD incidence in southern Africa has been advanced.

Domestic pigs are exceptionally efficient in excreting FMD virus and are often intensively farmed, resulting in high densities of diseased animals during outbreaks and large scale contamination of the surrounding areas (Donaldson, 1983). In contrast with cattle, pigs are equally susceptible to infection by the oral and respiratory routes (Terpstra, 1972). And circumstantial evidence suggests that infection of pigs usually result from the feeding of contaminated food (Strohmaier, 1987). Although sheep and goats are highly susceptible to infection and excrete virus at significant levels, little is known about the role of these animals in the epidemiology of the disease in southern Africa. Both species is however considered to be crucial in the spread of the disease in Turkey and some European countries (Pay, 1988).

Other secretions that contain appreciable quantities of virus are those of the nasal, pharyngeal and urogenital tract. The importance of urine and faeces in the transmission of FMD is not yet clear. Only low levels of virus have been found in these excretions and are not considered as a significant source of infection during outbreaks (Cottral, 1969;

Gainaru, 1986). Semen, however, may contain virus levels of up to $10^{4.7}$ MLD_{50}/ml even prior to the development of visible lesion (Sellers, 1983). Recovery of infectious virus from the genital tract of infected steers makes transmission via the sexual route a distinct possibility.

Although there is no known insect vectors implicated in the spread the virus, circumstantial evidences suggest that face flies (*Musca* spp.) may be involved in the mechanical transmition of FMDV via the conjunctiva (Sutmöller & McVicar, 1973). However, preliminary investigations failed to generate any supporting evidence of this (Thomson, 1985).

Cattle are immune to re-infection by homologous virus for one to three years on average and occasionally for up to four-and-a-half years (Bachrach, 1968; Brooksby, 1982). Circumstantial evidence suggests that immunity post-infection may be less for the SAT types (Thomson, 1994). Little is known concerning the duration of immunity in other species, but experimental evidence suggests that it may be shorter than in cattle (Bachrach, 1968).

1.3 The structure and antigenic properties of FMDV

1.3.1 Three-dimensional structure

The FMD virion is roughly spherical and contains no lipids. It is composed of 60 protein subunits consisting of the four structural proteins, VP4, VP2, VP3 and VP1. Five copies of VP1 are clustered around the five-fold axis, while VP2 and VP3 are positioned around the two- and three-fold axes of symmetry. These three proteins assemble into a protomeric subunit and twelve subunits form the icosahedral virus particle (Finch & Klug, 1959). VP4 of picornaviruses is completely buried at the base of the protomer.

The three-dimensional structure of FMDV has been resolved by X-ray crystallography (Acharya *et al.*, 1989). Each of the structural proteins is folded into an eight-stranded β-barrel. For VP2 and 3 the axes of the barrel are in the plane of the capsid while the axis of VP1 is turned slightly upwards in the direction of the 5-fold axis. Residues 141 to 160 of VP1 have been identified as the major antigenic region of the virus and are able to induce

neutralising antibodies (Bittle *et al.*, 1982; Pfaff *et al.*, 1982). This region has been shown to be situated within a connecting loop structure projecting out onto the surface of the virus. With the exception of this region the virion appears to be smooth with an average radius of between 100Å and extending to 150Å at the outer most limit. The viral genome is complete encapsulated by the icosahedra particle and not exposed at the surface at any point.

1.3.2 Genome structure and organisation

The FMD virus genome consists of a single-stranded RNA molecule of approximately 8700 base pairs (bp) in length. The RNA has a positive polarity and thus acts directly as messenger-RNA. The protein-coding region of the genome is flanked on each end by a non-translated region. In typical mRNA fashion, the viral RNA is polyadenylated at the 3' end. The function of the polyA tail is not clear, but RNA molecules of polioviruses with shorter polyA tracts have been shown to exhibit a lower specific infectivity (Spector & Baltmore, 1974; Hruby & Roberts 1976; Frey & Strauss, 1978). However, in contrast to several other picornaviruses, the specific infectivity of FMDV is not affected by the length of the 3' polyA tract (Grubman *et al.*, 1979). The polyA tail may further be responsible for RNA stability by protecting the genome against enzymatic degradation within the host cell.

The 5' end of the viral genome carries a small viral protein known as the VPg. This protein is covalently attached to the 5' terminal pU-pU sequence of the RNA through a phosphodiester linkage involving the phenolic hydroxyl group of a tyrosine residue (Wimmer, 1982). Three different forms of the VPg have been observed in association with the FMDV genome (Forss & Schaller, 1982.). This corresponds to the presence of a three VPg gene tandem present within the coding region of the genome (King *et al.*, 1980). The VPg appears to play an important role during viral replication and possible packaging of the virus. The 5' non-translated region can broadly be separated into three distinctive regions. The first 400bp, known as the S fragment is capable of forming a large hairpin structure of which the precise function is not known. The S fragment is separated from the larger portion of the genome by a polyC tract ranging in size from 100 to 420 residues (Harris *et al.*, 1976; Rowlands *et al.*, 1978). The function of the polyC tract is not entirely known, but it has been shown to be involved in determining virulence in other picornavirus models (Costa Giomi *et al.*, 1984; Duke & Palmenberg, 1989) The remainder of the 5'

non-coding region consists of approximately 700bp and contain a series of three pseudoknots and the internal ribosome entry site (IBRS) (Belsham & Brangwyn, 1990). The IBRS is linked to the cellular 57kDa protein and plays a crucial role during translation (Luz & Beck, 1990; Luz & Beck, 1991).

Four functional regions can be identified within the protein-coding region of the genome (Robertson *et al.*, 1985). The first of these is the L region starting at the 5' initiation codon and ending at the start of the structural-protein-coding region. This region encodes the leader protease, a distinctive feature of the aphthoviruses. The L region is directly followed by the P1 region encoding the precursor of the structural proteins. The P1 region is approximately 3kb in length and contains the coding sequences for all four of the structural proteins. The remaining 2/3 of the genome encodes all of the non-structural proteins and can be separated into the P2 and P3 functional regions respectively. The RNA contains an open reading frame encoding a single polyprotein. A cascade of proteolytic processing events carried out by three viral proteases result in 12 mature gene products. A graphical representation of the genome organisation and translational products are given in Figure 1.3.1.

1.3.3 Translational products

Translation of the viral RNA is initiated by the binding of ribosomes directly to the internal ribosome-binding site within the non-coding region and results in the formation of a polysome. Synthesis of a polyprotein is mediated by the use of the host cell protein synthesis machinery. Full-length (260kDa) polyprotein is ordinarily not formed, since the proteolitic processing thereof is initiated prior to the termination of translation.

Translation is initiated with equal frequency from of two in-frame AUG codons situated within the L region of the genome (Fross *et al.*, 1984) and facilitate the formation of two alternative forms of the leader protease (Lab and Lb) sharing a common carboxyl terminus (Clarke *et al.*, 1987; Sanger *et al.*, 1987). During the initial stages of viral replication there is a marked decrease in translation of host mRNA. This is a consequence of the proteolytic cleavage of the 220kDa subunit of the cellular cap-binding complex (eIF-4). The cleavage is not carried out by a viral coded protease but by a latent host protease activated by the leader protease (Devaney *et al.*, 1989).

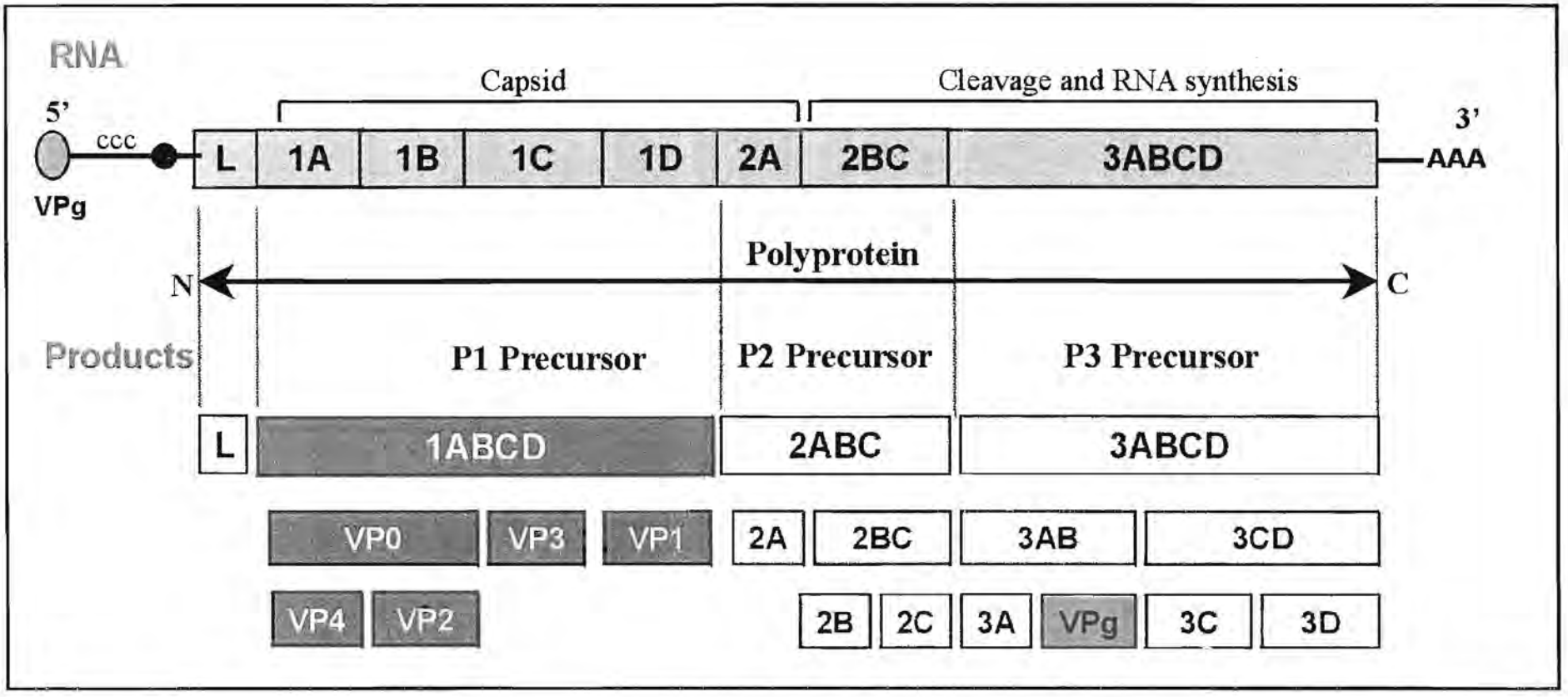

Figure 1.3.1 A graphical representation of the genome organisation and expression of FMDV. The three C's (ccc) at the 5' end of the genome indicates the position of the polycytilylic acid tract. The black dot at the same end indicates the position of the internal ribosyme entry site. The polyprotein is synthesised from left (N terminus) to the right (C terminus). All non-structural proteins are indicated in yellow. The four structural proteins that make up the proteinatious capsid structure are represented by the blue rectangles. The three A's (AAA) at the 3' end of the genome indicates the position of the poly A tail associated with the genome.

During the early processing event, two auto-catalytic cleavages occur involving the leader protease (Strabel and Beck, 1986) and the 2A protein (Clarke and Sanger, 1988) respectively, resulting in the formation of the mature leader protease and the separation of the P1 fragment from the rest of the gene products. It has been shown that a 19 amino acid sequence located within the 2A protein is sufficient to mediate the cleavage of the polyprotein immediately C-terminal to 2A (Ryan *et al.*, 1991). The 2A protein has furthermore been implicated in the shut off of host protein synthesis (Bernstein *et al.*, 1985).

The structural proteins 1A (VP4), 1B (VP2), 1C (VP3) and 1D (VP1), are contained within the P1 precursor and associate to form the protein coat subunits (protomers). The protomers of picornaviruses have been shown to be wedge-shaped with VP1, VP2 and VP3 occupying separate domains within the structure. Alkaline removal of VP2 and VP4 does not result in the dissociation of the protomer, suggesting that VP1 and VP3 contain important protomer-bonding domains which plays a curtail role in the stabilization of the virion (Katagiri *et al.*, 1971).

In addition to their structural role, the coat proteins are the major determinants of the antigenicity of the virus. Several findings indicate that VP1 is the major immunogenic protein of FMDV and is able to elicit neutralising antibodies in its purified form (Laporte *et al.*, 1973; Sobrino *et al.*, 1989). However, a relatively large dose of purified VP1 is required to induce similar levels of antibodies compared to levels induced by the intact virus, stressing the importance of conformational epitopes involving all the structural proteins (Laporte *et al.*, 1973; Bachrach *et al.*, 1975;).

Two conformational neutralising epitomes have been found within VP1 of type O viruses (Kitson, 1990) as well as a third epitope within the same protein of the A12 virus (Baxt *et al.*, 1989). In addition to the before mentioned conformational epitope, several liniar epitopes can also bee found within the VP1 protein (Crowther *et al.*, 1993). Experimental data showed that the major antigenic site is located in the region of amino acid residues 141 to 160. These amino acids coincide with the G-H loop loop structure, protruding from the particle surface (France *et al.*, 1994). Trypsin mediated removal of the loop structure not only results in the reduction of its antigenecity, but also causes a dramatic loss of

infectivity. The abolishment of infectivity is associated with the loss of ability to attach to susceptible cells, due to the removal of the RGD motif (amino acids 145 to 147). The RGD motif is thought to be the cell attachment site of the virus, reacting with cellular receptors on the surface of susceptible cells (Geysen *et al.*, 1985; Fox *et al.*, 1989).

In addition to the epitopes located on the VP1 protein, neutralizing epitopes can also be found on VP2 and VP3 proteins (Kitson *et al.*, 1990; Baxt *et al.*, 1989). VP4 is rarely exposed to the service of the particle and therefore in not subjected to immunogenic surveillance. For this reason VP4 is the least antigenic of the four structural proteins. However, a distinctive feature of the VP4 structural protein, as well as the P1 and VP0 precursors, is the N-terminal addition of a myristic acid to these proteins (Chow *et al.*, 1987). A cellular enzyme rather than an enzyme of viral origin mediates the addition of the n-tetradecanoic acid. The possible role of the myristate is not clear, but it is thought to be involved in the localization of the P1 precursor on the cellular membrane, assembly of the pentamenrs or the release of pentamers from the membrane (Rueckert, 1990).

The role of 2B and 2C is not quite clear, but initial investigations suggested that the 2B protein of rhinovirus 2 might be involved in host range determination (Yin & Lomax, 1983). Guanidine hydrochloride is known block the initiation of RNA synthesis and is able to inhibit viral replication of several picornaviruses at millimolar concentration (Caliguiri & Tamm, 1973; Crowther & Melnick, 1961; Rightsel *et al.*, 1961). Resistance to this drug have been mapped to the 2C region of FMDV (Saunders & King, 1982), as well as poliovirus (Anderson-Sillman *et al.*, 1984), suggesting that the 2C protein is involved in RNA synthesis (Li & Baltimore, 1988). It has also been shown that the introduction of site-specific mutations within the highly conserved 2C coding region, severely restricts the ability of the virus to replicate (Li & Baltimore, 1988). This result further point to the involvement of the 2C protein in RNA synthesis.

No information is available to clearly define the role of protein 3A in the picornavirus life cycle and its possible effect on the cellular metabolism. However, the 3AB polypeptide, the precursor of the genome-bound VPg protein (3B) and the 3A protein, stimulates the *in vitro* synthesis of poly(U) RNA, directed by the $3D^{pol}$ viral polymerase (Lama *et al.*, 1994) and is thought to act in a similar manner *in vivo*. The 3AB precursor is known to associate

with cellular membranes during replication and thereby direct the replication complex to the endoplasmic reticulum (Giachetti *et al.*, 1991; Xiang *et al.*, 1998). The association of the precursor with cellular membranes is mediated by a hydrophobic domain present in the 3A protein (Xiang *et al.*, 1998). Protein-membrane interaction of the 3A protein has been shown to cause a cytophatic effect due to an increase in membrane permeability, prevent surface expression of host proteins and to inhibit protein excretion from 3A expressing cells (Doedens & Kirkegaard, 1995; Doedens *et al.*, 1997).

Despite the fact that the 3AB precursor is required for effective viral replication of poliovirus, purified 3A protein does not show RNA binding or transactivation activity. This suggests that the 3A protein acts only as an anchor or carrier of the VPg protein, which in turn acts as a primer during viral RNA synthesis (Takegami *et al.*, 1983; Takada *et al.*, 1986; Giachetti *et al.*, 1991; Porter, 1993). If the 3AB precursor does in fact act as a co-factor of the viral polymerase, it would allow for the regulation of replication by means of proteolytic processing (Lama *et al.*, 1994). The recent emergences of naturally attenuated strains of FMDV have shed new light on the possible role of the 3A protein. In 1997, a devastating outbreak of FMD in Taiwan was caused by a type O virus exhibiting atypical virulence. It produced a high incidence of disease in swine but did however not affect cattle. Genetic characterization of this and other attenuated strains of FMDV, showed an altered state of the 3A protein to be the major determinant of the atypical epidemiology (Beard and Mason, 2000). It was therefore hypothesised that the 3A protein my play a significant role in determining host range specificity and virulence. The exact mechanism by which this could be effected is however not clear.

The 3CD precursor of FMDV can be cleaved in two different ways to yield either 3C + 3D or 3C' + 3D' (Doel *et al.*, 1978, Ruekert *et al.*, 1979). In the case of poliovirus the alternate cleavage occur at a Tyr-Gly site and is mediated by 2A (Toyda *et al.*, 1986). The 3C protease or its precursor 3CD, mediates the majority of the cleavages of the polyprotein with the exception of the auto-catalytic cleavages of the leader protease, the 2A protein and VP0 precursor (VP0 $\rightarrow$ VP4 + VP2) during maturation of the virion. This viral protease belongs to the family of cellular serine proteases. The catalytic triad was predicted to be His46-Asp84-Cys163 (Gorbalenya *et al.*, 1989) and confirmed by site-directed mutagenesis (Grubman *et al.*, 1995). In addition to the viral protein processing, the 3C

protease of FMDV cleaves the host-cell protein histone H3 and may be involved with shut-off process of host-cell transcription (Falk *et al.*, 1990). The final protein coded for by the virus, is the $3D^{pol}$ protein. This protein is cleaved form its precursor by the 3C protease and has as its main function the replication of the RNA genome. This RNA polymerase is required for copying the positive sense RNA strand to form the complementary minus stranded RNA and the consequential formation of additional the positive stranded viral RNA.

Besides the four capsid proteins, the VPg associated with the viral genome is the only protein of viral origin packaged into the mature virion. All other viral proteins are only produced during replication of the virus following the successful infection of a susceptible cell.

1.4 Variation within the FMDV genome

1.4.1 Overview of replication

Virus entry during FMD infection occurs due to receptor-mediated endocytosis. A crucial part of internalisation of the virion is the non-covalent interaction between the vitronectin host cell surface receptors (integrin $\alpha_v\beta_3$) and the Arg-Gly-Asp (RGD) protein ligand (Fox *et al.*, 1989). Internalisation are thought to occur due to the release of VP4 and the unfolding of hydrophobic regions of the VP2, 3 and 4 proteins that would result in the formation of a pore through which the RNA might be transferred to the cytoplasm (Evens & Almond, 1998).

Early on during infection the majority of viral RNA acts as mRNA, since viral replication is focused on protein synthesis. Levels of the replicative intermediate RNA increases 2 to 4 hours post-inoculation while the replicative form RNA accumulates slower and reaches lower levels during infection. RNA replication involves at least two viral coded proteins, the $3D^{pol}$ and the VPg. These two proteins as, well as a 67kDa host factor protein, is tightly associated with the smooth endoplasmic riticulum and forms the 27S replicon complex. RNA synthesis is template dependent and is thought to be primed by the VPg. The $3D^{pol}$ is a RNA dependent RNA polymerase and is responsible for the elongation of both positive and negative sense RNA during virus replication. Later, as RNA synthesis proceeds at a

constant rate, synthesis of positive-sense RNA predominates, and approximately 50% of these molecules are packaged into virions.

Shell assembly proceeds directly following the proteolytic cleavage of the P1 precursor. Immature protomers (5S subunits) are formed by the aggregation of VP0, VP3 and VP1. The 5S subunits assemble into pentamers followed by the assembly of 12 pentamers into either non-infectious RNA-containing provirion or naturally empty capsids. Formation of the infective 146S virions is achieved by the maturation cleavage of the VP0 precursor. Complete virus particles are ultimately released from the cell by infectious-mediated disintegration of the host cell.

It is important to note that the products of viral replication are rarely exact copies of the original virus involved in initial infection. The inherent error prone nature of viral replication is an important aspect of the biological evolution of the virus and contributes to the ability of the virus to adapt to the ever-changing cellular environment.

1.4.2 Mutations

Mutation is responsible for a significant portion of genetic variation observed after replication. Changes in the sequence of genomic nucleic acids occur at a higher frequency in single-stranded than double-stranded nucleic acids. Due to the absence of a complementary sequence, point mutations within single-stranded species become immediately effective. Mutation rates in RNA genomes are generally considerably higher than the rate observed for DNA (Steinhauer and Holland, 1987). High mutation rates observed for RNA are mainly due to the lack of an effective proofreading mechanism of RNA-dependent RNA polymerases.

Estimates of the mutation rate of FMDV were found to be in the order of $10^{-3.8}$ and $10^{-5.6}$ per base position depending on the method of investigation (Pringle, 1964; McCohen, 1981; McCohen, 1986). The rate at which mutations are fixed within the FMDV genome are heavily dependent on the time period and portion of the genome examined and were found to vary between 0.04×10^{-2} and 4.5×10^{-2} substitutions per nucleotide per year (Sobrino *et al.*, 1986). Mutations are most frequently observed in the three major capsid proteins *viz.* VP2, VP3 and VP1. In fact, mutation rates for the structural-protein-coding

regions of the genome are up to six-fold higher than for any other part of the genome (Sobrino *et al.*, 1986).

Studies on the accumulation of mutations in FMDV under field conditions revealed that in fully susceptible animals mutations only occur occasionally and that these mutations rarely persist in the progeny (McCohen, 1986). However, in partially immune populations mutations are rapidly accumulated. Nucleotide sequences of sequential isolates obtained from carrier animals revealed a similar pattern. However, specimens collected from carrier animals frequently yielded mixtures of mutated viruses (Gebauer *et al.*, 1988).

1.4.3 Quasispecies concept

In 1971, Eigen and his collaborators developed a mathematical model to describe the dynamics of RNA replication in which a RNA genome is considered to be a population of quasispecies. The model allows for the definition of each individual mutant within the population to be defined in terms of a probability. The first experimental evidence that RNA viruses may have a quasispecies distribution was generated using phage Qβ (Batschelet *et al.*, 1976, Domingo *et al.*, 1976). It was shown that up to 15 percent of individual isolates differed on average at 1 to 2 positions from the phage Qβ stock virus after 50 passages (Domingo *et al.*, 1978).

Despite the high level of heterogeneity within a virus population, high mutation rates are not always reflected in the pathology of infected individuals. Under constant conditions a single variant would dominate during infection causing a distinct pattern of pathology. However, any change in the conditions that in any way detrimentally affects the efficacy of replication of the dominant variant, would allow for the selection of a more fit quasispecies which would have the potential to replace the dominant variant (Hyslop, 1965; Domingo *et al.*, 1980, Matau *et al.*, 1987). The result is a highly flexible biological system, which allows the virus to evade constant immune surveillance and establish or mediate persistence through the continual selection of variant genomes. However, the effects of mutations are limited and rarely result in a large shift in the biological behaviour of the virus.

1.4.4 Recombination

Unlike mutation, recombination offers a mechanism by which large segments of genetic material can be exchanged between two viruses. Although recombination in DNA is well documented, evidence of RNA recombination is limited to only a few viral families. In 1977, Cooper reported that it was possible to construct a genetic map of the poliovirus genome using crosses between temperature sensitive mutants defective in coat protein formation and RNA synthesis. Recombination does not only occur under experimental conditions, but have also been shown to occur under natural infectious conditions. Kew and Nottay (1984) identified natural isolates of poliovirus that contained segments of all three virus serotypes. It is speculated that recombination occurs when a single cell is simultaneously infected with two or more FMD viruses (Pallansch *et al.*, 1980).

Biological evidence suggests that recombination of picornaviruses occur due to template switching during replication (Cooper *et al.*, 1974). The proposed mechanism by which this occurs, involves the detachment of the RNA polymerase from the original template and reattachment to an alternative template within close proximity to the replication machinery. The length of the picornavirus recombination map, 1.2% for poliovirus (Cooper, 1977) and 3% for aphthoviruses (McCohen, 1981), implies that recombination occurs with very high frequency. It is estimated that in the case of polio up to 20% of all viral genomes produced during a single cycle of replication may be recombinant. The frequency of recombination appears to be higher during synthesis of the negative strand (Kirkegaard and Baltimore, 1986). Although recombination between serotypes of FMDV do occur, crosses between different strains of the same serotype occur more readily (Tolskaya *et al.*, 1983; McCohen *et al.*, 1985).

King *et al.* (1988) showed that in addition to single cross-over events, double cross-overs were also occurring within the FMDV genome. The cross-overs were scattered throughout the genome, suggesting that it was a general rather than site-specific event (Fross *et al.*, 1984). Nucleotide sequence data of known cross-over points within the FMDV genome suggests that although perfect homology is not a prerequisite for recombination to occur, relatively high homology (70-80%) is required. Furthermore, cross-over points are generally situated within region of high secondary structure (Wilson *et al.*, 1988). Recombination can therefore probably occur throughout the genome.

Although there is substantial evidence that mutational and recombination events contribute substantially to the genetic heterogeneity of FMDV, the mechanisms that lead to the fixation of these mutation within successive populations are poorly understood. It is however clear that generation and the establishment of new variant populations play an important role in the evolution of the virus.

1.5 Aims of the Study

The study is composed of two distinct Sections. The first Section deals exclusively with the genetic characterisation of the structural-protein-coding region of a specific FMDV isolate originating from southern Africa. In the second Section, the relevance of the genetic variation observed for the 3A non-structural-protein-coding region of several African FMDV isolates originating from diverse geographical and host origins, is studied.

SAT 3 type viruses are rarely responsible for outbreaks of the disease in domestic livestock. This, coupled to the restricted geographical distribution of the serotype, are probable reasons why very little research on representative isolates of SAT 3 type viruses has been done. Nevertheless, SAT type 3 viruses pose a significant threat to commercial stock farming in southern Africa. This is especially pertinent in areas bordering wild life reserves where FMD is endemic. It has been shown that in the Kruger National park, South Africa, that a large portion of the resident African buffalo has significant titers of antibodies directed against SAT 3 type viruses. These animals represent a continual reservoir of virus capable of causing a severe outbreak among domestic animals in the surrounding areas.

The generation of genetic information with regards to the structural-protein coding region of the SAT type 3 viruses would greatly assist in the design of custom-made vaccines, which accurately reflects antigenic characteristics of prevalent isolates. The major objective of the study was to expand on the present knowledge concerning characteristics of the structural-protein-coding region of the SAT 3 type viruses by determining the nucleotide sequence of the entire P1 precursor of FMD virus KNP/10/90/3. This isolate was first isolated in 1990 from an oesophagus-pharyngeal sample collected during a routine survey of the endemic African buffalo population in the Kruger National Park of

South Africa and are currently being used in the formulation of vaccines to be used in southern Africa. We compared this region of the genome, consisting of the four structural proteins (VP1, VP2, VP3, and VP4), among different SAT serotypes and with that of representative isolates of serotypes A, O, and C.

The 3A non-structural protein of FMDV has been shown to be involved in viral RNA replication and death of infected cells. As previously discussed in Section 1.3.2, changes in 3A have been associated with altered host range and virulence in several picornaviruses. Although the 3A coding region has been shown to be highly conserved among European isolates, the sequence characteristics of the 3A of the SAT type viruses have not been investigated. In light of this, the nucleotide sequences of the 3A non-structural-coding sequence of several African FMD virus isolates were determined and comparatively analysed. We compared this region of the genome among different SAT serotypes and with that of European, South American and Asian isolates in order to assess the extent of genetic variation within the 3A coding region of naturally occurring viruses in sub-Saharan Africa.

Chapter 2:
Genetic characterisation of the structural-protein-coding region of foot-and-mouth disease virus KNP/10/90/3

2.1 Introduction

Nucleotide sequence determination has proven to be an invaluable tool in the genetic characterisation of foot-and-mouth disease virus (Palmenberg, 1989). Phylogenetic studies based on partial genomic sequences have aided scientists in their quest to understand the development of outbreaks and more accurately predict and combat the spread of the disease (Brown, 1987).

Phylogenetic characterisation of field isolates originating from sub-Saharan Africa, based on nucleotide sequence data, has revealed that these viruses evolve independently in different geographical regions complicating the design and production of universal vaccines which effectively address the antigenic variation within the SAT types (Vosloo *et al.*, 1995). The extension of effective FMD control to other countries in the region, as well as improved control in established areas thus requires the use of vaccines designed to address the antigenic characteristics of prevalent strains occurring in specific areas within the region (van Rensburg & Nel, 1999).

However, selection of prevalent, high producing and stable vaccine strains is a cumbersome and expensive process. To address this problem, new generation vaccines are being developed based on the ability of naked FMDV RNA to infect susceptible cells and result in the formation of intact virions (Rueckert, 1990). The first step in is approach is the construction of infectious full-length cDNA clone of a suitable vaccine strain. The antigenic characteristics of such a clone can easily be manipulated by exchanging the antigenic determinants of the virus i.e. the structural proteins with that of an alternative virus. The cDNA clone acts as a vector expressing the structural-protein cassette of an alternative virus inserted in place of the native structural-protein-coding region, resulting in the formation of a chimeric FMD virus exhibiting the antigenic characteristics of the alternative virus, while retaining the growth characteristics of the virus from which the clone was originally constructed (Sa-Carvalho *et al.*, 1997). The chimeric viruses can then be isolated and directly applied in current vaccine production protocols.

The major advantage of the approach is the ability to construct custom designed vaccines without having to adapt newly isolated strains to cell culture for the purpose of vaccine production. This strategy has previously been shown to be successful for both type A and

O serotypes (Zibert *et al.*, 1990; Reider *et al.*, 1993). Due to the genetic and phenotypic diversity of FMDV, a number of crucial factors that may influence the success of such a strategy in the southern African context however remain unknown. These include the role of proteolytic variation and cleavage site differences observed in isolates form different geographical areas.

Traditionally more emphasis has been placed on the characterisation of isolates occurring predominantly in Europe and South America. Although the structural-protein-coding region of several isolates representative of serotypes A, O and C have been successfully characterised (Forss *et al.*, 1984; Robertson *et al.*, 1985; Singh *et al.*, 1996), comparatively little is known of the SAT type viruses which are predominantly responsible for outbreaks in sub-Saharan Africa. To date only four reports have been published describing the genetic characteristics of the entire structural-protein-coding region of FMD viruses representative of the SAT serotypes (Knowles *et al.*, unpublished; Newman *et al.*, unpublished; Brown et al, 1989; van Rensburg and Nel, 1999; van Rensburg, unpublished). Of these only one deals with a representative isolate of SAT 3. Although the geographical distribution of the SAT type 3 viruses is restricted to southern Africa, it still poses a significant threat to the commercial farming in these regions. This chapter focuses on the genetic characteristics of KNP/10/90/3 and reports on the relevance of the date obtained towards elucidating the extent of genetic variation within the structural-protein-coding region of SAT 3 serotype. We furthermore briefly discuss the possible influence of genetic variation among the SAT isolates on the construction of custom-made vaccines based on the strategy described above.

2.2 *Materials and methods*[*]

2.2.1 Origin of the virus used in the study

The KNP/10/90/3 virus was first isolated in 1990 from an oesophagus-pharyngeal sample collected during a routine survey of the endemic African buffalo population in the Kruger National Park of South Africa. The virus was initially isolated on pig kidney cells (primary isolation), and cultured on IB-RS-2 cells for a further 3 passages. The virus was partially

[*] The recipes for the preparation of all buffers and reagents referred to in the texts are summarised in appendix A.4.

purified from the third passage and aliquots stored at –70°C. In addition to this isolate, the previously published structural-protein-coding region sequences of several other isolates were included in the analysis of the data and were obtained from the Genbank nucleotide sequence database. Details of all isolates used in the study are summarised in Table 2.2.1.

Table 2.2.1. List of viruses used in comparative studies of the structural-protein-coding region of the foot-and-mouth disease virus.

Virus	*Country*	*Year of Isolation*	*Species of Origin*	*Reference*
Serotype A				
A_{12}	United Kingdom	1932	Bovine	M10975
A_{22}	USSR	1964	Bovine	X74812
A_{10}	Argentina	1961		X00429
Serotype O				
O_1 Kaufbeuren	Germany	1965	Bovine	X00871
O_1 Campos/58	Brazil	1958		M89781
Serotype C				
C_1 Germany/c.26	Germany	1962		M90368
C_3 Resende/55	Brazil	1965	Bovine	M90381
Serotype Asia1				
India/63/72	India	1972	Bovine	Y09949
Lebanon/83	Lebanon	1983	Bovine	U01207
Serotype SAT1				
BOT/1/68/1	Botswana	1968	Bovine	Z98203
KNP/196/91/1*	South Africa	1991	Buffalo	AF283429
Serotype SAT2				
KEN/3/57/2	Kenya	1957	Bovine	AV0006
ZIM/7/83/2	Zimbabwe	1983	Bovine	AF136607
RHO/1/48/2	Zimbabwe	1948	Bovine	AV0006
Serotype SAT3				
BEC/1/65/3	Botswana	1965	Bovine	M28719
KNP/10/90/3#	South Africa	1990	Buffalo	AF286347

* The nucleotide and amino acid sequence data for KNP/196/91/1 have been submitted to the Genbank nucleotide sequence database and have been assigned the accession number AF283429. This sequence data was generously supplied by H.G. van Rensburg (unpublished results).

The nucleotide sequence has been submitted to the Genbank nucleotide sequence database and has been assigned the accession number AF286347.

2.2.2 Preparation of cell culture

All cell lines used in the study were subcultured as described by Freshney (1987). Briefly the procedure involved rinsing of the cell sheet, assisted detachment of the cells from the solid surface and re-incubation of the cells after the addition of fresh maintenance medium. A single Roux flask containing an IB-RS-2 cell line monolayer was initially rinsed with warmed (37°C) versene-trypsin after which the fluid was decanted and discarded, followed by the addition of 3ml fresh solution and the incubation of the cells at room temperature for approximately 5 minutes. Versene-trypsin acts as a detergent causing a lowering of the surface tension and thereby assists in the detachment of the cells from the glass surface. Following the incubation of the cells, 5ml Modified Eagle's maintenance medium, warmed to 37°C, was added to the flask and the cells resuspended by repeatedly dispensing the total volume of fluids using a 10ml glass syringe. Equal volumes (1ml) of the resulting cell suspension was transferred to sterile Roux flasks and 50ml, Modified Eagle's medium, containing 5% normal bovine serum, penicillin (1000 units/ml), neomycin (75µg/ml) and streptomycin (100µg/ml) added. The cells were incubated at 37°C for approximately four days, whilst continually being rocked before being used for virus propagation.

2.2.3 Propagation of the foot-and-mouth disease virus

A single Roux flask containing a confluent cell sheet was inoculated with 1ml of the stock virus (Section 2.1.1) at a low multiplicity of infection. The inoculated cells were overlaid with 50ml warmed Eagle's medium, containing penicillin (1000units/ml), neomycin (75µg/ml) and streptomycin (100µg/ml), but void of normal bovine serum. The flask was incubated at 37°C for approximately 24 hours, while gently being rocked. As soon as 100% cytopathic effect (CPE) was observed, 50ml of the medium now containing the viruses release from the cells during lysis, was pored off into a second Roux flask containing a fresh uninfected monolayer of IB-RS-2 cells. This flask was incubated at 37°C for 30 minutes to allow sufficient attachment of the virus to the uninfected cells. Following incubation, all fluids were decanted and discarded. Any cell debris and unattached virus was removed by adding 10ml phosphate buffer saline (pH 7.6 @ 25°C) and gently swirling the flask for 30 seconds. The rinsate was discarded before fresh maintenance medium (Modified Eagle's medium containing 1000units/ml penicillin, 75µg/ml neomycin and 100µg/ml streptomycin) was added to the monolayer. The infected cells were incubated for

3 to 4 hours at 37°C, until 80% CPE could be observed. At this point the virus was harvested as described in Section 2.2.4.

2.2.4 Partial purification of the foot-and-mouth disease virus

The FMD virus was recovered using a method similar to that described by Rico-Hesse *et al.* (1987) for the purification of poliovirus from cell culture. Infected cells were detached from the glass and resuspended as described above (Section 2.2.2). The suspension was then transferred to a 38ml polycarbonate ultracentrifuge tube (Beckman) and successively centrifuged for 10 minute periods at 2500, 5000 and 10 000 revolutions per minute respectively (Beckman SW28 rotor) to remove the cell debris. Subsequent to completion of each centrifugation step, the supernatant was transferred to a clean tube before proceeding to the next. At completion of the final centrifugation step the supernatant was again transferred to a clean tube and 2ml of a 30% sucrose solution (30% sucrose, 0.1M Tris/HCl, pH 7.6 @ 25°C) carefully placed at the bottom of the tube using a sterile 10ml glass syringe to form a distinct pellet within the aqueous phase. The tube was centrifuged at 28 000 rpm for 3 hours at 4°C using a Beckman SW28 rotor. The resulting supernatant was decanted and all residual fluid drained thoroughly using a P200 micropipet. A volume of 500µl ice-cold 1 x Tris/EDTA buffer (pH 7.6 @ 25°C) was added to the tube containing the pellet and incubated at 4°C for 5 minutes. On completion of incubation, the pellet was subjected to sonication for 5 seconds at an amplitude of 20 microns (Sonifier B-12 probe) whilST on ice. Aliquots of the resuspended pellet was transferred to a 1.5ml RNase-free micro-test tube and either stored at -70°C or processed immediately.

2.2.5 RNA extraction from purified foot-and-mouth disease virus

Volumes of 5µl β-mercaptoethonol and 50µl of a 10% sodium dodecyl sulphate solution were added to the sonicated pellet. The sample was briefly vortexed after which 250µl chloroform/isoamielalcohol solution at a volume-to-volume ratio of 24:1 and 250µl 0.1M Tris/HCl-saturated phenol were added. The contents of the tube were mixed by inversion and fractionated by centrifuging the tube at maximum revolutions for 5 minutes (Hettich bench top centrifuge, equal to 16000xg). The aqueous phase was transferred to a clean 1.5ml micro-test tube, taking care not to transfer any protein precipitate present at the interface. This was followed by the addition of 500µl fresh 0.1M Tris/HCl-saturated

phenol to the transferred aqueous phase. The contents of the tube were gently mixed by inversion and centrifuged at 16000xg for 4 minute. The aqueous phase was transferred to a clean 1.5ml micro-test tube and two volumes of diethylether added to it. The sample was centrifuged at maximum revolutions for 4 minutes to allow proper separation. The aqueous phase was transferred to a clean tube and incubated at room temperature for approximately 5 minutes to allow contaminating diethylether to evaporate. The viral RNA was concentrated by precipitation as describe below in Section 2.2.6.

2.2.6 Precipitation of RNA

The RNA was precipitated from the aqueous phase by the addition of 3M sodium acetate (pH 5.2 @ 25°C) to a final concentration of 0.3M and 2.5 volumes of absolute ethanol. The sample was incubated overnight at -20°C. The precipitate was concentrated by centrifugation (16000xg for 30 minutes). The resulting pellet was washed by discarding the supernatant and adding 100µl of an 80% ethanol solution. The ethanol was discarded after brief centrifugation. The excess ethanol was removed by evaporation under vacuum and the pellet resuspended in 50µl Tris/EDTA buffer (pH 7.6 @ 25°C) containing 5U RNase inhibitor (5U/µl, Promega).

The concentration and purity of the extracted viral RNA was determined by measuring the optical density of the solution at 260nm and 280nm, respectively. Only preparations with an OD_{260}/OD_{280} ratio of 1.6 to 1.9 were used for further manipulation. The RNA was aliquoted into smaller volumes of 5µl and stored as –70°C.

2.2.7 Preparation of cDNA

The oligonucleotide used to initiate the reverse transcription was previously described by Beck and Strohmaier (1987). It is designed to anneal to an area 33 nucleotides downstream of the 1D (VP1) structural-protein-coding region. This corresponds to a highly conserved portion of the FMDV genome within the 2AB region. The oligonucleotide was obtained commercially from GibcoBRL and designated WDA (5'AAGGGCCCAGGGTTGGACTC3').

The reaction was performed in a 0.5ml micro-test tubes (Roche Molecular Biochemicals). The reaction mix contained 4µl 5 x Reverse Transcriptase buffer (Promega), 0.5µl diluted

RNase inhibitor (5U/µl), 3µl deoxyribonucleotide 5'-triphosphate mixture (2.5mM), 2µl of the WDA oligonucleotide (10pmol/µl), 0.5µl saturated dimethylsulfoxide solution, 0.5µl hexanucleotides (200pmol/µl) and 5µl distilled H_2O. A volume of 5µl viral RNA was added to the reaction mix. The tube was briefly centrifuged, incubated at 70°C for 3 minutes in order to destabilise any secondary structure present in the conformation of the RNA and on completion directly transferred onto ice for an additional 2 minutes. Subsequent to this, 3µl Avian Myeloblastosis Virus Reverse Transcriptase enzyme (10U/µl, Promega) was added and the reaction mix incubated at 42°C for 1 hour. On completion of the polymerisation reaction, enzyme activity was terminated by incubating the reaction mix at 80°C for 2 minutes. The cDNA was stored at -20°C.

2.2.8 Amplification of the structural-protein-coding region

The P1 coding region of the FMDV genome was amplified by employing the polymerase chain reaction (PCR)(Erlich, 1989). The first oligonucleotide used to initiate amplification was designed to anneal to an area within the 5' non-coding region of the genome. This oligonucleotide was obtained commercially from Gibco BRL and designated NCR1 (5'-TACCAAGCGACACTCGGGATCT-3'). The WDA primer used to initiate the preparation of the cDNA was used as the second oligonucleotide in the amplification of the P1 region.

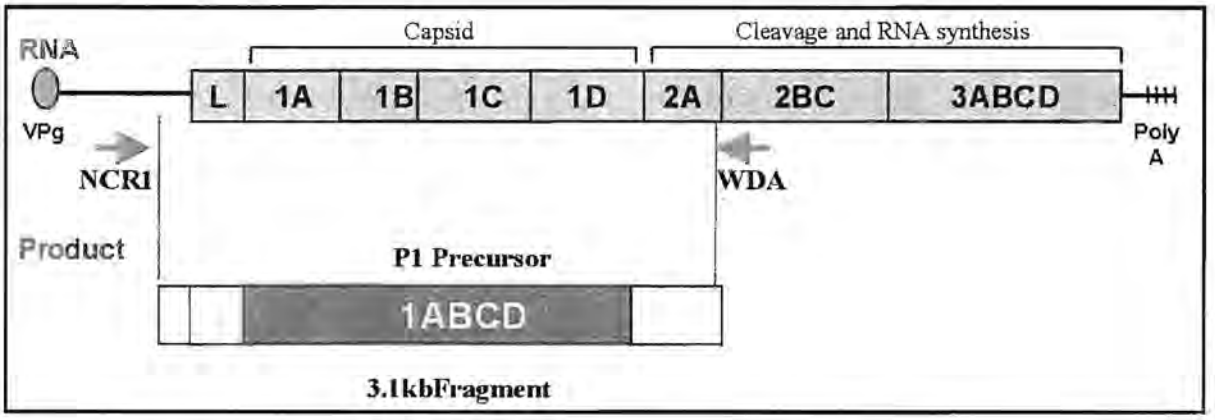

Figure 2.2.1 A graphical representation of the polymerase chain reaction strategy used to amplify the structural-protein-coding region of FMDV. The primers used in the reaction are represented by the red arrows (→).

The Expand™ Long Template PCR System (Roche Molecular Biochemicals) was used to amplify the region of interest. Two separate master mixtures were prepared as prescribed by the manufacturer. The first master mixture consisted of 5µl deoxyribonucleotide 5'-triphosphate mixture (2.5mM), 1.25µl of the NCR1 oligonucleotide (10pmol/µl), 1.25µl of the WDA oligonucleotide (10pmol/µl) and 4µl distilled H_2O. The second mixture contained 2.5µl 10 x Expand™ Long Template PCR Reaction Buffer (Roche Molecular Biochemicals), 0.5µl $MgCl_2$ (100mM), 0.4µl Expand™ Long Template enzyme mixture and 9.85µl dH_2O. The two master mixtures were added together, mixed thoroughly and overlaid with mineral oil. A volume of 2µl cDNA was added and the reaction mixture briefly centrifuged.

The PCR was performed using an OmniGene Hybaid Thermocycler. The sample was initially denatured at 95°C for 2 minutes followed by 10 cycles consisting of denaturation at 95°C for 30 seconds followed by a 30 seconds annealing period at 60°C and ending with polymerisation for 60 seconds at 68°C. This was followed by a further 20 cycles differing from the first 10 cycles only by the stepwise extension of the initial polymerisation period of 60 seconds by 20 seconds for each new repetition of a cycle.

After completion of the thermocycling the amplified fragments were separated in a 1.0% agarose gel by standard electrophoreses techniques (Sambrook, 1987). The gel was prepared by dissolving 1.0g of Multi-Purpose Agarose (Boehringer Mannhein) in 100ml 1 x TAE buffer. After allowing sufficient time for the gel-mix to cool to approximately 60°C, 2µl Ethidium bromide (10µg/ml) was added, the gel cast and left to set at room temperature. A volume of 5µl DNA loading buffer (40% sucrose solution coloured with bromo-phenol-blue) was added to the PCR reaction mixture and 50µl of loaded onto the set gel. A commercial DNA marker (*HindIII* restricted phage φX17 genomic DNA) was loaded onto the gel as a reference. The gel was run at 50V for approximately 45 minutes after which the fragment was visualised by exposing the gel to short wavelength ultra-violet light using a bench top transeluminator. Recovery of the appropriate amplified fragment from the agarose gel was done as described in Section 2.2.9.

2.2.9 Agarose gel purification of the P1 amplicon

The Amplified P1 fragment was recovered from the agarose gel according to the protocol set forward by the manufacturers of the Nucleospin Extract 2 in 1 DNA extraction kit (MacHerey-Nagel). The DNA fragment (≈3.1kbp) was excised from the gel with the minimum excessive agarose and transferred to a clean 2ml micro-test tube. The fragment was weighed and 300µl NT1 buffer (supplied in the kit) added for each 100mg gel excised. The mixture was incubated at 50°C for approximately 10 minutes and periodically inverted to ensure proper degradation of the agarose gel. The melted sample was loaded into the NucleoSpin column containing a silica filter (supplied as one unit in the kit). The column was inserted into a clean 2ml micro-test tube and centrifuged for 60 seconds at maximum revolutions using a Hettich bench top centrifuge. The resulting filtrate was discarded and 750µl NT3 buffer (supplied in the kit) added onto the filter. The column was again inserted into a 2ml micro-test tube and centrifuged for 60 seconds at maximum revolutions. The filtrate was discarded and the NucleoSpin column placed in a 1.5ml micro-test tube where after the DNA was eluted from the filter by adding 50µl 1M Tris/HCl buffer (pH 8.8 @ 25°C) and centrifuging the column for 60 seconds at 16000xg. The filtrate containing the purified DNA was stored at -20°C.

2.2.10 Ligation of the purified P1 fragment into a suitable plasmid vector

The pGEM-T-Easy vector system (Promega) was chosen as a suitable vector for cloning of the purified P1 PCR fragment (Figure A.1.1). This system has been optimised using a 1:1 molar ratio of insert to vector DNA. The appropriate amount of PCR product included in the ligation reaction was calculated by applying the following formula:

$$\textbf{Amount of insert (ng)} = \frac{\textbf{Amount of vector (ng) x Size of the insert (kbp)}}{\textbf{Size of vector (kbp)}} \textbf{ x insert:vector molar ratio}$$

Table 2.2.2 summarises all ligation reactions set up in during this experiment. Reactions were performed in 0.5ml micro-test tubes at a final volume of 10µl and incubated at 4°C for approximately 48 hours after which it was directly used in transformation reactions.

2.2.11 Preparation of competent ***Escherichia coli*** *cultures*

Competent cells were prepared using a method described by Tang *et al.* (1994). A pure culture of the *Escherichia coli* strain JM109 was inoculated into 100ml LB liquid medium and shaken at 37°C for approximately two hours. The optical density at 600nm was monitored every 20 minutes to examine the growth of the culture. After reaching a cell density of 5 x 10^7cells/ml ($OD_{600} \approx 0.3$), the cells were harvested by brief centrifugation and resuspended in 10ml of an ice cold solution consisting of 80mM calcium chloride and 50mM magnesium chloride. The resuspended cells were placed on ice for 10 minutes. This treatment was repeated twice before the cell pellet was finally resuspended in 5ml 0.1M calcium chloride. An equal volume of 50% Glycerol was added and 200ul aliquots of the cells frozen in liquid nitrogen. The aliquots were stored at –70°C.

Table 2.2.2. A summary of all reactions set up during the ligation of the purified P1 PCR fragment into the pGEM-T-Easy plasmid vector.

	Standard Reaction	*Positive Control*	*Background control*
10 x T4 Ligase Buffer	1µl	1µl	1µl
pGEM-T-Easy vector (50ng/µl)	1µl	1µl	1µl
Control Insert DNA (4ng/µl)	Not applicable	2µl	Not applicable
PCR Product (50ng/µl)	1µl	Not applicable	Not applicable
T4 Ligase (3U/µl)	1µl	1µl	1µl
dH_2O	6µl	5µl	7µl

2.2.12 Transformation of the ligated pGEM-T-Easy vector

A 200µl aliquot of competent *E. coli* JM109 cells, prepared as described in Section 2.2.11, was slowly thawed on ice for each of the respective ligation reactions (Table 2.2.2). The total volume of the ligation reaction (10µl) was added to each aliquot. Each sample was left on ice for a further 30 minutes and subsequently incubated at 37°C for 5 minutes. A volume of 800µl SOC culture medium was added to each sample after incubation, the contents of each tube gently mixed by inversion and incubated at 37°C for 30 minutes. A representative volume (100µl) of each sample was transferred to SOB-agar culture plates, containing ampicillin (100ng/ml), X-Gal solution (2%) and IPTG (10mM). The plates were incubated at 37°C for 14 hours after which all visible colonies with an estimated diameter

of at least 0.5mm were counted. The transformation efficiency was calculated using the data obtained by the transformation of the positive control ligation reaction by applying the following formula:

$$\textbf{Transformation efficiency (\# cfu/µg of DNA)} = \frac{\textbf{\# of transformants}}{\textbf{Quantity DNA transformed (µg)}} \times \frac{\textbf{Final Reaction volume (µl)}}{\textbf{Volume plated (µl)}}$$

Suspected recombinant colonies, indicated by the absence of blue pigment, were transferred to fresh SOB-agar culture plates, incubated at 37°C for 18 hours and stored at 4°C.

2.2.13 Isolation of recombinant plasmids

Cultures suspected of containing recombinant plasmids were inoculated into 200ml flasks containing 50ml TB liquid medium containing 50ng/ml ampicillin. The flasks were incubated overnight at 37°C while continuously being shaken at 26 revolutions per minute The cells were harvested from 3ml liquid medium by successively transferring 1ml of the cell suspension to a 1.5ml micro-test tube, centrifuging the tube at 8000 x g using a Hettich bench top centrifuge and subsequently discarding the supernatant. The resulting cell pellet was resuspended in 250µl SET buffer (pH 8.0 @ 25°C). To this was added 500µl of a freshly prepared solution consisting of 1% sodium dydocyl sulphate and 0.1N sodium hydroxide. The contents of the tube was mixed by inversion and incubated at 65°C. After a 30 minute incubation period, 375µl of a 3M sodium acetate solution (pH 5.2 @ 25°C) was added and the sample transferred onto ice for approximately 15 minutes. Cell debris was removed by centrifuging the samples at 16000xg for 15 minutes. The resulting supernatant was transferred to a clean 1.5ml micro-test tube and a volume of 300µl polyethylene glycol/sodium chloride solution added to it. The sample was left on ice for a further 2 hours and subsequently centrifuged at 16000xg for 15 minutes using a bench top centrifuge. The resulting supernatant was discarded, the pellet washed with 500µl 70% ethanol and resuspended in 400µl 1 x TE buffer (pH 7.6 @ 25°C). A volume of 400µl 5M ammonium acetate was added to the resuspended pellet followed by the addition of 600µl isopropanol. The sample was placed on ice for 30 minutes and the precipitated DNA concentrated by centrifuging the sample 16000xg. On completion of centrifugation the supernatant was carefully removed. Any remaining ethanol was removed by evaporation

under vacuum. The pellet was resuspended in 50µl 1 x TE buffer (pH 7.6 @ 25°C) and stored at -20°C.

2.2.14 Sequence determination of the cloned structural-protein-coding region

The nucleotide sequence of the cloned fragment was determined independently for both the positive and negative strand of the DNA using oligonucleotide primers designed to generate partially overlapping DNA sequences of the entire P1 genomic region. A list of oligonucleotides used in the sequence determination, as well as a graphical representation of their respective positions on the genome are summarised included in the appendix (Table A.1.1 and Figure A1.2). These primers were designed based on nucleotide sequence data available on Genbank as well as the data generated in this study, as sequence data became available.

Two independent methods were employed to determine the nucleotide sequence of the structural-protein-coding region of the FMDV isolate in question. The first method was previously described by Zimmerman and Kaesberg (1978) and later modified by Xie *et al.* (1987). This method of sequence determination is based on the chain termination due to the incorporation of 2,3-dideoxy-D-ribose analogues of the deoxyribonucleoside 5'-triphosphates during template directed polymerisation of deoxynucleotides (Sanger *et al.* 1977). The newly synthesised DNA is labelled by the incorporation of ^{32}P-labelled deoxyadenosine 5'-triphodphate. Nucleotide sequencing by chain termination involves the following methodology.

Purified plasmid DNA (5µl) was denatured by adding 5µl of a 0.4M sodium hydroxide solution containing EDTA at a concentration of 0.04M to it and incubating the mixture at 37°C for 30 minutes. The denatured plasmid was then precipitated by adding 1µl of a 3M sodium acetate solution (pH 5.2 @ 25°C) and 3 volumes of absolute ethanol to the mixture and incubating it overnight at –20°C. The precipitated DNA was concentrated by centrifuging the mixture at 16000xg for 10 minutes using a bench top centrifuge. The supernatant was removed with a P1000 micropipet and the pellet dried under vacuum to remove any excess fluids. The DNA was resuspended in 12µl dH_2O.

An annealing reaction consisting of 4µl of the resuspended plasmid DNA, 2µl 5 x Sequenase Reaction Buffer and 4µl of a oligonucleotide primer (10pmol/µl), was set up and incubated at 65°C for a period of 2 minutes before being slowly cooled to 35°C. The labelling reaction mix, containing 1µl DTT, 2µl diluted labelling solution, 0.5µl ^{32}P α-ATP (800Ci/mmol) and 2µl Sequenase enzyme, was prepared and 5µl thereof added to the annealing reaction mixture. The contents of the micro-test tube was mixed thoroughly and incubated at room temperature for 5 minutes.

The polymerisation reactions were performed in a 96 well polyethene plate. An 2.5µl aliquot of each dideoxynucleoside-triphosphate terminator (ddNTP) was transferred to separate wells and 3.5µl of the annealing reaction mixture added to each well. The contents of the wells were mixed by repeatedly dispensing the entire volume using P10 micropipet. The mixture was incubated at 37°C for 5 minute and polymerisation reaction terminated by adding 4µl formamide dye mix to each well. The constructed DNA was denatured by incubating it at 75°C and immediately processed.

Two micro-litres of the contents in each well was loaded onto a 12% polyacrylamided gel in order A, C, G and T according to the ddNTP present the respective wells. The gel was run at 60W for approximately 60 minutes, using 0.5 x TBE buffer (pH 8.3 @ 25°C) as conductive fluid, to effectively separate the fragments constructed during the polymerisation reaction. After completion of the run, the gel was carefully submerged in an acetic acid/methanol/water solution (1:1:8 v/v) for 10 minutes. It was then transferred onto 3MM filter paper (Whattman), dried at 80°C for 45 minutes using a gel drier and placed in a development cassette in direct contact with a X-ray film. The film was exposed to the gel for approximately 2 days at room temperature and subsequently developed using Ilford Phenisol developer and fixer at the prescribed concentrations. The nucleotide sequence was read manually from the developed X-ray film and transferred onto computer in order to construct a digital record thereof. The digital version was used computer assisted analysis of the nucleotide sequence.

Alternatively, the ABI PRISM™ BigDye™ Terminator Cycle Sequencing ready reaction kit (Perkin-Elmer Applied Biosystems) was used to determine the nucleotide sequence of the cloned fragment. The kit offers a non-radio active alternative to the method discussed

above by incorporating florescent labelled synthetic terminators. Sequencing reactions were set up as prescribed by the manufacturer and consisted of 2µl Terminator Ready Reaction Mix, between 250 to 300ng double-stranded DNA template and 20pmols of the appropriate oligonucleotide. The final volume of each reaction was adjusted to 20µl with the appropriate volume of deionised water. Thermocycling was performed on the GeneAmp 2400 PCR System (Perkin-Elmer Applied Biosystems) as specified by the manufacturer (ABI PRISM™ BigDye™ Terminat or Cycle Sequencing ready reaction kit protocol, 1997).

After completion of the thermocycling, 20µl 65% ethanol was added to the reaction mixture before incubation thereof at room temperature for a period of 1 hour. The tube was centrifuge for 30 minutes at 16000xg and the supernatant removed from it using a P20 micropipet. The resulting pellet was washed with 70% ethanol and dried under vacuum before being gently resuspended in 3µl sequencing loading buffer. The sample was submitted for processing to the Department of Genetics at the University of Pretoria. The completed sequence was supplied in both electropherogram format and digital text format. The data were verified manually and the necessary corrections made before being analysed as discussed in Section 2.2.15

2.2.15 Computer assisted analysis of the obtained nucleotide sequence

The nucleotide sequence of the entire P1 polyprotein-coding region of FMDV KNP/10/90/3 was constructed by computer assisted alignment of the partially overlapping DNA sequences generated as described in Section 2.2.14 (DAPSA ver.4.31: Harley, 1998). The reconstructed full-length nucleotide sequence was aligned to the corresponding region of several other FMDV isolates (Table 2.2.1) using the DAPSA program. The alignment was used to determine the phylogenetic relationship of these isolates. All phylogenetic analyses were performed using the GCG program (Wisconsin Package ver.10). More detailed analysis of each of the specific structural proteins coded for within the P1 genomic region was carried out using either the Antheprot (Delaege, 2000) or DNAMAN programs (Lynnon Biosoft) depending on the task performed.

2.3 Results

2.3.1 Amplification and cloning of the structural protein-coding region

Initial attempts to amplify the P1 region of the FMD virus KNP/10/90/3 yielded five distinct DNA products ranging in size from approximately 620 to 3100bp. Subsequent attempts to optimise the reaction conditions and thereby reduce non-specific amplification were unsuccessful (results not shown). However, the size of the largest fragment (3100bp) corresponded to the size of the product expected to contain the P1 region as predicted by computer assisted analysis of several published FMD virus genomes. In order to verify whether this product indeed contained the P1 coding region, it was purified from the agarose gel, cloned into the pGEM-T-Easy vector system (Promega) by means of T/A cloning and sequenced. Compotent *E. coli* JM109 cells were transformed as described in Section 2.2.12. The efficiency of the transformation was calculated to be 4 x 10^6 colony forming units per µg DNA transformed. Since the pGEM-T-Easy vector system contains a selection protocol based on the insertional inactivation of the *LacZ* operon, the absence of the distinctive blue colour generated by the catalytic conversion of XGal, was taken to be indicative of the successful incorporation of the desired fragment. The ratio of white to blue colonies was calculated to be 5:1.

Five white colonies were randomly selected and cultured for the purpose of plasmid extraction. The respective isolated plasmids were independently sequenced inwards from both cloning sites using the NCR1 and WDA primers. Sequences obtained for all five clones using the NCR1 oligonucleotide corresponded to the unpublished leader protease coding sequence of KNP/10/90/3 (van Rensburg, unpublished results). However, only the sequence of one clone generated using the WDA oligonucleotide, corresponded to the published VP1 genomic sequence of KNP/10/90/3 (GenBank accession number: AF0235). Results indicated that the remaining four clones contained a truncated version of the P1 region and lacked approximately 70 nucleotides at the 3' end of the VP1 coding region. The presence of the complete coding regions of both the leader protease and VP1 structural protein suggested that the entire structural-protein-coding region was contained within at least one of the clones. This clone, designate pGEM.SAT3P1, was used in determining the nucleotide sequence of structural-protein-coding region of KNP/10/90/3.

2.3.2 Nucleotide sequence of FMD virus isolate KNP/10/90/3

The nucleotide sequence of the cloned fragment was determined independently for both the positive and negative strand of the DNA using oligonucleotide primers designed to generate partially overlapping DNA sequences of the entire P1 genomic region. The reconstructed sequence translated into a single open reading frame of 2784 base pairs. The derived amino acid sequence contained recognisable sequences of the leader protease, the VP4, VP2, VP3, VP1 structural proteins and the 2A non-structural protein when compared to published sequence data of viruses from the O, A and C serotypes. The nucleotide as well as the deduced amino acid sequences of the structural-proteins-coding regions VP0, VP3 and VP1 is represented in Figures 2.3.1. *a* to *c*. The nucleotide sequence of KNP/10/90/3 was submitted to the NCBI nucleotide sequence database and was assigned the accession number AF286347.

Nucleotide frequency were calculated to be 19%, 30%, 26% and 25% for the bases T/U, C, A and G respectively (DNAMAN, Lynnon Biosoft). A striking characteristic of the coding region is the apparent bias in favour of G/C (55%) over A/T (45%). This type of bias has previously been shown to be common for at least four of the seven FMDV serotypes and are not restricted to the structural-protein-coding region but rather persist throughout the genome (Brown *et al.*, 1989). Similar biases are not commonly observed in other members of the picornavirus family (Racaniello and Baltimore, 1981) and seem to be to be a distinguishing characteristic of FMDV within the taxon. The bias towards G/C is further reflected in the codon usage during translation and is particularly evident in the third position of the codons were a 1.62:1 G/C:A/T ratio is well above the overall ratio of 1.2:1.

2.3.3 Computer assisted analysis of the structural proteins.

Analyses of the nucleotide and amino acid sequence of the P1 region were performed on a homologous region starting at the N-terminal amino acid of the VP4 structural protein and ending in the C-terminal amino acid of the VP1 structural protein. Sequence identity across all serotypes analysed are given in Table 2.3.1.

Nucleotide homology over the entire structural-protein-coding region when compared to representative isolates of the SAT 1 and 2 were calculated to be approximately 60% while

an expected higher homology was observed between the SAT 3 isolates (73.6 %). When the nucleotide sequence of the KNP/10/90/3 isolate was compared to the corresponding region of three European serotypes, the homology was found to be markedly lower (≈ 50%). This is in contrast with a previously study where the P1 region of another SAT 3 isolate, Bec/1/65/3, was compared to the same European isolates which found the sequence identity amongst these isolates to be in the region of 60% (Brown *et al.*, 1989). It is however important to note that both studies found the amino acid homology of the isolates to be approximately 60%. The intratypic and interpypic variation calculated for the individual structural proteins VP4, VP3, VP2 and VP1 are summurised in Table 2.3.1.

Table 2.3.1 Percentage homology between KNP/10/90/3 and 6 other FMDV serotype at nucleotide and amino acid levels.

	A_{10}	O_1 Kaufbeuren	C_1 Germany	Bot/1/68/1	Zim/7/83/2	Bec/1/65/3
VP4	75.2 (90.5)	76.0 (94.1)	74.8 (90.4)	81.0 (95.2)	81.9 (95.2)	81.1 (95.2)
VP2	62.9 (63.1)	62.7 (66.3)	61.7 (62.3)	63.7 (72.3)	69.4 (71.4)	79.0 (88.4)
VP3	57.4 (52.2)	48.6 (50.0)	55.5 (49.3)	57.1 (57.5)	68.3 (54.4)	75.6 (70.7)
VP1	38.4 (44.7)	42.4 (45.4)	47.3 (42.1)	49.1 (48.8)	48.0 (55.9)	75.6 (81.7)
Total[a]	52.1 (58.3)	46.4 (60.5)	51.8 (56)	61.3 (67.7)	61.6 (63.2)	78.6 (82.6)

Nucleotide and amino acid (in brackets) homology determined using DAPSA (Harley, 1998).
[a] Total homology was calculated for the polypeptide region spanning from the N-terminus of VP4 structural protein and C-terminus of VP1.

The intratypic and interpypic variation calculated for the individual structural proteins VP4, VP3, VP2 and VP1, are summurised in Table 2.3.1. The highest level of homology was found within the VP4 protein. In contrast to VP4, the VP1 structural protein showed remarkably low levels of homlogy.

The cleavage of the P1 precursor at the VP2/VP3, VP3/VP1 and P1/P2 junction sites are carried out by the 3C viral protease in a sequence dependent manner (Baxt *et al.*, 1989; Belshaml *et al.*, 1993). Using data generated in this study, comparative investigations of these sites revealed notable differences in at least one of the cleavage sites of the SAT isolates compared to the equivalent site of the European serotypes (Figure 2.3.4). The junction between VP0 and VP3 was marked by the substitution of the glutamate residue (E) by a glutamine residue (Q) at position 303 of the polyprotein.

2.3.4 Phylogenetic relationship of FMDV based on the P1 protein

Phylogenetic analyses were carried out for a data set containing 14 orthologous taxonomic units using neighbor-joining and maximum likelihood methods. Two representative sequences of each of the 7 FMDV serotypes were included in the analysis. Similar results were obtained in both cases and only results of the neighbor-joining method are depicted in Figure 2.3.2. Isolates grouped strictly according to serotype. Two distinct lineages were obtained, the first representing the European types (Green) and the second representing the SAT types of FMD viruses (Blue). Variation for the entire P1 region of the SAT 3 serotype was calculated to be 16.4% (N=2) and was found to be similar to the intratypic variation present within the same region of the SAT 2 serotype (16%, N=3, van Rensburg *et al.*, 1999). It is however important to note that only two isolates were used to determine variation within the P1 region of the SAT 3 serotype and that that the degree of intratypic variation for this serotype may change with the addition of more representatives. Results furthermore indicated that the variation for the P1 region of the SAT type viruses was higher compared to that of serotypes A (11%), O (7%) and C (9%) (van Rensburg *et al.*, 1999). This corresponded to results of previous studies that showed intratypic variation for the VP1 protein of SAT types to be as much as two to three times higher than variation observed within the corresponding protein of serotypes A, O and C (Bastos *et al.*, 1998).

2.5 Discussion

Nucleotide and amino acid sequence analyses showed the genetic makeup and protein organisation of KNP/10/90/3 to be typical of FMDV. As would be expected the isolate shared a high level of sequence homology with Bec/1/65/3 and was more closely related to the other SAT serotypes than to European FMD viruses.

The VP4 protein-coding region showed the highest level of homology among SAT3 isolates. This protein is concealed within the capsid and is therefore not exposed to antibody surveillance. Furthermore, it exists in close association with the viral RNA. It has been hypothesized that the autocatalytic processing of the VP0 precursor involves a serine residue situated in close proximity of the cleavage site within VP2 (Arnold *et al.*, 1987). However, the lack of an adjacent histidine residue required for serine protease activity suggested that an alternative proton donor is involved. It has been proposed that a base

within the viral RNA might fulfil this role (Rossnam *et al.*, 1985). A high degree of structural conservation is needed to consistently maintain this association. The applicability of the hypothesis towards the proteolytic processing of KNP/10/90/3 is however questionable, since no serine residues within VP2 can be found in close proximity to the VP4/VP2 cleavage site.

In contrast to VP4, the VP1 structural protein showed remarkably high levels of intratypic variation. The VP1 protein contains the major immunodominant region of the virion and is therefore under constant selective pressure. This contributes to the high levels of variation within this protein. As would be expected however, the homology for the VP1 protein within the SAT 3 serotype (75.6%) was noticeably higher than observed across serotypes (≈40% for European serotypes and ≈48% for SAT1 and 2). The VP2 and VP3 structural proteins have also been shown to contain neutralising epitopes, but to a lesser extent than VP1 and therefore displayed intermediate levels of homology.

In addition to the general composition of the P1 precursor, the biochemical characteristics of several distinct regions within the precursor are important to note. During an early post-translational processing event, a myristic acid is added to the N-terminus of the P1 precursor. This modification has been observed for viruses in all genera of the picronaviridae, as well as several other virus families (Chow *et al.*, 1987). The covalent linkage of the n-tetradecanoic acid is determined by the presence of the consensus moiety Gly-X-X-X-Ser/Thr/Ala (where X represents any other amino acids) at the N terminus of VP4 (Chow et al, 1987). The N-terminal amino acid sequences of the VP4 protein from the KNP/10/90/3 isolate conforms to this consensus and therefore could potentially be myristylated.

Virus entry during FMD infection occurs due to receptor-mediated endocytosis. A crucial part of internalisation of the virion is the non-covalent interaction between the vitronectin host cell surface receptors and the Arg-Gly-Asp (RGD) protein ligand (Fox *et al.*, 1989; Evens & Almond, 1998). The RDG sequence is located between the G and H β strands of VP1 and is conserved across all seven FMD serotypes. As would be expected this moiety was present within VP1 of KNP/10/90/3. In addition to containing the RGD moiety, the G-H loop also contains the major immunodominant epitope. The hypervariable region of

Bec/1/65/3 (amino acisd 142 to 165) corresponding to the immunodominant epitope of O_1 Kaufbeuren (amino acids 141 to 160) is characterised by the presence of five additional amino acids (Brown et al., 1989). The exact same insertions were found within the VP1 protein of KNP/10/90/3. Similar differences within this region were observed for both serotypes A and C involving 7 and 9 amino acids respectively when compared to KNP/10/90/3. The implications of the high homology of the additional amino acids observed within the G-H loop of both SAT 3 isolates however remain unexplored.

As mentioned earlier in the chapter, efforts are underway to circumvent problems associated with the selection of suitable strains of FMDV for vaccine production by constructing chimeric viruses (refer to paragraph 2.1.1). However, a number of crucial factors that may influence the success of such a strategy in the southern African context should be considered before the implementation of such a strategy. The most significant factor is the role of variation of the cleavage sites within the P1 precursor among different viruses.

Despite the high degree of variation found within the VP1 structural protein of different FMD virus isolates, the VP1/2A cleavage site was the least variable of all the junction sites within the P1 precursor. The two amino acids directly at the cleavage site (positions 739 and 740) were conserved across all isolates included in the analysis. The VP3/VP1 junction site (positions 523 and 524) of all the SAT types corresponded with that of serotypes A and C. In all SAT type viruses included in the study, the junction between VP0 and VP3 was marked by the substitution of the glutamate residue (E) by a glutamine residue (Q) at position 303 of the polyprotein. Structurally and spatially these two amino acids are similar, but differ significantly in the net charge of their respective side chains. Because of the carboxylate present in the side-chain of glutamic acid, this amino acid is negatively charged at neutral pH. However, the side-chain carboxyl group frequently bind NH_2 to form the side-chain amide group, yielding the analogue amino acid, glutamine (Campbell, 1995). In contrast with glutamate, glutamine has a polar side-chain that is electrically neutral at pH 7.

It is important to note that regions flanking the reactive sites could possibly play a role in the recognition of the cleavage sites by the protease and that variation within these regions

could further compromise the effective processing of the P1 precursor. Comparative analyses of the flanking regions indicated that these regions seem to be conserved within serotypes, but show a greater degree of variation across serotypes. The precise effect of the variation within the P1 precursor remain unknown, but in our opinion do not pose a potential problem in the construction of chimeras SAT type viruses using the cDNA backbone of the European isolates.

VP0 Structural protein precursor

Open Reading Frame: 1-909
Total amino acid number: 303, MW=33021Da

```
1          GGAGCGGGCCAGTCATCCCCTGCCACTGGATCCCAAAACCAATCTGGTAACACTGGCAGC
1           G  A  G  Q  S  S  P  A  T  G  S  Q  N  Q  S  G  N  T  G  S

61         ATAATTAACAACTACTACATGCAGCAGTACCAAAATTCAATGGACACTCAACTCGGCGAC
21          I  I  N  N  Y  Y  M  Q  Q  Y  Q  N  S  M  D  T  Q  L  G  D

121        AACGCCATAAGTGGAGGTTCCAATGAGGGTAGCACGGACACCACGTCCACGCATACCAAC
41          N  A  I  S  G  G  S  N  E  G  S  T  D  T  T  S  T  H  T  N

181        AACACCCAGAACAATGATTGGTTTTCCAAATTGGCACAATCTGGCATTTCAGGCCTTTTT
61          N  T  Q  N  N  D  W  F  S  K  L  A  Q  S  G  I  S  G  L  F

241        GGGGCCCTTCTGGGCAGACAAAAGGACAGAAGAAACAACCCTTCTGGAGGATCGCATTCTC
81          G  A  L  L  G  D  K  R  T  E  E  T  T  L  L  E  D  R  I  L

301        ACTACACGCCACAACACAACCACATTCACAACTCAGAGTTCTGTCGGCGTGACACACGGT
101         T  T  R  H  N  T  T  T  F  T  T  Q  S  S  V  G  V  T  H  G

361        TACGCATCAGCGGACCGTTTTCTACCTGGACCCAACACCAGTGGACTCGAGACACGCGTT
121         Y  A  S  A  D  R  F  L  P  G  P  N  T  S  G  L  E  T  R  V

421        GAACAAGCAGAGAGATTCTTCAAGGAGAAACTCTTRACTTGGACCGCGGCTCAAGAGTAC
141         E  Q  A  E  R  F  F  K  E  K  L  L  T  W  T  A  A  Q  E  Y

481        GCACACGTGCATCTGTTGGAGCTGCCGGTGGACCACAAAGGCATCTACGGTGCCATGCTG
161         A  H  V  H  L  L  E  L  P  V  D  H  K  G  I  Y  G  A  M  L

541        GACAGCCACGCATACGTGCGCAACGGCTGGGATGTGCAGGTTTCAGCAACCAGCACACAG
181         D  S  H  A  Y  V  R  N  G  W  D  V  Q  V  S  A  T  S  T  Q

601        TTCAACGGCGGCACCCTTCTTGTCGCCATGGTCCCCGAGTTGCACACTTTGGACAAGCGT
201         F  N  G  G  T  L  L  V  A  M  V  P  E  L  H  T  L  D  K  R

661        GACGTTTCACAACTCACGCTCTTTCCTCACCAGCACATCAACCCACGTACCAACACCACT
221         D  V  S  Q  L  T  L  F  P  H  Q  H  I  N  P  R  T  N  T  T

721        GCCCACATCGTGGTACCCTACGTGGGTGTCAACAGACACGACCAGGTAAAAATGCACAAG
241         A  H  I  V  V  P  Y  V  G  V  N  R  H  D  Q  V  K  M  H  K

781        GCATGGACACTTGTGGTGGCAGTGCTCGCACCGCTCACCACCTCCAACATGGGGCAAGAC
261         A  W  T  L  V  V  A  V  L  A  P  L  T  T  S  N  M  G  Q  D

841        AACGTTAGGTGTACGCAAACATCGCCACCAACCAACGTGTATGTTGCTGGAGAGAAGCCG
281         N  V  R  C  T  Q  T  S  P  P  T  N  V  Y  V  A  G  E  K  P

901        ACAAAACAA
301         T  K  Q
```

Figure 2.3.1.a The nucleotide and deduced amino acid sequence of the VP0 precursor of FMD virus isolate KNP/10/90/3. (U residues present in the viral RNA are shown as T.) The VP4/VP2 cleavage site is indicated in red bold italics. *The molecular weight of the precursor was m predicted by computer modeling and not determined empirically.*

VP3 Structural-protein-coding region

Open Reading Frame: 910-1632
Total amino acid number: 221, MW=24422

```
910         GGCATCGTCCCCGTGGCTTGCCACGACGGCTACGGCGGTTTCCAGAACACA
304          G  I  V  P  V  A  C  H  D  G  Y  G  G  F  Q  N  T

961         GACCCCAAGACCGCGGACCCTATCTACGGCCTAGTGTCCAACCCGCCACGCACAGCGTTT
321          D  P  K  T  A  D  P  I  Y  G  L  V  S  N  P  P  R  T  A  F

1021        CCCGGCAGGTTCACCAACCTGTTGGGACGTCGCTGCAGCGTGCCCAACTTTCTTTGGACT
341          P  G  R  F  T  N  L  L  G  R  R  C  S  V  P  N  F  L  W  T

1081        TCGACGGGACACCGTACGTTAAGACCAAGCACAACAGTGGCAGCAAGATACTCACACACA
361          S  T  G  H  R  T  L  R  P  S  T  T  V  A  A  R  Y  S  H  T

1141        TTGACTTGGCATTTGGGCACAAGAAGTTTTAAGAACACTTACGTGGCAGGGCTCGCGCAG
381          L  T  W  H  L  G  T  R  S  F  K  N  T  Y  V  A  G  L  A  Q

1201        TACTACGCCCAATACAGTGGTTCTCTGAATCTGCACTTCATGTACACCGGTCCCACGCAG
401          Y  Y  A  Q  Y  S  G  S  L  N  L  H  F  M  Y  T  G  P  T  Q

1261        TCAAAAGCACGCTTCATGGTTGCGTACATACCACCAGGGACCGATCTGTTTCGACACTCC
421          S  K  A  R  F  M  V  A  Y  I  P  P  G  T  D  L  F  R  H  S

1321        CGAGGCAGCAGCACACTGCTCCACTCGGAGTGGGACACCGGACTGAACTCCAAGTTCACG
441          R  G  S  S  T  L  L  H  S  E  W  D  T  G  L  N  S  K  F  T

1381        TTCACGGGCCAACCGACATGGAGAGCAGACTTTGCCTATACATATTGTGATGAGCCTGAA
461          F  T  G  Q  P  T  W  R  A  D  F  A  Y  T  Y  C  D  E  P  E

1441        CAGGCCTCTGCACAAGGATGGGTGACAGTGTATCAAATCACTGACACACATGACCCTGAT
481          Q  A  S  A  Q  G  W  V  T  V  Y  Q  I  T  D  T  H  D  P  D

1501        TCAGCAGTGCTTGTCTCGGTCAGCGCTGGCGCTGACCTTGAGTTTCGGCTACCAATCAAC
501          S  A  V  L  V  S  V  S  A  G  A  D  L  E  F  R  L  P  I  N

1561        CCTGTGGCACAA
521          P  V  A  Q
```

Figure 2.3.1.b The nucleotide and deduced amino acid sequence of the VP3 structural-protein-coding region of FMD virus isolate KNP/10/90/3. (U residues present in the viral RNA are shown as T.) The molecular weight of the precursor was predicted by computer modeling and not determined empirically.

VP1 Structural-protein-coding region

Open Reading Frame: 1572-2217
Total amino acid number: 215, MW=23681

```
1572      ACAACCAGCGCAGGTGAGGGTGGTGACGTCGTGACGACCGATGTCACG
525        T  T  S  A  G  E  G  G  D  V  V  T  T  D  V  T

1621      ACACACGGCGGAGCTGTGGACACACCGCGCCGGCAGCACACCAACGTGGAGTTCCTGCTG
541        T  H  G  G  A  V  D  T  P  R  R  Q  H  T  N  V  E  F  L  L

1681      GACAGGTTCACACACATTGGTACGATCACCGGCTCAAAAACCATTGACCTCATGGACACG
561        D  R  F  T  H  I  G  T  I  T  G  S  K  T  I  D  L  M  D  T

1741      AAGGAACACACGCTAGTGGGCGCAATCTTGCGTTCTGTCACGTACTACTTTTGTGACTTG
581        K  E  H  T  L  V  G  A  I  L  R  S  V  T  Y  Y  F  C  D  L

1801      GAAGTTGCCGTCTTAGGCACTGGGTGGGTCGGGTGGGTTCCCAACGGCTGCCCACACACC
601        E  V  A  V  L  G  T  G  W  V  G  W  V  P  N  G  C  P  H  T

1861      ACCCGCGTGGAGGACAATCCAGTTGTTCACGCCAAGGGAGGTGTTACCCGTTTTGCTTTG
621        T  R  V  E  D  N  P  V  V  H  A  K  G  G  V  T  R  F  A  L

1921      CCATACACGGCACCACACGGAGTTCTTGCCACAGTTTACAATGGAAACTGCAAATACTCC
641        P  Y  T  A  P  H  G  V  L  A  T  V  Y  N  G  N  C  K  Y  S

1981      AAGACCCAACACGTTGTACCGCGCCGCGGCGATCTCGCCGTGCTGGCGCAGCGCGTTGAG
661        K  T  Q  H  V  V  P  R  R  G  D  L  A  V  L  A  Q  R  V  E

2041      AATGAAACCACAAGATGCAGACCCACAACATTCAACTTCGGGAGACTGTTGTGTGACACC
681        N  E  T  T  R  C  R  P  T  T  F  N  F  G  R  L  L  C  D  T

2101      GGTGACGTCTACTACCGGATGAAGAGGGCTGAGCTGTACTGTCCACNCGCCTTACGGGTT
701        G  D  V  Y  Y  R  M  K  R  A  E  L  Y  C  P  X  A  L  R  V

2161      AGGTACGCGCACACCACTGATCGGTACAAGACCAAGCTGGTAGCACCTGACAAACAA
721        R  Y  A  H  T  T  D  R  Y  K  T  K  L  V  A  P  D  K  Q
```

Figure 2.3.1.c The nucleotide and deduced amino acid sequence of the VP1structural-protein-coding region of FMD virus isolate KNP/10/90/3. (U residues present in the viral RNA are shown as T.) The molecular weight of the precursor was predicted by computer modeling and not determined empirically.

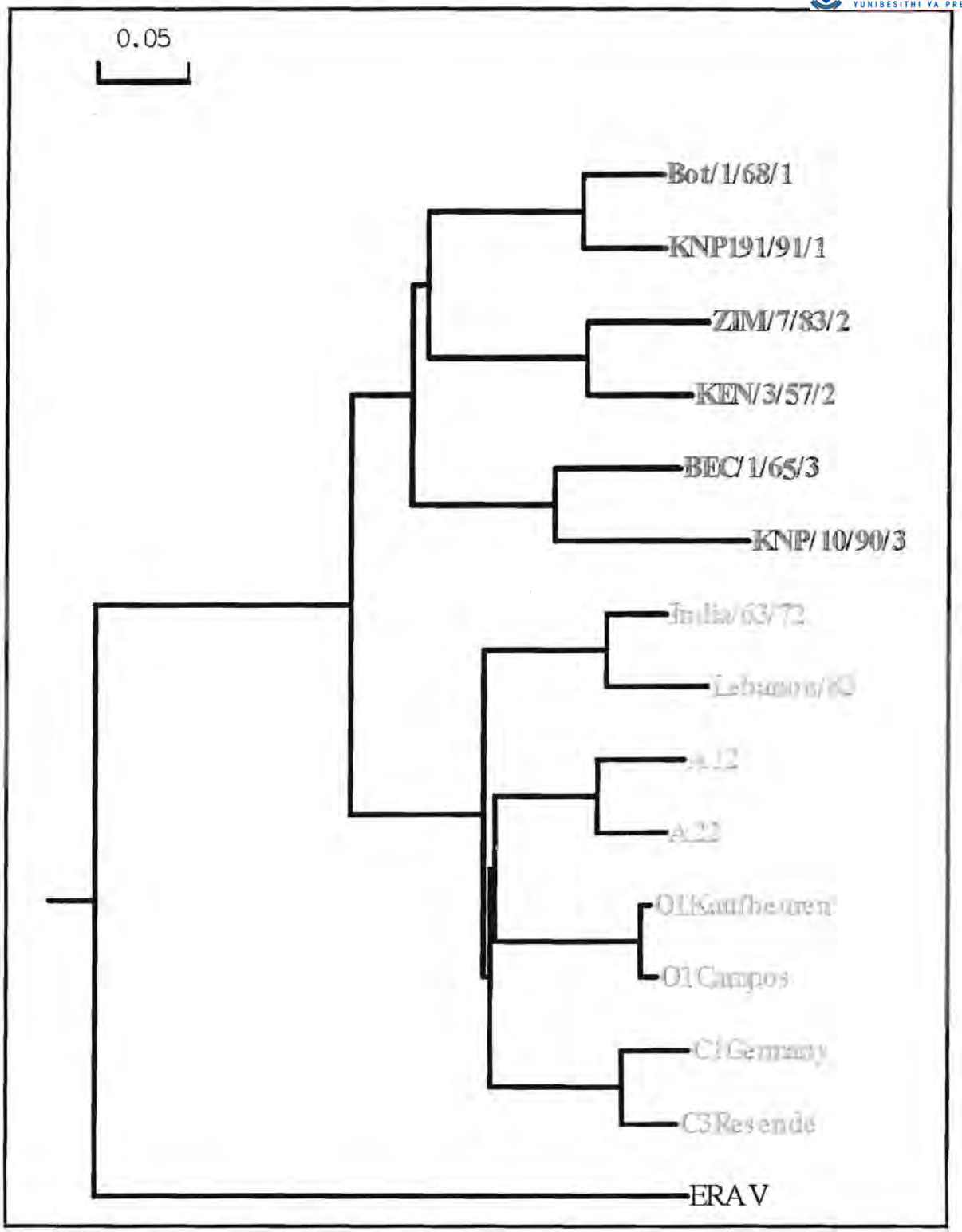

Figure 2.3.2 A Neighbor-joining tree depicting the structural-protein gene relationships between fourteen FMD virus isolates representative of the seven serotypes. A trial number of 1000 was applied. The analysis was performed using the deduced amino acid sequences of the structural-protein-coding region of the genome. Isolates originating from Europe, Asia and South America are indicated in green, while isolates originating from southern Africa are indicated in blue. Equine Rhinitis virus A was used as an outgroup.

	VP0/VP3 Junction			VP3/VP1 Junction			VP1/2A Junction		
	294		313	514		534	730		752
A12	YVHVAGELPSK	**EG**	IFPVACSDG	ELRLPIDPRS	**QT**	TATGESADPV	KQKIIAPG	***KQ***	LLNFDLLKLAGDV
A22	H..........	**EG**	.V.......	....	**QT**	.S........	.......A	**KQ**	....
O1 Kaufbeuren	N......F...	**EG**	....	.....V.A.A	**ET**	.SA.......	....V..V	**KQ**	T............
O1 Campos	N......F...	**EG**	....	.....V.A.A	**ET**	.SA.......	....V..V	**KQ**	T............
C1 Germany	N.Y........	**EG**	..S......	.....V.A.Q	**QT**	.T........	..PL...A	**KQ**	....
C3 Resende	N..........	**EG**	.....-...	.....V.A.L	**QT**	.T........	..RL...A	**KQ**	.S...........
India/68/73	H..........	**EG**	.V....A..	.F...V.A.R	**ET**	.TT....C..	..E....E	**KQ**	V............
Lebanon/83	....	**EG**	.V....A..	.F...V.A.Q	**QT**	.TA.......	..E....E	**KQ**	V............
Bot 1/68/1	N.Y....K.V.	**QG**	.L...V...	.F.M..S.SR	**QT**	.SA..G....	.TTLVK.A	**KQ**	.S...........
KNP 196/89/1	N.Y....K.A.	**QG**	.L...V.V.	.F.M..S.SR	**QT**	.SA..G.E..	RTA.TK.V	**KQ**	.C...........
Ken 3/57/2	N.Y......G.	**QG**	.V....A..	.F.F....VR	**QT**	.SA..G.EV.	FDAP.GVA	**KQ**	....
Zim 7/83/2	N.F....K.A.	**QG**	.I....F..	.F.F.V..VR	**QT**	.SS..G..V.	FDSP.GVE	**KQ**	.C...........
Bec 1/65/3	N.YR...R...	**QG**	......N..	......N.AT	**QT**	.SA..G..V.	.TPLVK.D	**KQ**	MC...........
KNP 10/90/3	N.Y....K.T.	**QG**	.V....H..	.F....N.VA	**QT**	.SA..GG.V.	.T.LV..D	**KQ**	.C...........

Figure 2.4.1. A multiple alignment of the deduced amino acid sequence of the proteolytic cleavage sites VP0/VP3, VP3/VP1 and VP1/2A of twelve FMDV isolates representative of the seven FMDV serotypes. Isolates originating from Europe, Asia and South America are indicated in green, while isolates originating from southern Africa are indicated in blue. The precise catalytic sites are indicated in bold. Residues within the cleavage sites represented in red italics indicate amino acid differences compared to the deduced sequence of the A12 isolate.