Chapter 5

FMDV Capsid Stability and Variation Analysis

5.1. Introduction

The capsid of the FMD virus consists of 60 copies of a four chain protomer derived from polypeptide P1 (Fig. 5.1) and is ca. 300Å in diameter. The P1 polypeptide is cleaved into three parts by the $3C^{pro}$ or $3CD^{pro}$ protease complex which results in the VP3, VP1, and VP0 peptides. Autocatalytic cleavage of VP0 into VP4 and VP2 is the last step in capsid assembly. One of the 60 protomers consists of chains VP1-4 encoded by the 1A-D coding regions on the FMDV genome. VP1-3 each consists of a β-barrel (8-stranded) in a jelly-roll topology (Fig. 5.1; Acharya *et al.*, 1989). The pentamers formed by five protomers, associate through an α-helix situated in the VP3 protein (Ellard *et al.*, 1999). This helix associates with its reciprocal helix as well as with His 142 in the opposite pentamer. Curry and co-workers have proposed that His 142 is vital in keeping the protomers together (Curry *et al.*, 1995). His 142 in VP3 of one pentamer associates with the positive dipole formed at the one end of the α-helix in VP2 of the opposite pentamer. It was speculated that the protonation of His 142 may prevent capsid assembly. Other histidine residues (His 145 on VP3 and His 21 on VP2) in close proximity were thought to also play a role in capsid assembly and uncoating. Mutation studies showed that if His 142 is replaced with an arginine, there is almost no capsid formation (Ellard *et al.*, 1999).

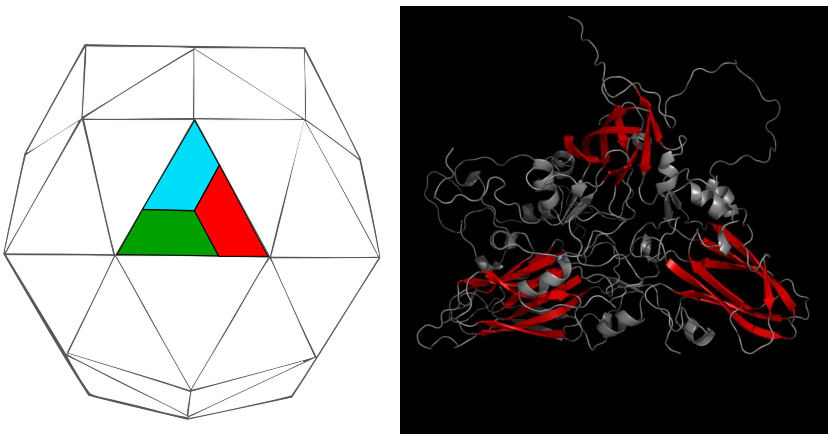

Figure 5.1: Left: A schematic representation of the FMDV icosahedral capsid. The capsid consists of 60 copies of each of the protomers. VP1: blue, VP2: green, VP3: red. Right: The 8 stranded β-barrel in a jelly-roll topology in VP1-3. β-barrels are coloured red.

The arrangement of the structural proteins in the capsid provides the antigenic sites important for eliciting neutralizing antibodies following infection or vaccination. VP1-3 forms the outside of the capsid while VP4 is completely buried inside the capsid. The FMDV capsid, unlike other Picornaviridae, also functions as a general scaffold to keep the RNA protected from the *in vivo* environment and mediates the binding to cellular receptors during cell entry. Interaction with cellular receptors is via the flexible βG-βH loop of VP1. This exposed loop contains an RGD (Arg-Gly-Asp) motif (Logan *et al.*, 1993) involved in binding integrin-receptors of which the αvβ1, αvβ3, αvβ6 and αvβ8 are known to be utilized by FMDV. In the well-studied O serotype, this GH-loop contains a cysteine residue at its base that allows the formation of a disulphide bond with a cysteine in VP2. This adds some stability to the loop and may aid in the receptor preference of the virus. Although field viruses use the integrin-receptors for infection, cell culture-adapted viruses obtain the ability to utilize an alternative receptor *i.e.*. heparan sulfate proteoglycans (HSPG) to enter cells (Fry *et al.*, 1999).

Previous structural work on FMDV includes crystallization of serotype O, A and C capsid at various resolutions (Acharya *et al.*, 1989;Curry *et al.*, 1992; Fry *et al.*, 1993; Lea *et al.*, 1994). These structures were used to identify the important areas such as the RGD-containing GH-loop in VP1. Later studies (Fry *et al.*, 1999; Fry *et al.*, 2005) used crystal structures to identify binding sites for HSPG on the capsid. The HSPG binding

site was identified as a shallow depression at the junction of VP1, VP2 and VP3. Residue 56 of VP3 was identified as being important in the interaction with HSPG (Jackson *et al.*, 1996). In the wild type, residue 56 is a histidine but upon cell culture adaptation, this changes to an arginine which associates with high affinity to HSPG. Curry and co-workers (1996) postulated that the GH-loop flexibility affects the movement and interaction of VP3 and in turn mutations in VP2 affect the GH-loop on VP1.

Dr. F. Maree and co-workers have shown *in vitro* with FMDV that two of the SAT2 serotype capsids (ZIM/5/83/2, ZIM/7/83/2) differ by only six residues (located on the surface), yet ZIM/5/83/2 has more infectious particles and is more stable following treatment at pH 6.0 than ZIM/7/83/2 (unpublished work). At pH 6.0 infectious particles of ZIM/5/83/2 could still be detected while ZIM/7/83/2 lost infectivity. ZIM/7/83/2 also adapted to using HSPG to infect cells and kills cells with a high efficiency. In contrast, ZIM/5/83/2 does not use HSPG to infect cells and has a low cell-killing efficiency. The HSPG adaptation is a known result from viral passage through cultured cells (Sa-Carvalho *et al.*, 1997; Fry *et al.*, 1999). This is important in their vaccine research work and implies that small mutations on the capsid can play a vital role in capsid stability and infectivity. The work presented here will try to characterize the variation and link it to the structure of the capsid proteins in an attempt to explain the results seen *in vitro.*

5.2. Methods

The complete modelling of a virus capsid is very time consuming and resource intensive. An alternative approach is to use a protomer (in this case the assembly formed by VP1-4) and then, using symmetry operations, generate the complete capsid assembly. 1ZBE (Fry *et al.*, 2005) from the PDB was used to generate the complete capsid. This resulted in a complete virus capsid which showed the interactions between the different chains. It also showed the pore structures that are involved in ion movement into and out of the virus.

5.2.1. Capsid Protomer

Protomer models of six SAT2 strains were constructed (ZAM/7/96/2, ZIM/14/90/2, ZIM/17/91/2, ZIM/5/83/2, ZIM/7/83/2, SAU/6/00/2) based on the crystallographic coordinates of O1BFS (1FOD) (Logan *et al.*, 1993). With the exception of SAU/6/00/2, the remaining strains have been found to be prevalent serotypes of the SAT2 family in the western and northern geographical regions of Southern Africa. The sequence data for the strains were provided by Dr. Francois Maree from the TADP, Agricultural Research Council, South Africa. Alignments for all the models were done with ClustalX using the default parameters, the modelling scripts were generated using the Structural module in FunGIMS and models were built using Modeller 9v1 (Fiser and Sali, 2003).

A PROPKA (Li *et al.*, 2005) analysis of each protomer (ZIM/5/83/2 and ZIM/7/83/2) was also done to assist in identifying major protonation states affected by a pH of 6.0. Yasara was used to analyze any hydrogen bond networks present.

5.2.2. Capsid Pentamer

A model of a pore (capsomer, 5-fold symmetry) consisting of five protomers was selected from the generated capsid model by deleting all unnecessary chains. This was used as a basis to investigate the effect of the different mutations found in the various strains and the way in which they influence chain-chain interactions as well as protomer-protomer interactions. Pentamer models of ZIM/5/83/2 and ZIM/7/83/2 were also built using the 1ZBE-generated capsid (strain A1061) as template. Alignments were done with ClustalX using the default parameters, modelling scripts were generated with the Structural module of FunGIMS and models with Modeller 9v1. The template lacked certain residues and for the modelling process these residues had to be removed from the targets due to no template matching (residues 140-158 from VP1, the first 7 residues of VP2, the first 14 residues from VP4 and residues 40-59).

To investigate pH-dependant differences between ZIM/5/83/2 and ZIM/7/83/2 pentamers, a molecular dynamics simulation was done for ~2.5ns. Yasara was used to do the dynamics. The simulation was run at a pH of 6.0, water density of 0.997 g/ml, a NaCl

concentration of 0.9%, using the Amber99 forcefield with periodic boundary conditions at a temperature of 298K. These simulation conditions were applied to both the respective protomers as well as the pentamers. A molecule consisting of two protomers (henceforth called the dimer) was also generated to analyze the interface between two pentamers.

5.3. Results and Discussion

5.3.1. Capsid Modelling

The complete capsid was generated with symmetry operations and used as a template for the investigation of the various proteins involved in capsid assembly (Fig. 5.2). The pore is located at the 5-fold axis (Fig. 5.3) and is comprised of five protomers. One VP1 chain from each protomer forms the pore. This was used as a basis for investigating the interactions between the five protomers and the chains in the protomers. Figure 5.3 shows the interaction between the VP2 and VP3 chain in the five protomers.

After analysis and structural mapping of the variation it is clear that the core of the capsids is quite conserved. The observed variation is probably the result of the quasi-species nature of the FMDV genome and positive selection pressure exerted on phenotype level. Variation seemed to occur mostly on surface areas and areas close to protein-protein interfaces. Although most of the variation is neutral, some of the variable residues result in the addition or loss of interactions. These specific differences may change the capsid assembly and disassembly dynamics slightly but none of the conserved amino acids identified as playing a role in capsid stability were affected. A far more detailed study of variation and structure would be required to identify individual interactions deemed to be important in capsid structure.

5.3.2. Protomer Modelling and Variation Mapping

Recently there has been considerable interest in the structural basis of the effect of pH on FMDV (Curry *et al.*, 1995). Furthermore, Doel and Baccarini (1981) reported a direct correlation between thermal stability of 146S particles and the protective ability of a

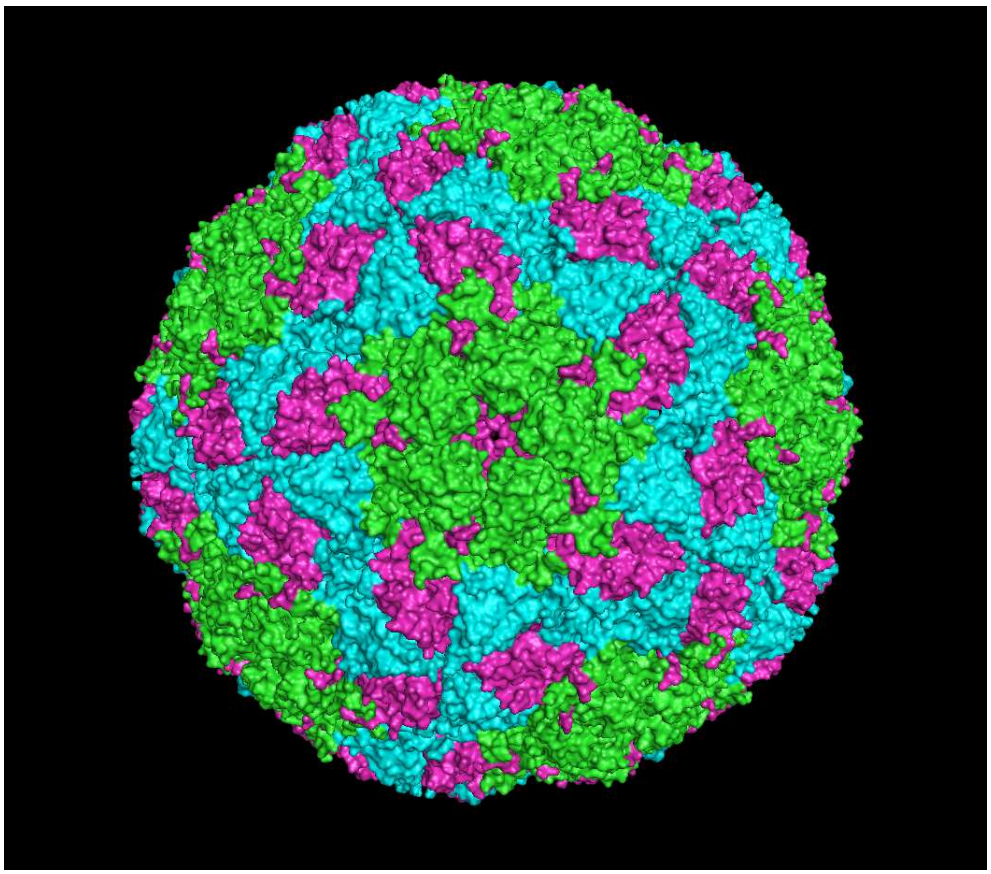

Figure 5.2: The capsid as generated from 1ZBE using symmetry operations. Green: VP1, Cyan: VP2, Magenta: VP3, VP4 - hidden on the inside of the capsid. The pore at the 5-fold axis is the in the centre of the image surrounded by 5 VP1 chains (green).

vaccine. Dr. F.F. Maree and colleagues examined the stability of SAT2 viruses from different topotypes in southern Africa (Haydon *et al.*, 2001; Bastos *et al.*, 2003) as well as a SAT3 virus to different pH environments to compare the phenotypic variance within these serotypes. The southern Africa SAT2 and SAT3 isolates can be divided into three lineages based on 1D phylogenetics, supporting a southern, western and northern clusters (Haydon *et al.*, 2001; Bastos *et al.*, 2003). Two of the viruses, i.e. SAT2/ZIM14/90 and SAT2/ZIM/17/91, belong to the western lineage of SAT2 viruses. A third SAT2 virus, belonging to the northern lineage of southern Africa SAT2 isolates, i.e. SAT2/ZAM/7/93 was included in this study. Also available was a SAT3 virus from the same geographical region, designated as SAT3/ZAM/4/96. Treatment of the SAT2 and SAT3 viruses with a buffer of pH 6.0 revealed differences toward there stability in mild acidic environment even within a serotype. Both the SAT2 and SAT3 Zambian isolates lost their infectivity

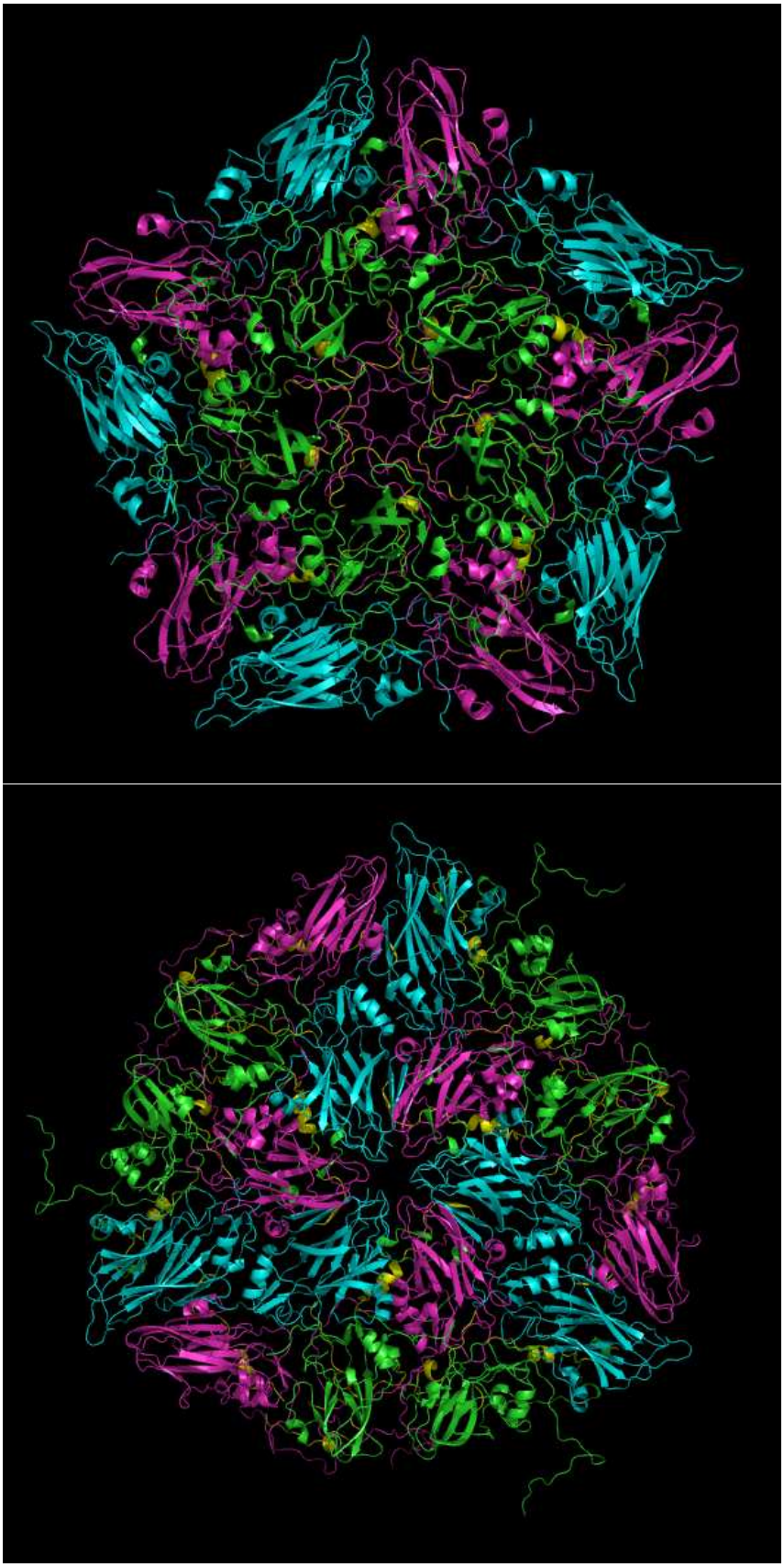

Figure 5.3: Top: A complete 5-fold pore assembly comprised of 5 protomers. Bottom: A 6 protomer complex surrounding the 3-fold pore showing the association between VP2 and VP3. Green: VP1, Cyan: VP2, Magenta: VP3, Yellow: VP4.

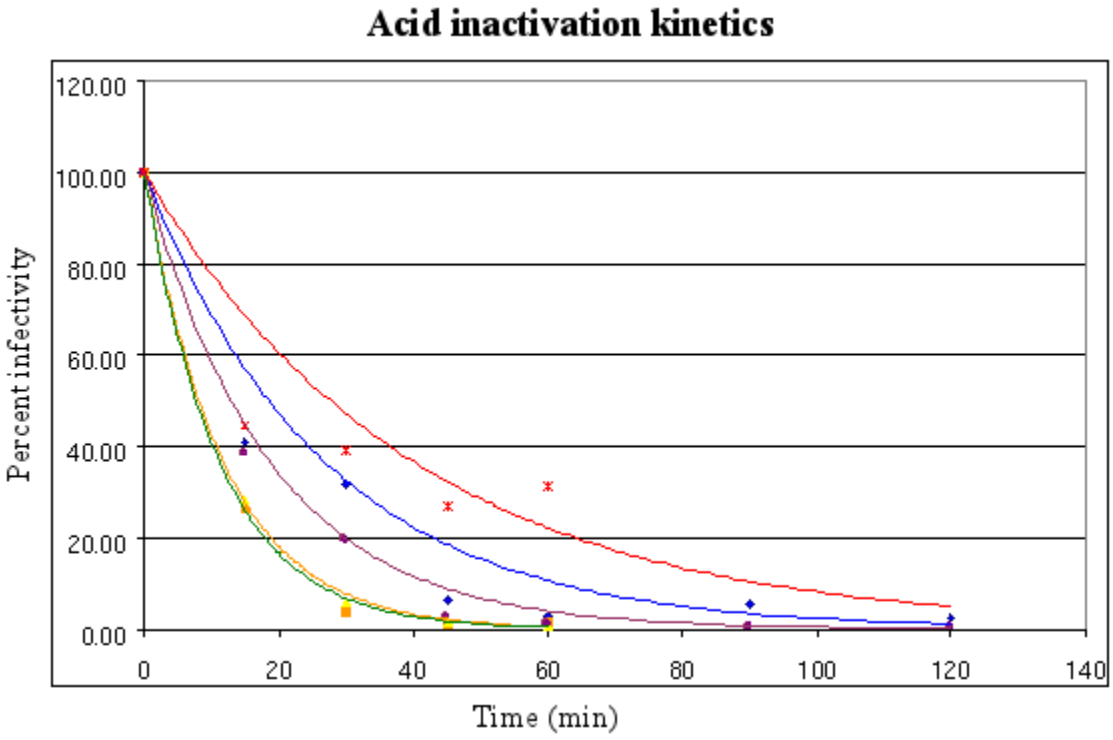

Figure 5.4: The sucrose density gradient purified viruses at an approximate titre of 4-9[106] were treated at pH 6.0 for different lengths of time following a 1:50 dilution in the appropriate NET buffer (150mM NaCl, 10mM EDTA and 100mM Tris). The percentage of infectious particles remaining after treatment was determined by plaque titrations on BHK-21 cells and plotted against time. The exponential declines were used to calculate the inactivation rate constants (described by Mateo *et al.*, 2003).The acid inactivation kinetics of the viruses were reflected by the inactivation rate constants at pH 6.0, which were 0.025, 0.044, 0.065, 0.090 and 0.085 for ZIM/7/83, ZIM/14/90, ZIM/17/91, ZAM/7/96 and SAT3/ZAM/4/96, respectively. Red: ZIM/7/83. SAT2: blue - ZIM/17/91, purple - ZIM/14/90, green -ZAM/7/96. SAT3: Yellow - ZAM/4/96. Data courtesy of Dr F.F. Maree.

completely at a pH of 6.0 with 30 minutes treatment (Fig. 5.4). In contrast all three the SAT2 isolates from the western lineage of southern Africa isolates revealed significant drop in titres following 30 minutes at pH6.0 but in one instance, i.e. ZIM/7/83, at least 40% infectivity was still present after 1h incubation. Since FMDV relies on acid induced disassembly of the capsid proteins for infection and release of RNA this variation in the acid stability of SAT2 virions was further investigated by mapping the amino acid variation on the modelled 3D structure of a SAT2 virion.

Protomer models for six SAT2 strains, summarised in Table 5.1 - 5.3, were built using the alignments in Figure 5.5. The resulting models were compared to the pore model as well

as to one another. All differences were classed into three categories: no effect (normal variation or surface exposed without any change in local structure or interactions), effect on intra-protomer association and effect on inter-protomer association. The results are summarized in Tables 5.1, 5.2 and 5.3. As can be expected, most of the variation was found in the VP1 chain, the most variable of the capsid proteins. VP4 did not show any significant differences. Most of the differences seen could have a possible effect on protomer-protomer interaction although in isolation, single differences might have a very small effect. Overall it seems that most of the differences in the chain could have have a small effect on the inter-protomer interaction and to a far lesser degree, intra-protomer interaction.

The variation in the capsid was also mapped to a model of the pentamer (Fig. 5.6). The variation only included mutations that would change the type or amount of interaction. This showed that such variation mostly occurs on interfaces and, significantly, around the pore and pore wall at the 5-fold axis. Most of the mutations do not appear to influence the structure, but some of the variation around the pore wall could have effects with regard to other virus functions such as adhesion and ion movement.

5.3.3. Pentamer Molecular Dynamics

SAT2/ZIM/7/83 was considered an efficient vaccine strain for many years in the southern Africa region in view of the fact that it produced high yields of 146S antigen, was considered to be a stable virus in the production process and elicited a strong immune response (Esterhuysen *et al.*, 1988). The recent inability to produce sufficient yields of 146S particles in cell culture monolayers lead us to investigate genetic changes that may affect the stability of the virus. ZIM/5/83/2 and ZIM/7/83/2 showed a difference of 6 residues (Table 5.4) and this resulted in a differing stability at pH 6. The pH_{50} can be described as the half-way point in the transition of 146S infectious particles into 12S pentamers and was described by Curry *et al.*, 1995 as a measure of pH sensitivity. The pH_{50} for both the SAT2 viruses (Fig. 5.7) were similar at pH 6.6 and comparable to serotype A viruses (Curry *et al.*, 1995). Nevertheless, between pH 5.8 and 6.3 the infectious particles deteriorate rapidly, probably as a result of break down into 12S

Table 5.1: The results from a comparison of the VP1 chain of the 6 SAT2 strains used in this study. Differences that do not have an influence on interaction were ignored (e.g. Ile -> Val). Strains: 1: ZAM/7/96, 2: ZIM/14/90, 3: ZIM/17/91, 4: ZIM/5/83, 5: ZIM/7/83, 6: SAU/6/00. The ZIM/7/83 proteome sequence was used as a reference sequence.

VP1	**Strains**						
#	**1**	**2**	**3**	**4**	**5**	**6**	**Effect**
6	E	E	G	E	E	E	Possible ionic interaction with VP2 (intra-protomer), Gly would disrupt this interaction.
21	R	S	S	A	A	N	Interaction with VP2 (inter-protomer), Arg, Asn might show stronger interaction with VP2.
23	V	A	M	T	T	Q	Interaction with VP2 (inter-protomer), different side chains might have different interaction strengths, Gln introduces a charge.
28	M	M	M	M	V	K	Interaction with VP3 (inter-, intra-protomer), Lys introduces a $\delta+$ charge.
39	F	F	F	F	F	S	Ser completely lacks the hydrophobic interaction present in other strains.
43	H	H	H	L	L	H	Exposed to surface. His may gain ionic interaction with VP1 (inter-protomer), Leu may disrupt interaction with VP1 (inter-protomer).
57	K	N	N	N	N	K	Ionic residue can interact with VP3 (inter-protomer), Lys lacks a $\delta-$ charge.
83	E	D	T	E	E	D	Situated on the exposed outer edge of the pore. Interacts with receptors in conjunction with residue 85. Thr lacks the $\delta-$ charge.
85	A	K	K	E	E	T	Situated in the wall of the pore, exposed to surface and might interact with receptors. Ala lacks any charge, Lys only presents $\delta+$ charge while Glu presents $\delta-$ charge.
101	G	R	R	R	R	G	Pore wall, ionic interaction with VP1, VP3 (inter-protomer), Gly lacks any side chain charge.
111	K	S	S	N	N	G	On the outer edge of the pore, longer side chains such as Lys may gain interaction with VP1 (interprotomer). Gly loses all possible charged interactions.
129	R	R	R	R	R	V	Val loses the charged interaction with between VP1 and VP2.
147	R	R	W	R	R	R	On surface, interactions with VP2, VP3 (intra-protomer).
200	S	G	G	A	A	T	Interaction with VP3 (inter-protomer), Ser might have one more hydrogen bond.

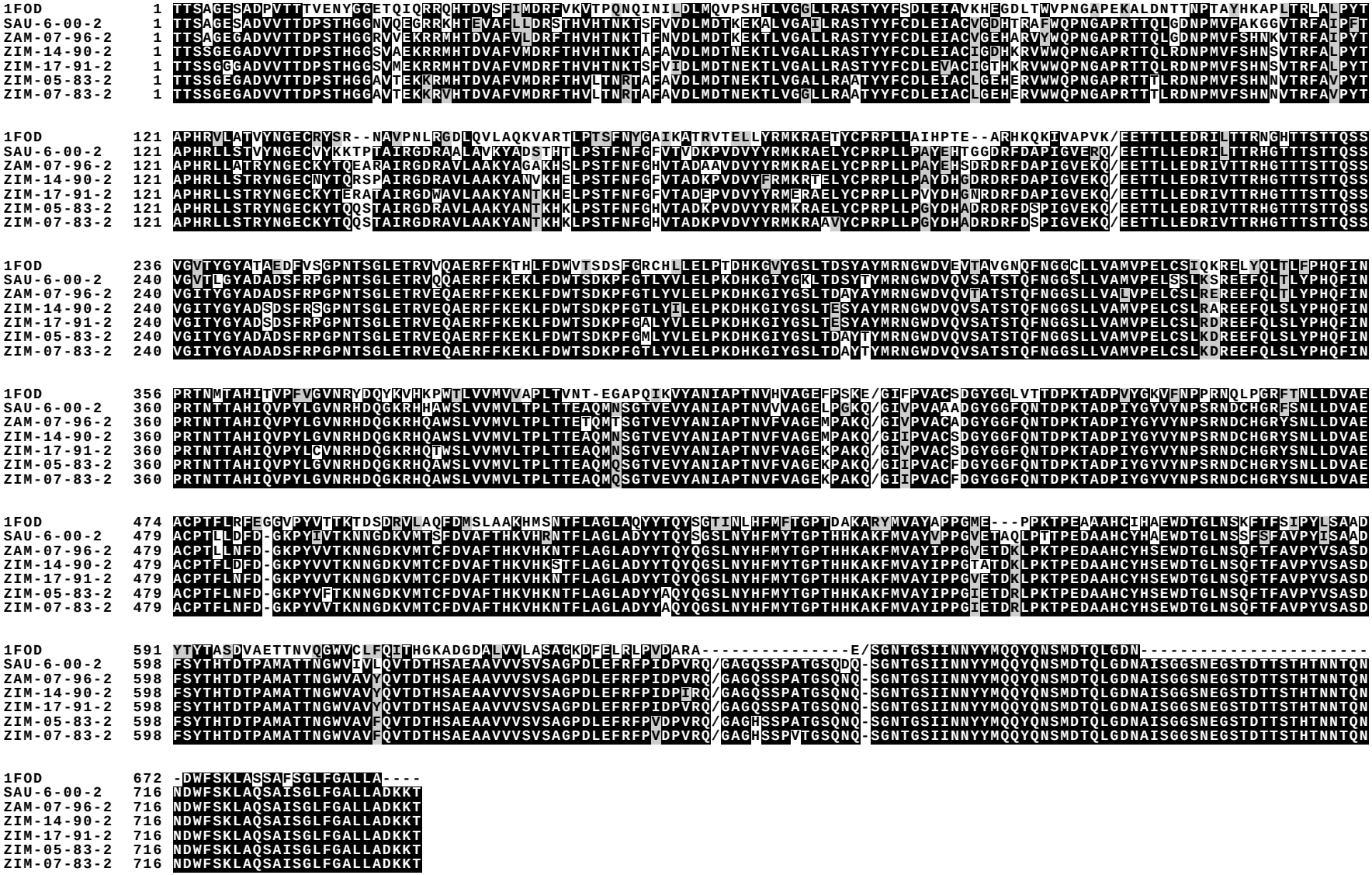

Figure 5.5: The alignments used to model the six SAT2 capsid protomers. The identity between the targets and template are all around 60% with a variation of 1%.

Table 5.2: The results of a comparison of the VP2 chain of the 6 strains used in this study. Differences that do not have an influence on interaction were ignored (e.g. Ile -> Val). Strains: 1: ZAM/7/96, 2: ZIM/14/90, 3: ZIM/17/91, 4: ZIM/5/83, 5: ZIM/7/83, 6:SAU/6/00. The ZIM/7/83 proteome sequence was used as a reference sequence.

VP2	**Strains**						
#	**1**	**2**	**3**	**4**	**5**	**6**	**Effect**
51	E	E	E	E	E	Q	Charged Gln can affect protomer interaction.
88	S	S	S	S	S	K	Charged Lys can interact more strongly with other protomers.
91	A	S	S	A	A	S	Interaction with VP2 (inter-protomer), Ser might induce an extra hydrogen bond.
93	A	A	A	T	T	T	Interaction with VP3 (intra-protomer), Thr might induce an extra hydrogen bond.
188	T	N	N	Q	Q	N	Interaction with VP3 (inter-protomer), Thr might disrupt the ionic interactions seen in Gln and Glu.
209	M	M	K	K	K	L	Interaction with VP2 (inter-protomer), Met, Leu might disrupt the ionic interactions seen in Lys.

pentameric units (Curry *et al.*, 1995; Knipe *et al.*, 1997; Mateo *et al.*, 2003). The rate of loss of infectious particles was not equal for ZIM/7/83 and ZIM/5/83 at the low pH range (Fig. 5.7), with the infectivity of ZIM/7/83 deteriorating more rapidly than ZIM/5/83 below pH 6.2. Although the starting titer of the two viruses was normalized, the ZIM/5/83 repeatedly end up with approximately 10-80 infectious particles at pH 5.8 and 5.6 respectively, while no ZIM/7/83 infectious particles were present below pH 6.0. However, no infectious particles was repeatedly observed for ZIM/7/83 at these pH conditions. The biological significance of these difference were investigated using models of the 12S pentamers.

Molecular dynamics simulations of the ZIM/5/83/2 and ZIM/7/83/2 pentamers were run for ∼2.5ns. Figure 5.8 shows the RMSD variation over time for each of the two different pentamers over the simulation time at pH 6.0. The protomer simulations of ZIM/5/83/2 and ZIM/7/83/2 were run for ∼2.2ns. The RMSD variation over time are shown in Figure 5.8. There was no significant difference in the RMSD of either the pentamers or the protomers of ZIM/5/83/2 and ZIM/7/83/2. Any significant difference such as a pentamer dissociation would have shown highly divergent RMSD values.

The PROPKA results showed that there were four interesting His residues to investigate.

Table 5.3: The results from a comparison of the VP3 chain of the 6 strains used in this study. Differences that do not have an influence on interaction were ignored (e.g. Ile -> Val). Strains: 1: ZAM/7/96, 2: ZIM/14/90, 3: ZIM/17/91, 4: ZIM/5/83, 5: ZIM/7/83, 6:SAU/6/00. The ZIM/7/83 proteome sequence was used as a reference sequence.

VP3	**Strains**						
#	**1**	**2**	**3**	**4**	**5**	**6**	**Effect**
3	V	I	V	I	I	V	Forms part of the central pore, has an effect on the size of the pore. Other serotypes have a Phe in this position.
8	A	S	S	F	F	A	Situated in the pore opening. The Phe will close up the pore and might be a compensatory mutation for position 3. Other serotypes have an Ala in this position. The Ala and Ser is smaller in size and thus allows for a slightly bigger pore.
54	L	F	F	F	F	L	Hydrophobic interactions with VP2 (intra-protomer), Leu lacks ring which reduces hydrophobicity.
64	V	V	V	F	V	V	Surface exposed but possible interaction with VP2 (inter-protomer). Phe might disrupt sheet formation slightly.
87	N	S	N	N	N	N	Surface exposed, interaction with VP1 (inter-protomer). The Ser interaction might be slightly less due to the OH group.
98	T	T	T	A	A	T	Ionic interaction with VP1 (intra-protomer). The Ala lacks an OH group to form hydrogen bonds.
129	V	T	V	I	I	V	In combination with site 130, this forms the binding area for Heparan sulfate, interaction with VP2 (inter-protomer). The Thr OH group might form extra interactions thus compensation for the Ala in position 130.
130	E	A	E	E	E	E	In combination with site 129, this forms the binding area for Heparan sulfate, interaction with VP2 (inter-protomer). The Ala will result in a loss of hydrogen bonds when compared to Glu.
137	K	K	K	K	K	T	Thr disrupts charged inter-protomer interaction.

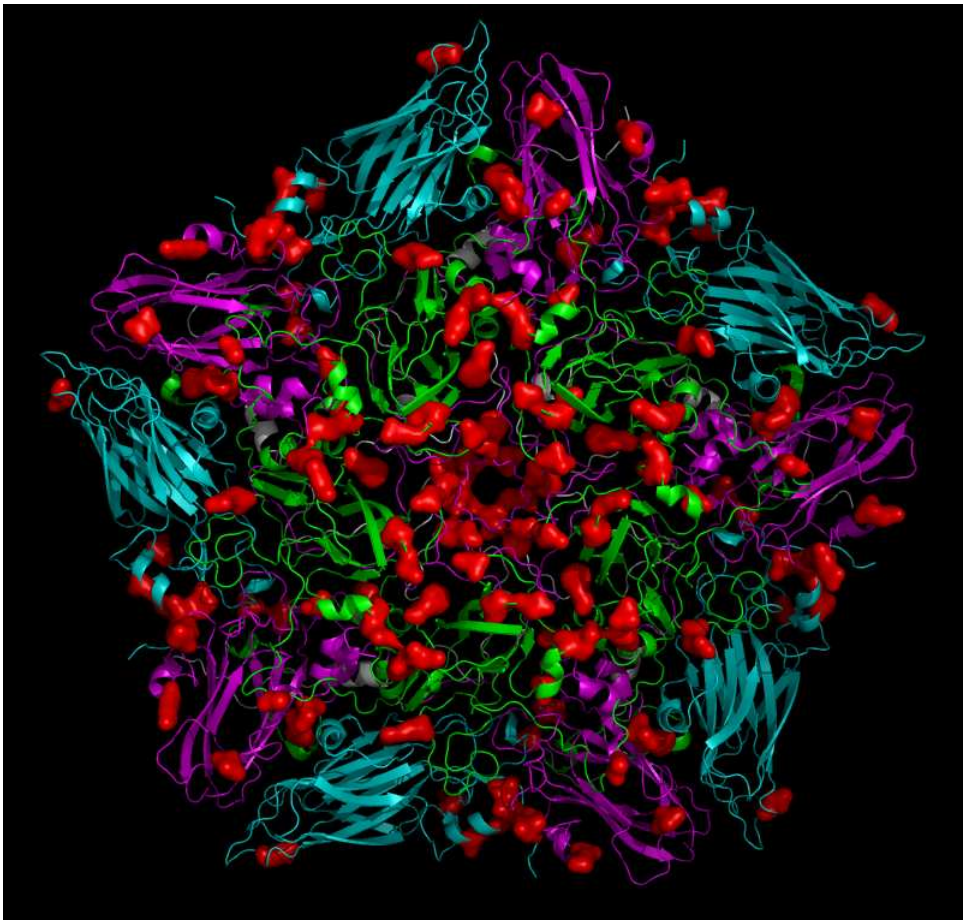

Figure 5.6: The variation seen in VP1-3 mapped to a 5-fold axis model of the protomers. Variable positions are coloured red. Green: VP1, Cyan: VP2, Magenta: VP3.

These had a shift from below a pKa of 6.0 to above a pKa of 6.0 between ZIM/5/83/2 and ZIM/7/83/2. The four His residues were: His 511 (81), His 545 (115), His 575 (145) and His 602 (172). The numbers in brackets are the residue numbers as referred to in Ellard *et al.*, 1999 (Table 5.5). The PROPKA results for the generated dimer showed different results as most of the His residues identified in the protomers were buried in the dimer interface. This difference in result was due to the fact that when the pentamers associate to form the dimer, the His residues identified are buried in the interface and thus excluded from any water contact. This changed the solvent environment around the His residues.

Molecular dynamics simulation was done for each for each of the pentamers of ZIM/5/83/2 and ZIM/7/83/2. The pentamer models were both built on the same template and thus

Table 5.4: The differences between the P1 peptide of ZIM/5/83/2 and ZIM/7/83/2.

Res #	ZIM/5/83/2	ZIM/7/83/2
28	Met	Val
64	Ala	Gly
186	Glu	Ala
187	Leu	Val
287	Thr	Met
493	Phe	Val

Table 5.5: The pKa values for the four His residues identified by PROPKA as undergoing protonation changes at pH 6.0. The protomer of both ZIM strains were used for the respective pKa predictions and a diner generated from the ZIM/5/82/2 protomer.

	pKa		
His #	ZIM/5/83/2	ZIM/7/83/2	Dimer
511	3.21	7.07	3.21
545	6.15	5.94	6.15
575	5.92	7.62	-1.12
602	6.51	5.27	6.51

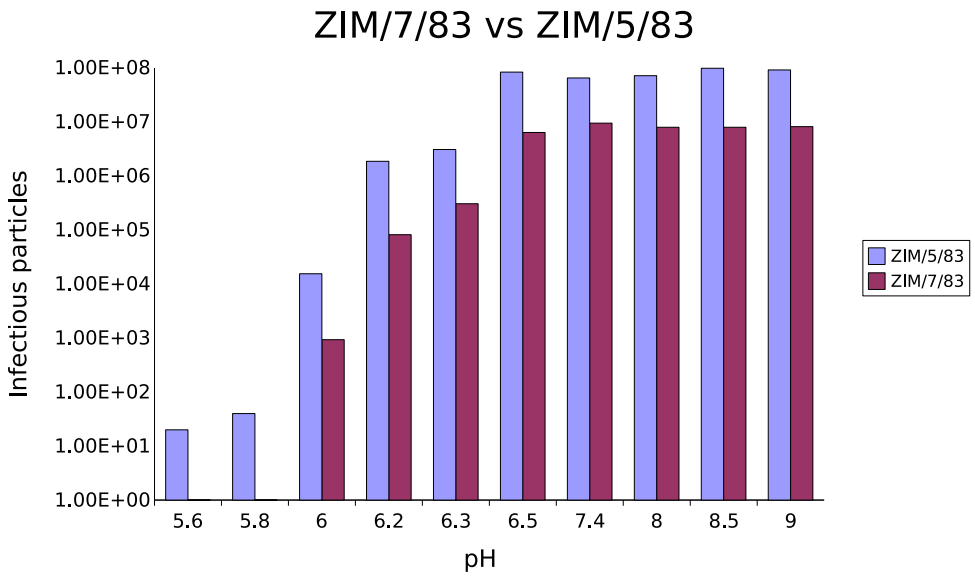

Figure 5.7: The sucrose density gradient purified ZIM/7/83 and ZIM/5/83 infectious 146S particles were incubated in buffered solutions spanning a pH range of 5.6 to 9.0. Following 30 min incubation the pH of the solution was restored and the amount of infectious particles remaining determined by titration on BHK-21 cells. Both SAT2 infectious particles were stable at a wide range of pH conditions from 6.5 to 9.0 for a period of 30 min and up to 2 hours (data not shown). At pH 6.5 at least 35% and 38% of infectious particles for ZIM/7/83 and ZIM/5/83 respectively, were still present after 30 minutes. Data courtesy of Dr F.F. Maree.

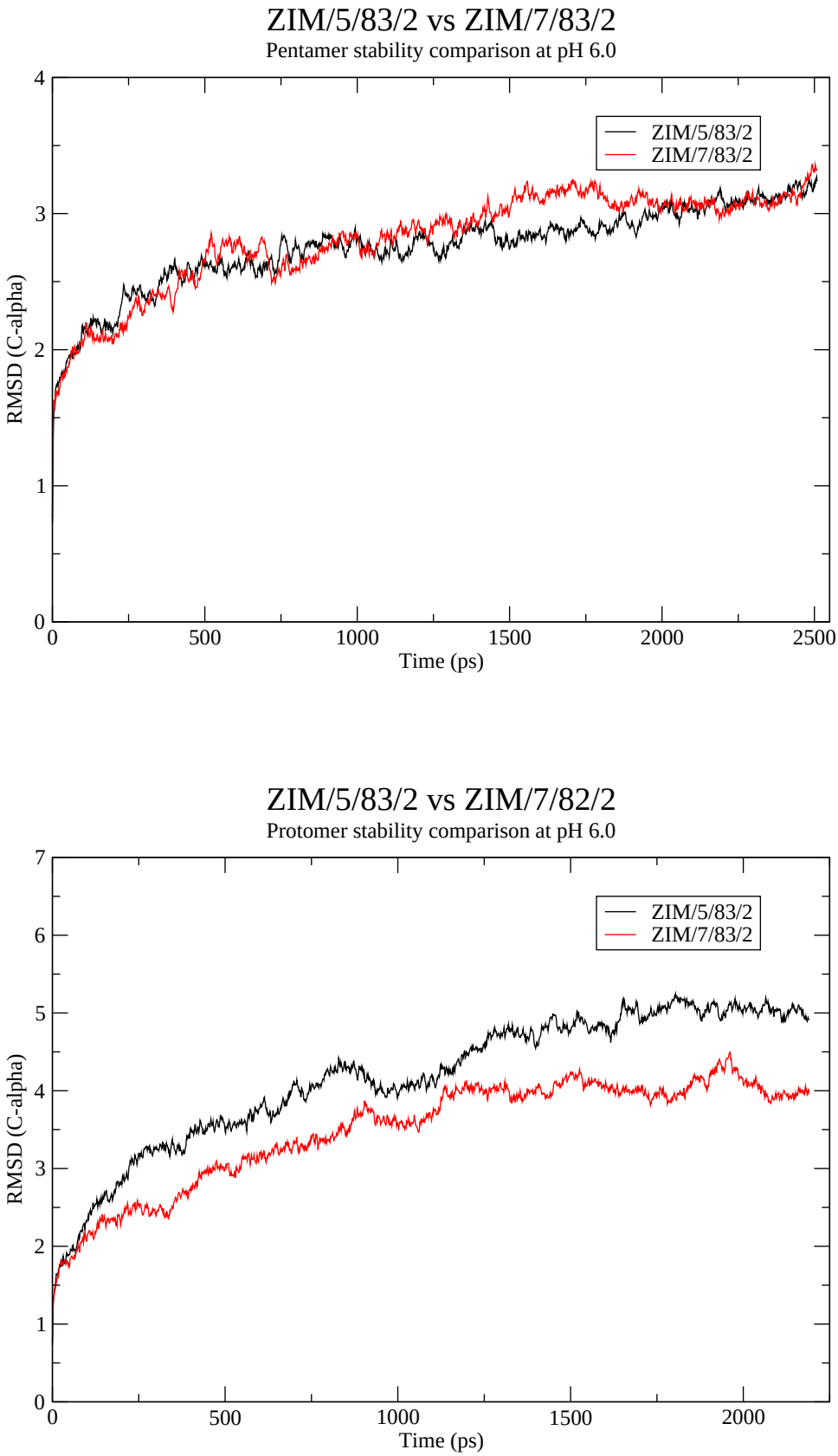

Figure 5.8: The Cα RMSD variation of ZIM/5/83/2(black) and ZIM/7/83/2 (red) over the $\sim$2.5ns simulation time at pH 6.0. Top: Pentamer stability. Bottom: Protomer stability.

after a dynamics simulation, the results in terms of RMSD deviation from the model could be compared. After the 2.5ns simulation run, it was seen that there was no significant difference in RMSD between the pentamers. The graphs showed that the RMSD deviation started to flatten out and thus is was concluded that the pentamers were stable. *In vitro* evidence showed at pH 6.0 ZIM/7/83/2 was less stable than ZIM/5/83/2. There was only a 6 residue difference between the two pentamers and none of these residues were predicted to show any change in protonation state from pH 7.0 to pH 6.0. Thus it was speculated that the lower pH may disrupt general association between the protomers in the pentamers. The simulation showed that no major changes occurred as can be seen in Figure 5.8. A major change, such as the pentamer dissociating, would have showed prominently on a graph plotting RMSD. In order to investigate whether the disruption occurs at protomer level, molecular dynamics simulations were run on the respective protomers as well. The results showed that the protomers were stable and thus the residues had no effect on protomer stability (Fig. 5.8). The difference between the RMSD levels of the two plots are as a result of the presence of loops in the protomer. The movement of these loops will affect the RMSD calculations but not to such an extent as to mask an unstable protomer. A factor to consider is that some residues, mainly in VP4, could not be resolved from the electron density maps during structure determination and could thus not be modelled. This simulation showed that when considering RMSD, the protomers stayed relatively stable with no major increase in RMSD as would have been expected for a protomer dissociating. This implies that the pH disruption occurs at another level. This dissociation may be investigated in the future by using binding interaction studies on the individual components using equipment such as a biosensor.

It was decided to perform a pKa prediction on both protomers as well as the dimer to see whether there is any change in pKa. The PROPKA pKa prediction results for the protomers indicated four interesting His residues which change protonation states around pH 6.0. These residues were inspected manually. When considering the pattern of binding by these His residues, it would appear that ZIM/5/83/2 and ZIM/7/83/2 do not gain or lose a nett amount of bonds (Table 5.6). However when the residues are mapped to the

Table 5.6: The changes in pKa for the four His residues in the protomer identified by PROPKA as undergoing protonation changes at pH 6.0. All residue numbers refer to the residues in the full model.

His #	**ZIM/5/83/2**	**ZIM/7/83/2**	
511	3.21	7.07	This His is exposed to the surface and thus the solvent environment would affect this pKa massively. No conclusions can be drawn about this residue.
545	6.15	5.94	In ZIM/5/83/2 the His is pointing towards solvent, whereas in ZIM/7/83/2 it is pointing inwards towards the protein. It also interacts with the adjacent pentamer.
575	5.92	7.62	This His is pointing towards the interface with another protomer. It is implicated in interprotomer association.
602	6.51	5.27	Exposed to the surface.

structure a different picture emerges. From the structure it can be seen that all these changes occur on the VP3 chain (Fig. 5.9).

His 575 (145 in chain C) in VP3 is associated in interprotomer interaction (Curry *et al.*, 1995; Ellard *et al.*, 1999). Hydrogen bond analysis of the dimer molecule indicated that His 575 (145) interacts with Ala 571 (141 on chain C of protomer 1) and with Lys 273 (63 on chain B of protomer 2). The PROPKA results for the dimer molecule show the pKa for His 575 to be -1.12. It must be kept in mind that this is a statistical calculation (in a non-water environment) and implies that the residue is deprotonated most of the time. The fact that it is not exposed to solvent influences the pKa predictions as well. Thus, from these results it appears that a pH below 6.0 would prevent the formation of the capsid, as pentamers cannot assemble. It appears that a significant proportion of His 575 (145) needs to be neutral for pentamers to assemble into a capsid. This confirms the work done by van Vlijmen and co-workers (1998) in which they calculated that His 575 (145) may play a role in capsid disassembly (van Vlijmen *et al.*, 1998) and the work of Twomey and co-workers (Twomey *et al.*, 1995) on FMDV vaccine stability. Various authors (Curry *et al.*, 1995; Ellard *et al.*, 1999) also showed that His 572 (142) was important in association between the pentamers. The hydrogen bond analysis showed that His 575 (145) made a hydrogen bond with the backbone of Ala 571 (141), which is

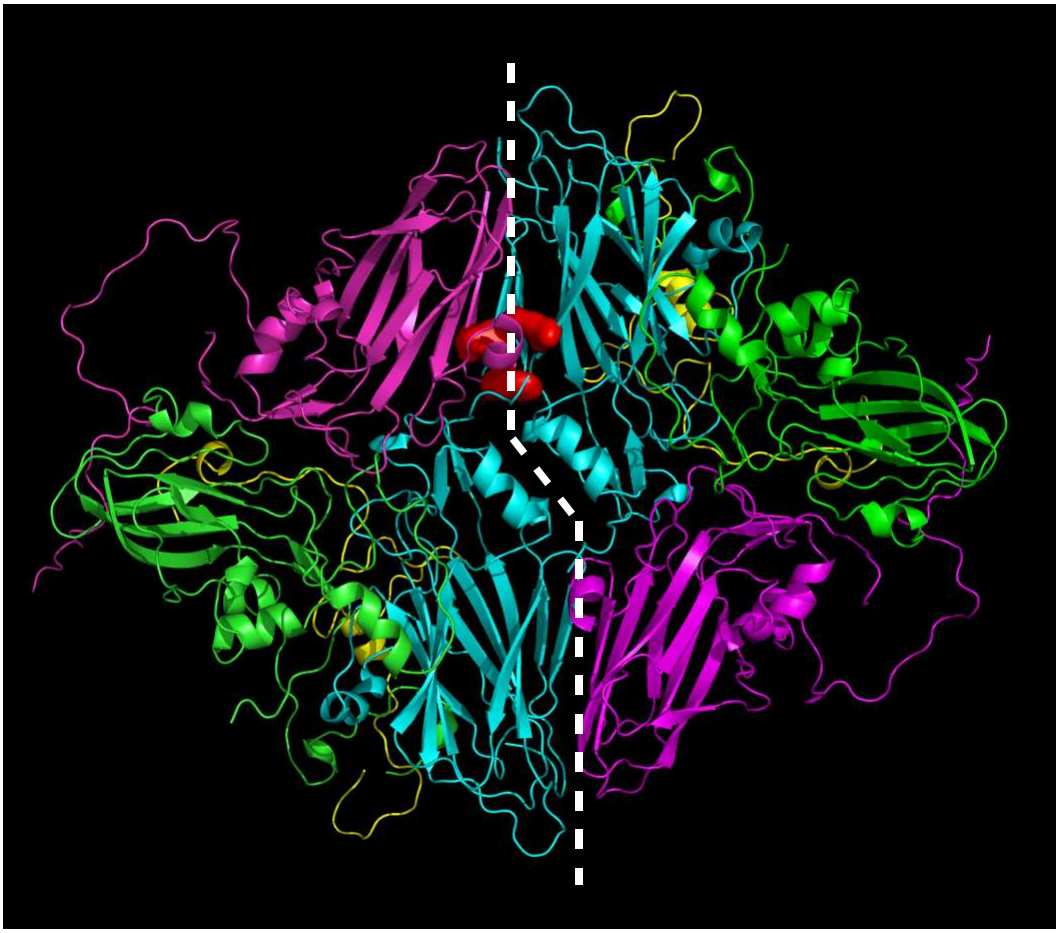

Figure 5.9: The interaction interface between two pentamer sections. One protomer of each pentamer is shown. The dashed line indicates the interaction surface. His 575 (145), His 572 (142) and Lys 273 (63) are coloured red using Van der Waals surfaces.

located right next to His 572 (142) (Fig. 5.10). This backbone hydrogen bond seems to be important is helping to orientate the His 572 (142) containing loop correctly to form the association with the charged dipole of the α-helix. His 575 (145) also makes a hydrogen bond with Lys 273 (63) in the adjacent pentamer, thus providing extra interaction and stabilization between the pentamers. The loss of the hydrogen bonds with either Lys 273 (63) or Ala 571 (141) would have a significant effect on the interaction interface.

The PROPKA pKa analysis predicts that the V493F mutation affects the pKa values of the ZIM/5/83/2 and His 575 and thus makes it neutral above pH 5.92, while ZIM/7/83/2 is neutralized at a higher pH. Although the distance between Phe 493 (63 on chain C) and His 575 (145) is $\sim$17.5 Å, long distance effects transmitted through the β-sheet cannot

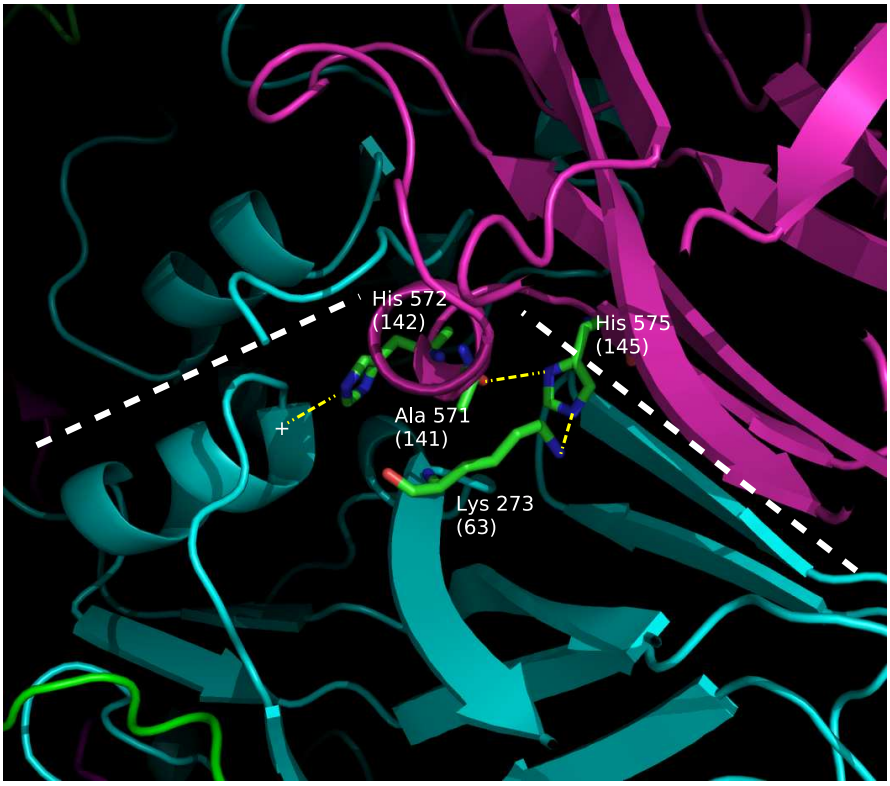

Figure 5.10: The hydrogen bond network found in the pentamer interface. When His 575 (145) is neutral, it makes a hydrogen bond with Lys 273 (63) and Ala 571 (141). The neutral state seem to prevent pentamer association through His 572 (142) and His 575 (145). Yellow dashed lines indicate hydrogen bonds and white dashed line indicates pentamer interface. The "+" indicates the charged dipole of the α-helix.

be ruled out (Fig. 5.11). Thus the neutral His 575 (145) seems to be vital for pentamer assembly. This result is similar to the one noted by Curry and co-workers (Curry *et al.*, 1995) in which it was found that subtype A10 was more stable than A22 by 0.5 pH units and shows that there is variation . It must be kept in mind that these results are based on mostly statistical predictions and that experimental work is required to confirm the results.

5.4. Conclusion

Protein-protein recognition mediates many fundamental biological processes. A detailed knowledge of these processes requires the determination of the structural, energetic, and functional roles of individual amino acid residues and interactions in protein-protein in-

Figure 5.11: A side-on view of VP3 with the location of Phe 493 (63) in relation to His 575 (145). The distance between the residues are ∼17.5 Å. VP1: green, VP2: cyan, VP3: magenta, Phe 493: red and His 575: orange.

terfaces. These studies have been generally undertaken by using small protein-ligand complexes or oligomeric proteins of moderate size (Reguera *et al.*, 2004). In contrast, for multimeric protein complexes, such as viral capsids (Liljas, 1986; Hadfield *et al.*, 1997) or large cellular assemblies, little is known about the specific molecular determinants of protein association and stability. Mutational studies of virus capsids, generally focused on a few specific amino acid residues, have provided important insights (Ellard *et al.*, 1999; Mateo *et al.*, 2003). However, exhaustive experimental studies on the relative importance of residues and molecular interactions in viral capsid assembly, disassembly, and or stability are still limited. These studies contribute also to the understanding of protein structure-function relationships and they could be exploited possibly in the design of thermostable vaccines and antiviral agents promoting capsid disassembly or interfering with assembly (Wien *et al.*, 1996; Hadfield *et al.*, 1997; Diana *et al.*, 1997; Belnap *et al.*, 2000).

Many viruses, including viruses of medical or veterinary significance, have capsids of

icosahedral symmetry (Reguera *et al.*, 2004). FMDV is a small non-enveloped virus with a pseudo T=3 icosahedral capsid formed by 60 copies each of four nonidentical polypeptide chains, i.e. VP1, VP2, VP3 and VP4. There has been considerable interest in the structural basis of the effect of pH on FMDV (Curry *et al.*, 1995). Multiple evidence on the structural data of FMDV had been gathered by high resolution X-ray crystallography in recent years that allow the identification of residues involved in stabilising the virion structure. Assembly of the picornaviral capsid proceeds in several steps (Rueckert, 1996). The capsid proteins VP0 (1AB), VP3 (1C), and VP1 (1D) are translated as a polyprotein precursor (P1), may fold co-translationally (Rossmann and Johnson, 1989), and are proteolytically processed by $3C^{pro}$ (Birtley *et al.*, 2005) to yield the mature protomer. Five protomers are assembled to form a pentameric intermediate, and finally, 12 pentamers are assembled to form the icosahedral capsid (Fig. 5.1). After encapsidation of the RNA genome most VP0 molecules are processed to give VP4 (1A, the N terminus of VP0) and VP2 (1B). Disassembly of the FMDV virion *in vivo* begins with its dissociation into pentamers (Vasquez *et al.*, 1979) by acidification in the endosomes (Carrillo *et al.*, 1984). Furthermore, Doel and Baccarini (1981) reported on a direct correlation between thermal stability of 146S particles and the protective ability of an antigen/vaccine. It was found that mild heating of FMDV virions leads to irreversible dissociation into stable pentamers (Rueckert, 1996), an event that appears as the main cause for the need of a cold chain to preserve FMD vaccines. Analysis of the crystal structure of the FMDV capsid (Acharya *et al.*, 1989; Lea *et al.*, 1994; Lea *et al.*, 1995; Curry *et al.*, 1996; Fry *et al.*, 1999) indicates that the pentameric intermediate subunits interact mainly through a relatively limited number of electrostatic interactions; a role of His-142 of VP3 in the acid-induced disassembly of FMDV has already been demonstrated (Ellard *et al.*, 1999).

A variety of approaches have been used to study the effects of acid. X-ray crystallographic techniques have been used to determine acid-induced structural changes in mengo virus (Kim *et al.*, 1990) and HRV (Giranda *et al.*, 1992). Amino acid changes which affect acid lability, have been identified by the generation and sequencing of acid stable mutants of HRV (Giranda *et al.*, 1992; Skern *et al.*, 1991). Another approach involved computer modelling of the effects of pH on electrostatic interactions within poliovirus and HRV

(Warwicker, 1992). In the present study, we did a side-by-side comparison of the pH stability of SAT2 and SAT3 viruses. The results revealed that SAT2 infectious particles showed similar or even more stability in mild acidic conditions than was previously described for viruses belonging to the A, O and C serotypes, stable in solutions with high ionic strength, but was sensitive to heat (Maree *et al.*, unpublished). Even though the SAT2 viruses used in this study differed by less than 11% in there amino acid sequence of the capsid proteins, the SAT2 virions had a diverse range of sensitivities toward mild acidic conditions. A SAT3 isolate from the same geographic distribution were much more sensitive to acidic environment (Maree *et al.*, unpublished).

Using the tools provided by the Structural module, it was possible to construct models as well as run molecular dynamics simulations. The variation mapping showed that most of the variation in the protomers occurs in areas on the surface as well as close to interface areas. Despite the differences, each individual difference plays only a small part in the overall interaction. The molecular dynamics results showed no real difference in the stability of the pentamers or the protomers at pH 6.0. However a pKa analysis showed that the difference in pH stability of ZIM/5/83/2 and ZIM/7/83/2 was due to the change in pKa of His 575 (145) in VP3. These results indicated that although pentamer association is mediated by many different interactions, there are usually one or two very important residues in the interaction interface. In the case of FMDV the role of His 142 has been proven (van Vlijmen *et al.*, 1998; Ellard *et al.*, 1999). This work predicts that His 145 is important in the initial association of the pentamers and also shows the effect of pH on pentamer association. The simulations conducted also support the current theories that the protonation states of His 142 and His 145 are the determining factors in pentamer assembly.

The work done here provides the local vaccine researchers with data about the predicted behavior of the capsid proteins under certain pH conditions. It provided possible explanations for their results as well as opened up new avenues of research into designing stable vaccines by exploiting the knowledge gained in analysing capsid interactions.

Chapter 6

Concluding Discussion

Structural biology forms the basis of our understanding of the relationship between the structure and function of a protein. The one cannot be studied without the other. Traditional structural biology involved time-consuming experiments to characterize a protein and its structure. In the modern age of structural biology this task has been made easier by the presence of databases that contain a vast amount of data related to the structure and function of a protein. These databases are usually specialized for a specific function such as the PDB, which accepts only three dimensional coordinates of protein structures. Although most of these databases are available on the Internet, they are underutilized by biologists. The opposite is also true in that computational biologists does not always utilize all the data and expertise of experimental biologists.

Structural biology consists of two parts: the experimental part in which structures are determined and proteins are characterized and the computational part in which computers are used to analyze and interpret structures. Experimental biologists tend to shy away from the computational side, citing reasons such as the complexity of the programs and the vast amount of data that is available. In an effort to alleviate these problems, a web-based system known as FunGIMS was designed.

FunGIMS is a Functional Genomics Information Management System that consists of various modules, each specialized for a specific type of data, yet integrating the different data types in a transparent manner. FunGIMS currently consists of modules for Structure, Sequence, Genomics and Small molecules. This study focused on the Structural module, its design and the way in which it can help experimental biologists enrich and

guide their experiments to achieve more successful results. In the future the system may include more aspect to educate the users about the limitations inherent in the specific tools that they use.

FunGIMS was designed for ease of use by both programmers and biologists. During the design phase, it was decided to use the MVC architecture for FunGIMS, which allows for easy expansion of the program as well as addition of new programs and analysis methods. This type of architecture separates the display of data, control functions and data management into three separate sections, allowing for easy maintenance or upgrading of a section of FunGIMS. The design architecture was applied not only to the overall FunGIMS section but also to the more specific Structural module.

The main focus during the design of FunGIMS was not the programmers but the end users of the program. To alleviate the problems encountered by biologists when using structural biology programs, the interfaces were designed to be intuitive and easy to use. All the syntax and specific subtleties of running a program have been hidden from the user and only the basic information is required. A user can access this information by simply uploading files or using data from the databases already present in FunGIMS. The program is then run and the results presented to the user in a clean interface with the option to download or save the results. Security was also of concern as some users preferred to keep data private or share it with only a certain subset of users. To overcome this issue a system was created whereby users belong either to a single or multiple groups and every data entry belongs to a certain group. Public data are visible to all users and belong to a World group. Whenever a user saves data, he/she can decide to which group the data belongs and thus share it with the members of that group while preventing any other user from accessing it.

Easy access to data is important and this was well catered for in FunGIMS. It provides a search function that allows the user to search across all data or a selected subset with either keywords or a specific entry identifier. When searching, access rights to data entries are taken into consideration and a user will only be able to view results which he has access to.

The data are stored in a relational database that allows for the creation of complex

queries to return specific results. The database is populated by parsing public data from the PDB, MSD and GenBank as well as storing user-generated data. Links between the data are also generated to allow for better integration between the data types.

The Structural module caters exclusively for structural and protein data as well as the analysis of proteins. It provides access to all the known protein structure files in the PDB as well as the enhanced data from the MSD. This allows a user to explore the protein structure in detail while also presenting an interactive display of the protein in the browser. Jmol is used in this regard and allows the user to interact with the protein in a three dimensional environment inside his web-browser. Data such as the secondary structure composition, SCOP, Pfam and other relevant information are presented to the user in a clear and consistent format.

The Structural module also provides protein structure and sequence analysis tools. There are tools for predicting transmembrane helices (TMHMM), for predicting protein families on the basis of sequence (Hmmer search against Pfam) as well as searching for conserved motifs in a sequence (Prosite). In addition to the analysis tools there are also tools that allow the user to build homology models and generate scripts for molecular dynamics simulations. In the homology modelling section a user simply enters the basic required information and thereafter generate scripts for homology modelling using Modeller or WHAT IF. Homology models can also be built online using Modeller, with the user providing a protein sequence and a template PDB structure id as well as refinement levels. The Structural module then proceeds to do an automated alignment and model building using Modeller. The resulting model, alignment file and script file used are then supplied to the user to download or save in the system.

Due to the computationally intensive nature of molecular dynamics simulations, the Structural module only provides a script generation capability. Scripts can then be run on a local machine. The user simply enters the required information and can then select to generate a script for either CHARMM, NAMD or Yasara. Thereafter the script is prepared and the user can download it or save it in the system.

The functionality of the Structural module was used in three investigations on FMDV. The first objective was related to proteome differences between different serotypes of

FMDV. Using the tools in the Structural module, each proteome was analyzed for various features such as secondary structure, conserved motifs and hydrophobicity. The results were then compared on an individual protein level between the different serotypes. Various differences were found such as changes in the hydrophobicity patterns on proteins 2A and 3A. These changes may affect the way in which certain proteins associate with each other as well as with membranes such as the ER and hence may have an influence on replication rates.

The second hurdle encountered by the local researchers was related to differences in the replication rate and plaque morphology between the different serotypes. Experimental evidence pointed to variation in the 3C protease and 3D RNA polymerase proteins of FMDV. Using the Structural module, models of the 3C and 3D proteins were built and differences between various SAT serotypes were mapped to the structure. For 3C 51 SAT serotype sequences were used and for 3D 16 SAT serotype sequences were used. After the differences were mapped to the protein models, it was found that a region in 3C, previously believed to be invariant, contained 9 differences. When locating the differences on the protein model, it was found that although these differences did occur, the hydrogen bond network in the local area was preserved. This preservation allows 3C to accept these differences without a major change in the activity of the protein.

Previous studies showed that 3D contained four invariant regions. After mapping the differences to the structure it was found that three of the four invariant regions were also conserved in the SAT serotypes. However, one region showed some variation. When mapping these differences to the protein model, it was found that these differences did not affect the structure since the different amino acids involved all have the same physiochemical characteristics and size. These changes will also not have a major effect on the activity of the protein, but subtle differences may explain the differences seen in replication rate and plaque morphology.

A third problem faced by the researchers during FMDV vaccine design, was the stability of two FMDV SAT2 subtype capsids. There were five differences between the proteins making up the capsid, but during experiments it was seen that one capsid was consistently more stable at pH 6.0 than the other. To investigate this observation, the Structural

module was used to search for relevant structure and to construct homology models of the capsid protomers. Molecular dynamics simulation scripts for the Yasara program were also generated to investigate at which level of capsid assembly the difference had an effect. After building the models and running simulations of capsid protomers as well as capsid pentamer assemblies, it was found that there were no differences between the stability of the protomer and that of the pentamer. This prompted other avenues of investigation that resulted in performing pKa predictions of the residues predicted to be involved in the pentamer association interface. The pKa predictions showed that the pKa value of His145 on chain 1C, involved in interpentamer interactions (Ellard *et al.*, 1999), changed when a Val493Phe mutation occurred on chain 1C, structurally close approximation to His145. This resulted in a pKa shift of 0.5 units and thus made the ZIM/5/83/2 slightly more stable at pH 6.0 than ZIM/7/83/2.

The results obtained for FMDV allow researchers to understand the results reflected in their experimental work with regard to slight differences in FMDV replication rates. It also allows for a new understanding of the interaction between the different protein chains in the capsid as well as understanding the effect of seemingly innocuous differences in the amino acids sequence. In conclusion, these small differences in the capsid protein sequence affect pentamer-pentamer association and not the assembly of protomers or pentamers.

Introducing the local researchers to these tools, allowed them to become more comfortable with using structural biology tools and lead to the use of more advanced programs. Throughout the various chapters in this study, it was seen that structural biology plays a vital role in understanding the biological world. By providing easy access to structural data and analysis tools, biologists can now explore a new world that was previously considered to be a complex environment and so improve and guide future experimental work. This work expanded the knowledge of local researchers by providing new information about conserved patterns and features in local SAT strains, variation levels and effects in SAT 3C and 3D enzymes as well as providing new avenues for improving vaccine design based on viral capsid interaction analysis.