Expression, Purification, and Characterisation of the Alpha-Helical and Beta-sheet Domains of Rotavirus VP6

by

MILAAN STRACHAN

submitted in accordance with the requirements for the degree of

MASTER OF SCIENCE IN LIFE SCIENCES

in the

COLLEGE OF AGRICULTURE AND ENVIRONMENTAL SCIENCES
DEPARTMENT OF LIFE AND CONSUMER SCIENCE

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROFESSOR S GILDENHUYYS
CO-SUPERVISOR: MR T MASHAPA

August 2023
Declaration

I, Milaan Simone Strachan, hereby declare that the dissertation, with the title: “Expression, Purification, and Characterisation of the Alpha-Helical and Beta-sheet Domains of Rotavirus VP6” which I hereby submit for the degree of Master of Science in Life Sciences at the University of South Africa, is my own work and has not previously been submitted by me for a degree at this or any other institution.

I declare that the dissertation /thesis does not contain any written work presented by other persons whether written, pictures, graphs or data or any other information without acknowledging the source.

I declare that where words from a written source have been used the words have been paraphrased and referenced and where exact words from a source have been used the words have been placed inside quotation marks and referenced.

I declare that I have not copied and pasted any information from the Internet, without specifically acknowledging the source and have inserted appropriate references to these sources in the reference section of the dissertation or thesis.

I declare that during my study I adhered to the Research Ethics Policy of the University of South Africa, received ethics approval for the duration of my study prior to the commencement of data gathering, and have not acted outside the approval conditions.

I declare that the content of my dissertation/thesis has been submitted through an electronic plagiarism detection program before the final submission for examination.

Student signature: [Signature]

Date: 20 August 2023
Abstract
The capsid protein VP6 is of paramount importance to the stability and infectivity of Rotaviruses. Through interactions of VP6s’ beta-sheet (VP6β) and alpha-helical (VP6α) domains with the viral particle’s outer- and innermost layers, respectively, VP6 stabilises matured Rotaviruses and activates transcription of the viral genome upon cell entry. This study focused on the individual domains of Rotavirus VP6. The aim of the study was to probe the structure and stability of VP6β and VP6α when expressed independently of each other. The objectives of the study were: (1) optimise the bacterial expression of VP6β and VP6α through modulation of the expression conditions (2) solubilise and then purify the domains by immobilised metal chromatography (IMAC), (3) characterise by means of spectroscopy (mass spectroscopy, far-UV circular dichroism (CD), and intrinsic tryptophan fluorescence spectroscopy) and gel electrophoresis (native-PAGE), the primary, secondary, tertiary, and quaternary structures of the domains, (4) characterise the conformational stability by means of spectroscopy (far-UV CD and intrinsic tryptophan fluorescence) of VP6β and VP6α when thermally and chemically challenged, and (5) determine the melting temperature by differential scanning calorimetry (DSC). To this end, two Escherichia coli strains BL21(DE3) and NiCo21 (DE3) were transformed with pET15a plasmids containing the codon-optimised DNA consensus sequences of VP6β and VP6α. The expression of VP6β and VP6α was done at different temperatures (37°C and 20°C), inducer concentrations (1 × and 10 × IPTG), and post-induction incubation times (2 h – 7 h, and 16 h) in both E. coli strains and the outcomes were visualised by SDS-PAGE. All conditions tested produced the domains in an insoluble form and though expression levels appeared to be comparable between strains, the NiCo21 (DE3) was ultimately selected for further expression of the domains as expression could be induced with the lowest concentration of IPTG. The insoluble domains were subjected to a solubilisation study where the domains were frozen in various Tris-HCl buffers differing in pH (7 – 10) and urea concentration (0 M, 2 M, and 5 M) and thawed. The results of the solubilisation study showed that both domains could effectively be solubilized in 2 M urea, provided that the pH of the freezing buffer was at least one unit higher than the pI of the domain. The solubilised VP6β and VP6α were purified by nickel affinity chromatography in yields of 13.32 mg and 25 mg from 1 L of NiCo21 (DE3) culture, respectively and were confirmed by mass spectroscopy, far-UV CD, and intrinsic tryptophan fluorescence spectroscopy, to have native-like sequences and structural features. The quaternary analysis revealed that VP6β existed as a single
monomeric species in solution while VP6α formed different-sized structures in solution. The conformational stability of VP6β and VP6α was demonstrated as the domains had resisted structural changes up to 46°C and 50°C, respectively and the DSC analysis revealed melting points of 67.94°C for VP6β and 68.55°C for VP6α. The domains were noted to aggregate extensively which prevented the recovery of their native structures upon cooling. The chemical unfolding study was done in 1 M – 5 M guanidine hydrochloride (GdCl) and 1 M – 8 M urea, and revealed the chemical stability of the domains and their respective unfolding pathways. Approximately 1.5 M and 2.25 M GdCl were needed to denature 50% of VP6β and VP6α, respectively. Urea concentrations of 4.5 M (VP6β) and 4 M (VP6α) also resulted in a 50% loss of native structures. The observation of non-cooperative unfolding pathways that differed between the spectroscopic probes suggested a complex unfolding process involving the formation of one or more intermediates. Though native-like structures could be recovered upon denaturant removal, the refolding and unfolding pathways differed, which was indicative of irreversibility. Overall, VP6β and VP6α were easily producible and purifiable in quantities suitable for further studies. Further investigations could highlight the potential applications the domains could have in vaccine development and drug-delivery. Non-cooperative folding indicated the necessity of interactions between VP6β and VP6α in the full-length protein for cooperative folding.

**Keywords**

Research Outputs

1. Poster Presentation at SASBMB 2022
Strachan, MS, Mashapa, T, and Gildenhuys, S.

2. Manuscript Submitted to Heliyon
Spectroscopic Analysis of the Bacterially Expressed Head Domain of Rotavirus VP6.
Strachan, MS, Mashapa, T, and Gildenhuys, S

3. Manuscript in Progress
Expression, solubilisation, purification, and characterisation of the Rotavirus VP6 alpha-helical domain.
Strachan, MS, Mashapa, T, and Gildenhuys, S
Dedication

I dedicate this work to my parents, Warren and Belinda Strachan, whose unwavering support and encouragement illuminated my path as I walked towards my dream. Their belief in my potential, countless sacrifices, and boundless love have been the guiding light that fuelled my journey and made this achievement possible.

*Perseverantia omnia vincit.*
*Perseverance conquers all.*
Acknowledgements

I am deeply grateful to:

My Family - To my dad Warren, mom Belinda, sister Chinique, brother Caleb, and nephew Gabriel, your unwavering support has been the bedrock of my journey. Your love and encouragement, often silent but profound, have carried me through the toughest times.

Natalie Vymetal and Tshepo Sekele - My lab mates and mentors. Your guidance and willingness to share knowledge and experiences have been invaluable.

Ms. Thato Manyaapelo - Your tireless dedication in the Biotechnology lab has set a remarkable example for all of us. Your support and encouragement have been instrumental in my growth as a scientist.

Ms. Lesego Masango - Your kindness and support, extending far beyond academia, have been a constant source of strength. Your belief in me has been a driving force.

Marlanka Prinsloo - My dear friend, your unwavering support and encouragement have been a comforting presence in my life. Your belief in my abilities has been a constant motivator.

Raylene Human - Your friendship and encouragement have been a ray of sunshine on cloudy days. Your unwavering support has meant the world to me.

Mr. Tshepo Mashapa - My co-supervisor, your patience and understanding in training me have been remarkable. Your contributions to my learning journey have shaped me as a scientist.

Professor Samantha Gildenhuys - My supervisor, your guidance and support have been the pillars of my academic growth. Your mentorship and the opportunity to work in your team have been transformative.

The National Research Foundation (NRF) - I extend my gratitude to NRF for their financial support, which has made my academic pursuit possible.
# Table of Contents

Declaration.......................................................................................................................... ii

Abstract.................................................................................................................................. iii

Keywords ................................................................................................................................ iv

Research Outputs ................................................................................................................ v

Dedication .............................................................................................................................. vi

Acknowledgements ............................................................................................................... vii

List of Figures ....................................................................................................................... xii

List of Abbreviations .......................................................................................................... xiv

List of Buffers ....................................................................................................................... xv

Chapter 1 Introduction.......................................................................................................... 1

1.1 Diarrhoeal Disease ......................................................................................................... 1

1.2 Global Burden of Rotavirus Infection ........................................................................... 1

1.3. Rotaviruses .................................................................................................................... 1

1.3.1 Rotavirus Species ........................................................................................................ 1

1.3.2 Rotavirus A Strains ..................................................................................................... 2

1.3.3 Rotavirus A Structure ................................................................................................. 2

1.4 Pathophysiology and Replication Cycle ....................................................................... 4

1.4.1 Pathophysiology of Rotavirus Infections ................................................................ 4

1.4.2 Rotavirus Replication Cycle ....................................................................................... 5

1.5 Rotavirus Vaccines ......................................................................................................... 7

1.6 Rotavirus Capsid Protein VP6..................................................................................... 8

1.6.1 A General Introduction to Protein Structure and Protein Domains ....................... 8

1.6.1.1 Protein Structure ..................................................................................................... 8

1.6.1.2 Protein Domains ................................................................................................... 10

1.6.2 VP6 and its Domains ................................................................................................. 12
1.6.2.1 Structure of VP6 ........................................................................................................ 12
1.6.2.2 The Beta-Sheet Domain........................................................................................ 12
1.6.2.3 The Alpha-Helical Domain.................................................................................. 14

1.7 Bacterial Protein Expression and the Formation of Inclusion Bodies .............. 15

1.8 Solubilisation of Inclusion Bodies.............................................................................. 16

1.9 Previous Expression and Purification of Rotavirus VP6 ..................................... 17

1.10 Probing the Structure and Conformational Stability of Proteins ..................... 19

1.11 Differential Scanning Calorimetry........................................................................... 24

1.12 Aim and Objectives.................................................................................................... 24

Chapter 2 Methodology................................................................................................. 25

2.1 Materials ....................................................................................................................... 25

2.2 Bacterial Transformation.............................................................................................. 25

2.3 Glycerol Stock Preparation.......................................................................................... 25

2.4 Protein Expression Studies.......................................................................................... 26

2.5 Isolation and Solubilisation of Inclusion Bodies ....................................................... 26

2.6 Purification ................................................................................................................... 27

2.7 Protein Gel Electrophoresis ......................................................................................... 28

2.7.1 SDS-PAGE............................................................................................................... 28

2.7.2 Native-PAGE............................................................................................................ 29

2.8 Spectrophotometry ....................................................................................................... 30

2.8.1 Spectrophotometric Determination of Protein Concentration ......................... 30

2.8.1.1 Absorbance Spectrometry.................................................................................. 30

2.8.1.2 Bradford Assay................................................................................................... 30

2.8.2 Mass Spectroscopy .................................................................................................. 31

2.8.3 Circular Dichroism .................................................................................................. 31

2.8.4 Intrinsic Tryptophan Fluorescence Spectroscopy ................................................. 31

2.8.5 Thermal Stability Studies......................................................................................... 32

2.8.6 Chemical Stability Studies ...................................................................................... 32

2.8.6.1 Normalising Fluorescence Spectra .................................................................... 32
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.8.6.2 Conversion to Fraction Unfolded</td>
<td>33</td>
</tr>
<tr>
<td>2.9 Differential Scanning Calorimetry</td>
<td>33</td>
</tr>
<tr>
<td>2.10 AlphaFold</td>
<td>33</td>
</tr>
<tr>
<td>2.11 Molecular Dynamics Simulation</td>
<td>34</td>
</tr>
<tr>
<td>2.12 Software and Online Tools</td>
<td>34</td>
</tr>
<tr>
<td>Chapter 3 Results</td>
<td>35</td>
</tr>
<tr>
<td>3.1 The Beta-Sheet Domain</td>
<td>35</td>
</tr>
<tr>
<td>3.1.1 Bacterial Growth Curves and Recombinant Protein Expression</td>
<td>35</td>
</tr>
<tr>
<td>3.1.1.1 Bacterial Growth Curves</td>
<td>35</td>
</tr>
<tr>
<td>3.1.1.2 Recombinant Protein Expression</td>
<td>35</td>
</tr>
<tr>
<td>3.1.2 Solubilisation Study</td>
<td>40</td>
</tr>
<tr>
<td>3.1.3 Purification</td>
<td>40</td>
</tr>
<tr>
<td>3.1.4 Characterising the Structure of VP6(\beta)</td>
<td>43</td>
</tr>
<tr>
<td>3.1.4.1 Primary Structure</td>
<td>43</td>
</tr>
<tr>
<td>3.1.4.2 Secondary Structure</td>
<td>46</td>
</tr>
<tr>
<td>3.1.4.3 Tertiary Structure</td>
<td>46</td>
</tr>
<tr>
<td>3.1.4.4 Quaternary Structure</td>
<td>51</td>
</tr>
<tr>
<td>3.1.5 Characterising the Conformational Stability of VP6(\beta)</td>
<td>53</td>
</tr>
<tr>
<td>3.1.5.1 Chemical Conformational Stability</td>
<td>53</td>
</tr>
<tr>
<td>3.1.5.2 Thermal Conformational Stability</td>
<td>58</td>
</tr>
<tr>
<td>3.1.5.3 Differential Scanning Calorimetry</td>
<td>64</td>
</tr>
<tr>
<td>3.2 Alpha-Helical Domain</td>
<td>66</td>
</tr>
<tr>
<td>3.2.1 Bacterial Growth Curves and Recombinant Protein Expression</td>
<td>66</td>
</tr>
<tr>
<td>3.2.1.1 Bacterial Growth Curves</td>
<td>66</td>
</tr>
<tr>
<td>3.2.1.2 Recombinant Protein Expression</td>
<td>66</td>
</tr>
<tr>
<td>3.2.2 Solubilisation Study</td>
<td>71</td>
</tr>
<tr>
<td>3.2.3 Purification</td>
<td>73</td>
</tr>
<tr>
<td>3.2.4 Characterising the Structure of VP6(\alpha)</td>
<td>73</td>
</tr>
<tr>
<td>3.2.4.1 Primary Structure</td>
<td>73</td>
</tr>
<tr>
<td>3.2.4.2 Secondary Structure</td>
<td>77</td>
</tr>
<tr>
<td>3.2.4.3 Tertiary Structure</td>
<td>79</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Structure of the <em>Rotavirus</em>.</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td><em>Rotavirus</em> Replication Cycle.</td>
<td>6</td>
</tr>
<tr>
<td>1.3</td>
<td>Structures of the VP6 Trimer and Monomer.</td>
<td>13</td>
</tr>
<tr>
<td>1.4</td>
<td>Location of Tryptophan in VP6.</td>
<td>23</td>
</tr>
<tr>
<td>3.1</td>
<td>Bacterial Growth Curves.</td>
<td>36</td>
</tr>
<tr>
<td>3.2</td>
<td>Bacterial Expression of VP6__β.</td>
<td>37</td>
</tr>
<tr>
<td>3.3</td>
<td>Uninduced Cultures.</td>
<td>38</td>
</tr>
<tr>
<td>3.4</td>
<td>Solubilisation of VP6__β.</td>
<td>41</td>
</tr>
<tr>
<td>3.5</td>
<td>Purification of VP6__β.</td>
<td>42</td>
</tr>
<tr>
<td>3.6</td>
<td>Quantification of Purified VP6__β.</td>
<td>44</td>
</tr>
<tr>
<td>3.7</td>
<td>VP6__β Peptides Detected by Mass Spectroscopy.</td>
<td>45</td>
</tr>
<tr>
<td>3.8</td>
<td>Far-UV CD Spectra of VP6__β.</td>
<td>47</td>
</tr>
<tr>
<td>3.9</td>
<td>VP6__β Tertiary Structure.</td>
<td>49</td>
</tr>
<tr>
<td>3.10</td>
<td>Molecular Dynamics and Scatter.</td>
<td>50</td>
</tr>
<tr>
<td>3.11</td>
<td>Native-PAGE of VP6__β.</td>
<td>52</td>
</tr>
<tr>
<td>3.12</td>
<td>VP6__β Unfolding with Urea.</td>
<td>54</td>
</tr>
<tr>
<td>3.13</td>
<td>VP6__β Unfolding with Guanidine Hydrochloride.</td>
<td>56</td>
</tr>
<tr>
<td>3.14</td>
<td>VP6__β Thermal Studies in 0 M Urea.</td>
<td>59</td>
</tr>
<tr>
<td>3.15</td>
<td>VP6__β Thermal Studies in 2 M Urea.</td>
<td>62</td>
</tr>
<tr>
<td>3.16</td>
<td>DSC Thermogram of VP6__β in 0 M urea.</td>
<td>65</td>
</tr>
<tr>
<td>3.17</td>
<td>Bacterial Growth Curves.</td>
<td>67</td>
</tr>
<tr>
<td>3.18</td>
<td>Bacterial Expression of VP6__α in BL21 (DE3).</td>
<td>68</td>
</tr>
<tr>
<td>3.19</td>
<td>Bacterial Expression of VP6__α in NiCo21 (DE3).</td>
<td>69</td>
</tr>
<tr>
<td>3.20</td>
<td>Bacterial Expression of VP6__α at 20°C.</td>
<td>70</td>
</tr>
<tr>
<td>3.21</td>
<td>Solubilisation of VP6__α.</td>
<td>72</td>
</tr>
<tr>
<td>3.22</td>
<td>Purification of VP6__α.</td>
<td>74</td>
</tr>
<tr>
<td>3.23</td>
<td>Quantification of Purified VP6__α.</td>
<td>75</td>
</tr>
<tr>
<td>3.24</td>
<td>VP6__α Peptides Detected by Mass Spectroscopy.</td>
<td>76</td>
</tr>
<tr>
<td>3.25</td>
<td>Far-UV CD Spectra of VP6__α.</td>
<td>78</td>
</tr>
<tr>
<td>Figure 3.26</td>
<td>VP6α Tertiary Structure.</td>
<td>80</td>
</tr>
<tr>
<td>Figure 3.27</td>
<td>VP6α Quaternary Structure.</td>
<td>82</td>
</tr>
<tr>
<td>Figure 3.28</td>
<td>VP6α Unfolding with Urea.</td>
<td>83</td>
</tr>
<tr>
<td>Figure 3.29</td>
<td>VP6α Unfolding with Guanidine Hydrochloride.</td>
<td>86</td>
</tr>
<tr>
<td>Figure 3.30</td>
<td>VP6α Thermal Studies in 0 M Urea.</td>
<td>89</td>
</tr>
<tr>
<td>Figure 3.31</td>
<td>VP6α Thermal Studies in 2 M Urea.</td>
<td>92</td>
</tr>
<tr>
<td>Figure 3.32</td>
<td>DSC Thermogram of VP6α in 0 M Urea.</td>
<td>95</td>
</tr>
</tbody>
</table>
List of Abbreviations

AmpR  Ampicillin resistant  
CD    Circular dichroism  
DLP   Double-layered Particle  
dsRNA Double-stranded ribonucleic acid  
E. coli Escherichia coli  
EGFP  Epidermal growth factor protein  
ENS   Enteric nervous system  
HBGA  Histo-blood group antigen  
HIC   Hydrophobic interaction chromatography  
IEC   Ion-exchange chromatography  
IMAC  Immobilised-metal affinity chromatography  
IPTG  Isopropyl β-D-thiogalactopyranoside  
LAV   Live attenuated vaccine  
LB    Luria-Bertani  
NSP   Non-structural protein  
PDB   Protein Data Bank  
RdRp  Ribonucleic acid-dependent ribonucleic acid polymerase  
rER   Rough endoplasmic reticulum  
RNAP  Ribonucleic acid polymerase  
RV    Rotavirus  
SDS-PAGE Sodium-Dodecyl Sulphate Polyacrylamide Gel Electrophoresis  
SEC   Size exclusion chromatography  
ssRNA Single-stranded ribonucleic acid  
TLP   Triple-layered particle  
Tris-HCl Tris(hydroxymethyl) aminomethane hydrochloride  
VP    Viral Protein  
VP6α  VP6 alpha-helical domain  
VP6β  VP6 beta-sheet domain  

The IUPAC-IUBMB three and one-letter codes for the amino acids were used (Recommendations, 1983).
### List of Buffers

<table>
<thead>
<tr>
<th>Buffer Name</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wash Buffer A</td>
<td>100 mM Tris-HCl pH 7.00 with 1% Triton X-100</td>
</tr>
<tr>
<td>Wash Buffer B</td>
<td>100 mM Tris-HCl pH 7.00</td>
</tr>
<tr>
<td>Freezing Buffer A</td>
<td>100 mM Tris-HCl pH 9.00 with 2 M Urea</td>
</tr>
<tr>
<td>Freezing Buffer B</td>
<td>100 mM Tris-HCl pH 7.40 with 2 M Urea</td>
</tr>
<tr>
<td>Equilibration Buffer A</td>
<td>100 mM Tris-HCl pH 9.00, 2 M urea, 150 mM NaCl, 40 mM imidazole, and 0.02% (w/v) sodium azide.</td>
</tr>
<tr>
<td>Column Wash Buffer A</td>
<td>100 mM Tris-HCl pH 9.00, 2 M urea, 300 mM NaCl, 40 mM imidazole, and 0.02% (w/v) sodium azide.</td>
</tr>
<tr>
<td>Elution Buffer A</td>
<td>100 mM Tris-HCl pH 9.00, 2 M urea, 300 mM NaCl, 600 mM imidazole, and 0.02% (w/v) sodium azide.</td>
</tr>
<tr>
<td>Equilibration Buffer B</td>
<td>100 mM Tris-HCl pH 7.40, 2 M urea, 150 mM NaCl, 40 mM imidazole, and 0.02% (w/v) sodium azide.</td>
</tr>
<tr>
<td>Wash Buffer B</td>
<td>100 mM Tris-HCl pH 7.40, 2 M urea, 300 mM NaCl, 40 mM imidazole, and 0.02% (w/v) sodium azide.</td>
</tr>
<tr>
<td>Elution Buffer B</td>
<td>100 mM Tris-HCl pH 7.40, 2 M urea, 300 mM NaCl, 500 mM imidazole, and 0.02% (w/v) sodium azide.</td>
</tr>
<tr>
<td>Dialysis Buffer 1</td>
<td>20 mM Sodium phosphate (dibasic-monobasic) pH 7.40 with 2 M urea and 0.02% (w/v) sodium azide.</td>
</tr>
<tr>
<td>Dialysis Buffer 2</td>
<td>20 mM Sodium phosphate (dibasic-monobasic) pH 7.40 and 0.02% (w/v) sodium azide.</td>
</tr>
<tr>
<td>SDS-PAGE Running Buffer</td>
<td>25 mM Tris pH 8.3, 7.2% (w/v) glycine, and 0.5% (w/v) SDS</td>
</tr>
<tr>
<td>Native-PAGE Running Buffer</td>
<td>25 mM Tris pH 8.3 with 192 mM glycine</td>
</tr>
<tr>
<td>SDS-PAGE Sample Buffer</td>
<td>62.5 mM Tris-HCl pH 6.8, 10% (v/v) glycerol, 2% (v/v) SDS, 5% (v/v) β-mercaptoethanol, and 0.05% (w/v) bromophenol blue</td>
</tr>
<tr>
<td>Native-PAGE Sample Buffer</td>
<td>62.5 mM Tris-HCl pH 6.8, 40% (v/v) glycerol, 0.05% Coomassie Brilliant Blue R250, and 0.05% (w/v) bromophenol blue</td>
</tr>
<tr>
<td>VP6β</td>
<td>62.6 mM Tris-HCl pH 6.8, 40% (v/v) glycerol, and 0.05% (w/v) bromophenol blue</td>
</tr>
</tbody>
</table>
Chapter 1 Introduction

1.1 Diarrhoeal Disease

Diarrhoeal disease is a significant global health issue, it ranks as the second leading cause of death among young children and claims the lives of approximately 525000 children annually (World Health Organization, 2023). Diarrhoea is the passing of loose or liquid stools at volumes and frequencies that are higher than usual for an individual (Powell, 1995; Fine, Krejs & Fordtran, 1998; World Health Organization, 2017). Diarrhoea usually results from intestinal infections caused by various bacteria, viruses, or protozoans that are introduced to the body following the consumption of spoiled food, faecal-contaminated water, or poor hygiene practices (Andra-Michel & Giannella, 1999). Dehydration is the most severe consequence associated with diarrhoeal diseases (Butler et al., 1987; Zodpey et al., 1999; van der Westhuizen et al., 2019). It is defined as the irrecoverable loss of water and essential electrolytes through vomiting, passing liquid stool, sweating, and urination (Paediatrics & Child Health, 2003; World Health Organization, 2023). Diarrhoeal disease, though a burden on a global scale, is most prevalent in low-to-middle-income countries (Walker-Fischer et al., 2012; World Health Organization, 2023; The Lancet, 2020). In low-to-middle-income countries, infections leading to diarrhoea are commonly caused by Rotavirus (RV) (Tate et al., 2016; Crawford et al., 2017).

1.2 Global Burden of Rotavirus Infection

Rotaviruses cause intestinal infections that result in diarrhoea in children aged five and younger, with children aged between six months to two years being most susceptible to RV infections (Bishop et al., 1973; Crawford et al., 2017; World Health Organization, 2020). It has been reported that RV infections were responsible for approximately 200000 deaths in 2013 and this number had reduced to 129000 in 2019 (Du et al., 2022; Tate et al., 2016; Parashar et al., 2003; Crawford et al., 2017; World Health Organization, 2020). An overwhelming majority of RV morbidity and mortality occurs in developing countries in Africa and South-East Asia (Crawford et al., 2017; World Health Organization, 2020).

1.3. Rotaviruses

1.3.1 Rotavirus Species

Rotaviruses, named for their wheel-like structure, belong to the family Reoviridae, a group of non-enveloped viruses that house segmented double-stranded RNA (dsRNA) genomes within
icosahedral capsids (Bishop et al., 1973; Mathieu et al., 2001; Desselberger et al., 2009; Desselberger, 2014; Afchangi et al., 2019). Rotaviruses are grouped into nine species (A – I, and J) based on the antigenicity of the capsid protein VP6 (Mathieu et al., 2001; Matthijnssens et al., 2008; Desselberger et al., 2009; Desselberger, 2014; Afchangi et al., 2019). Some species of RV may preferentially infect a particular host, for example, Rotavirus J, was observed predominantly in bats (Bányai et al., 2017). Other species, such as Rotavirus A, cause infections in a range of hosts including humans and the young of simian, bovine, murine, canine, feline, and avian species (Connor & Ramig, 1996; Estes, 2001; Ciarlet et al., 2002).

1.3.2 Rotavirus A Strains
Rotavirus A, the species of RV that humans are primarily infected by, possesses eleven segments of dsRNA enclosed in a triple-layered icosahedral capsid (Mathieu et al., 2001; Desselberger et al., 2009; Desselberger, 2014; Afchangi et al., 2019). These segments encode six structural proteins (VP1 – 4 and VP6-7) and six non-structural proteins (NSP 1-6) (Desselberger, 2014; Asensio-Cob et al., 2023). While non-structural proteins contribute to pathogenicity and replication (Hu et al., 2012), the six structural proteins collectively form the triple-layered capsid of RV particles (Crawford et al., 1994). Rotavirus A is further divided into serotypes or strains which are defined by the capsid protein VP7 and the spike protein VP4 which demonstrate significant variability (O’Ryan, 2009; Desselberger et al., 2009; Aoki et al., 2011; Patton 2012; Desselberger, 2014). The G serotype is determined by the RV capsid glycoprotein VP7 and P serotype is determined by the protease-sensitive RV spike protein VP4 (Patton 2012). Due to the fragmented nature of the RV genome, the genes encoding VP7 and VP4 can segregate independently and consequently result in RV strains with various P and G combinations (Patton 2012; Hoxie and Dennehy, 2021). There are 36 G types and 51 P types in Rotavirus A, however, over 90% of human RV infections are caused by the following six genotypes: G1P[8], G2P[4], G3P[8], G4P[8], G9P[8] and G12P[8] (Bányai et al., 2012; Matthijnssens et al., 2012; Dóró et al., 2014; Rakau et al., 2021).

1.3.3 Rotavirus A Structure
The Rotavirus A viral capsid is formed by three concentric layers of proteins that encapsulate the viral dsRNA genome (Figure 1.1). The outer layer of the viral capsid is a smooth coat formed by 260 trimers of the glycoprotein VP7 (Shaw et al., 1993; Yaeger et al., 1994; Ludert et al., 2002).
Figure 1.1: Structure of the Rotavirus.

Diagram showing the structure of a mature Rotavirus particle. Indicated are the structural proteins that form the outer, intermediate, and innermost layers of the viral capsid, making up what is referred to as the triple-layer particle (TLP). This diagram was constructed using biorender.com based on information from Mathieu et al. (2001), Desselberger et al. (2009), Aoki et al. (2011), and Desselberger (2014).
Embedded in the VP7 layer are 60 spikes of VP4, a protease-sensitive protein comprising two domains namely VP5 and VP8 (Shaw et al., 1993; Yaeger et al., 1994; Mertens et al., 2000; Ludert et al., 2002). The intermediate layer of the capsid consists of 260 homotrimeric VP6 molecules (Desselberger et al., 2009; Desselberger, 2014; Long & McDonald, 2017). The capsid protein VP6 is important in maintaining the organisation of virion and ensuring the structural integrity of the double-layered particle (DLP). The innermost layer of the capsid is formed by 120 VP2 decamers (Ludert et al., 2002). The VP2 layer encapsulates single-stranded RNA (ssRNA) and RNA processing enzymes VP1 and VP3. The single-ssRNA serves as a template for synthesising the dsRNA genome and mRNA transcripts (Desselberger, 2014). Structural protein VP1 functions as an RNA-dependent RNA polymerase (RdRp) responsible for synthesizing the dsRNA genome and mRNA transcripts (Desselberger, 2014). Additionally, VP3 caps mRNA transcripts at their 5' end, safeguarding them from degradation by host ribonucleases (Desselberger, 2014; Chanfreau, 2017).

1.4 Pathophysiology and Replication Cycle

1.4.1 Pathophysiology of Rotavirus Infections

Rotavirus particles enter the body when food or water containing faecal matter is ingested. This virus specifically targets the enterocytes, which are specialized cells responsible for digestion and nutrient absorption, located on the villi in the small intestine (Lundgren & Svensson, 2001). The entry of human RV strains into the target cell is mediated by the interaction of RV spike protein VP4, with the histo-blood group antigens (HBGAs) on the surface of the enterocytes (Erk et al., 2003; Aoki et al., 2011; Desselberger, 2014; Shanker et al., 2017). Upon entry, the outermost layer of the TLP is lost and the resulting DLP is released into the cytoplasm. The DLP is transcriptionally active, meaning that it synthesizes and releases mRNA into the cytoplasm to be translated by the translational machinery of the host cell (Crawford et al., 2017). The newly synthesised RV proteins have various functions in the replication cycle of the virus, however, there is evidence that implicates NSP4 in the pathogenicity of Rotaviruses. It has been noted that NSP4 stimulates the release of calcium ions from the endoplasmic reticulum (Ramig, 2004). The increased calcium concentration in the cytoplasm has a number of effects including a decrease in the expression of cell surface digestive enzymes, inhibition of absorptive pathways, disruption of the enterocyte cytoskeleton, induction of necrotic pathways, and stimulation of the enteric nervous system (ENS) (Ball et al., 1996; Crawford et al., 2017).
These effects result in the destruction of the villus enterocytes, decreased digestive capabilities, malabsorption, and increased secretion in the epithelial cells at the base of the villus (Rao & Wang, 2010; Crawford et al., 2017). The watery diarrhoea associated with RV infections is a result of the unabsorbed nutrients in the intestine and the stimulation of the ENS (Ball et al., 1996; Crawford et al., 2017). Rotavirus-infected enterocytes cannot optimally absorb nutrients from the chyme (a paste formed from ingested food following digestion in the stomach that is rich in nutrients) as it moves through the small intestine (Ramig, 2004; Hsu et al., 2020). This means that there is a higher concentration of nutrients in the intestine than there is inside the villi. This causes an osmotic gradient that stimulates the underlying mucosa to secrete water (Ramig, 2004). In addition to this, the virus also stimulates secretion by activating the ENS thereby resulting in the release of large volumes of water during defecation (Ramig, 2004; Crawford et al., 2017).

1.4.2 Rotavirus Replication Cycle

The RV replication cycle begins when the triple-layered particle (TLP) attaches to the surface of the target cell (Figure 1.2). Viral attachment occurs when the VP8 domain of the VP4 spike protein binds to the histo-blood group antigens (HGBAs) on the surface of the enterocyte (Ludert et al., 2002; Erk et al., 2003; Jayaram et al., 2004; Aoki et al., 2011; Desselberger, 2014). The interaction of VP4 with the membrane receptors allows the virus to enter the enterocytes by receptor-mediated endocytosis (Jayaram et al., 2004; Aoki et al., 2011; Desselberger, 2014). As the TLP enters the cell, it is enclosed within a vesicle with a low concentration of calcium ions. This low calcium concentration causes the vesicle to become permeable, resulting in the dissociation of the TLP’s outermost layer and releasing the double layered particle (DLP) into the cytoplasm (Jayaram et al., 2004; Aoki et al., 2011; Desselberger, 2014). The loss of the outermost layer of the capsid results in a conformational change in the underlying VP6 trimers. This conformational change activates the transcriptional activity of VP1 in the viral core (Desselberger, 2014). The transcriptionally active DLP then releases mRNA into the cytoplasm where translation occurs (Jayaram et al., 2004). In the infected cell, the translation of host mRNA is repressed by NSP3, and this results in the upregulation of viral protein synthesis (Suguna & Rao, 2010; Hu et al., 2012).
Figure 1.2: Rotavirus Replication Cycle.

The Rotavirus replication cycle includes processes such as virion attachment to host cell receptors and endocytosis, uncoating and release of the DLP, transcription and translation of viral mRNA, RNA synthesis and viral assembly. This diagram was constructed using biorender.com based on information from Ludert et al., 2002; Erk et al., 2003; Jayaram et al., 2004; Aoki et al., 2011; Desselberger, 2014; and Crawford et al., 2017.
With the exception of VP4, VP7, and NSP4 that accumulate in the membrane of the rough endoplasmic reticulum (rER), the newly synthesised structural and non-structural proteins concentrate in a membrane-less structure in the cytoplasm near rER known as a viroplasm (Aoki et al., 2011; Desselberger, 2014; Papa et al., 2021). The viroplasm requires NSP2 and NSP5 to form and is the site of dsRNA synthesis and viral assembly (Fabbretti et al., 1999). The VP1/3 complex is associated with a single-stranded sense RNA and the first step of the assembly process is when this RNA-VP1/3 complex is encapsulated by VP2 to produce the viral core (Desselberger, 2014; Long & McDonald, 2017). This is a necessary step because host RNases recognise, and rapidly degrade dsRNA. Therefore, dsRNA synthesis does not occur prior to viral core formation (Desselberger, 2014). Once the core is formed, dsRNA is synthesised using the ssRNA associated with the VP1/3 complex as a template (Desselberger, 2014). The core is encapsulated by VP6 trimers and the double-layered viral particle is formed (Desselberger, 2014; Long & McDonald, 2017). The DLP enters the rER by endocytosis mediated by the interaction of VP6 with the NSP4/VP4 complex in the membrane of the rER and this marks the beginning of TLP assembly (Desselberger, 2014). The maturation of the TLP, that is, the attachment of VP4 and VP7 to the DLP, occurs inside rER. It is not clear how newly formed TLPs exit the host cell, but it is proposed to occur either by the lytic pathway (that is, the release of mature viral particles following lysis of the host cell) or non-classical vesicular transportation (Crawford et al., 2017).

1.5 Rotavirus Vaccines
Vaccines are defined as biological preparations that provide acquired immunity to an infectious disease (Centers for Disease Control & Prevention, 2021). At present, there are two live attenuated vaccines available globally for RV infections. A live attenuated vaccine (LAV) is a vaccine that contains a viable pathogen that has been altered such that its virulence is significantly reduced (Plotkin, 2009). When an LAV is administered, it stimulates an immune response similar to that of the unattenuated pathogen without causing severe disease (Plotkin, 2009). Vaccines that are currently available are RotaRix™ and RotaTeq™. RotaRix™ is a monovalent vaccine that is administered twice (at ages two months and at four months) whereas Rotateq™ is a pentavalent vaccine that is administered thrice (at ages two, four, and six months) (Burnett et al., 2016; World Health Organization, 2006).
These vaccines were introduced in 2006 after the first RV vaccine Rotashield, a live, attenuated rhesus rotavirus based tetravalent vaccine, had to be withdrawn from the market as it was linked to the onset of intussusception, an unusual event where the intestine folds into itself (Cale & Klein, 2002). RotaRix™ and RotaTeq™ have not been associated with any adverse side effects to date. Despite the introduction of these vaccines, RV infections remain the leading cause of death of young children in developing countries (Crawford et al., 2017; Steele et al., 2016). Indeed, the vaccines on the market at present demonstrate low efficacies (50% – 60%) in developing countries (Burnett et al., 2016). This was found to be a peculiarity since the vaccines are highly effective (79% - 100%) in developed countries (Burnett et al., 2016). Though the exact reason for this observation is not known, genetic, microbiological, and socio-economic factors have been considered as possible explanations for the reduced efficacy in developing countries (Desselberger, 2017; ROTA Council, 2017). In China and Vietnam, Rotavin-M1 and LLR vaccines have reached the market, however, there is insufficient data regarding the safety of these vaccines and the impact of their use in these countries (Vetter et al., 2022). Two other Rotavirus vaccines, Rotavac and Rotasiil, have been investigated in animal models in India and the results thus far have been promising (Vetter et al., 2022). Since presently available vaccines do not demonstrate optimal efficiencies in developing countries there is a need for the development of inactivated vaccines (vaccines made from non-viable pathogens) or subunit vaccines (vaccines made from the immunogenic components of a pathogen) (Ward and McNeal, 2010; Li et al., 2014). Due to its abundance in the viral capsid and its ability to elicit an immune response, the RV structural protein VP6 has long been an attractive candidate for the development of novel RV vaccines.

1.6 Rotavirus Capsid Protein VP6

1.6.1 A General Introduction to Protein Structure and Protein Domains

1.6.1.1 Protein Structure

Proteins are biological macromolecules that have a myriad of functions in the cell including signal transduction, structural support, maintaining homeostasis, enzymatic catalysis, and transporting nutrients. This section provides an overview of the protein structural hierarchy. The hierarchy of protein structure refers to the conformation of proteins at increasing levels of complexity. The hierarchy includes four structural levels namely: primary, secondary, tertiary, and quaternary structures.
The primary structure of a protein is simply the sequence of its amino acids as determined by the genetic information obtained from the genome (LePelusa & Kaushik, 2022). This structure is linear and mainly stabilised by peptide bonds (LePelusa & Kaushik, 2022). The amino acid sequence may be the simplest structure but, it contains powerful information that determines the three-dimensional structure and function of the protein (Anfinsen, 1973). The distinct physiochemical properties of the different amino acids also determine the physical and chemical properties of the protein such as molecular weight, solubility, and reactivity (Katchalski-Katzir et al., 2006). The secondary structure of a protein refers to the local folding of the peptide backbone (Rehman et al., 2022). There are two types of secondary structures that are commonly seen in proteins namely, alpha-helices and beta-sheets. Alpha helices are stabilized by hydrogen bonds between the backbone amide and carbonyl groups (Brandt, 2015). Beta-sheets are formed when at least two segments of a polypeptide chain align and form hydrogen bonds between them (Cheng et al., 2013). When the individual strands are oriented parallel to each other it means that the N- and C-termini of both segments are on the same side whereas the anti-parallel conformation occurs when the N-terminus of one segment is on the same side as the C-terminus of the other segment (Cheng et al., 2013). The hydrogen bonds form at an angle in the parallel conformation and perpendicularly in the anti-parallel conformation (Cheng et al., 2013). Both alpha helices and beta-sheets are key for the protein to assume its correct three-dimensional structure. The tertiary structure of a protein refers to the three-dimensional conformation adopted by a single polypeptide chain (Engelking, 2015). This structure is formed when the secondary structures fold into more complex conformations and there are a variety of molecular interactions that stabilise this conformation including side chain interactions, electrostatic interactions, and hydrophobic interactions (Engelking, 2015). A major driving force in the folding of globular proteins is the hydrophobic effect. The hydrophobic effect describes the tendency of nonpolar molecules to avoid contact with water molecules in their local environment (Camillioni et al., 2016). When proteins fold, the amino acids with nonpolar or hydrophobic side chains are buried in the core of the protein while the hydrophilic amino acids are exposed to the aqueous environment (Camillioni et al., 2016). At the tertiary level of structure, most proteins are considered functional. In some cases, proteins have to associate with one or more (identical or different) proteins to become functional. The quaternary structure of a protein refers to the association of two or more proteins into a larger protein complex (Alberts et al., 2002).
This higher-order structure is stabilised by interactions between the individual subunits such as hydrogen bonding, electrostatic interactions, and hydrophobic interactions (Alberts et al., 2002). The quaternary structure is the most complex conformation proteins can assume. As not all proteins are functional at the tertiary level of structure, the quaternary structures allow multiple protein subunits to associate into larger structures to gain function or to be multifunctional. For example, the *Rotavirus* VP6 needs to trimerize in order to form and maintain the structural integrity of the viral particle and to activate processes such as the transcription of viral mRNA (Crawford et al., 2017).

### 1.6.1.2 Protein Domains

In the early 1940s, Beadle and Tatum hypothesized that the ratio of genes to proteins was 1:1, meaning that each gene was responsible for the synthesis and regulation of a single protein (Ponomarenko et al., 2016). The human genome was found to comprise ~ 20000 genes therefore if the one gene one protein hypothesis is applied, it would mean that there are roughly 20000 unmodified proteins in humans (Ponomarenko et al., 2016). However, this number does not correlate with the human proteome which was estimated to contain 10000 proteins in 2003 and this number increased significantly to several billion in 2013 (Smith & Kelleher, 2013; Ponomarenko et al., 2016). The latter approximation of the proteome accounts for the variety of proteoforms, brought about by events such as (1) post-translational modifications, (2) single nucleotide polymorphisms and the single amino acid polymorphisms they may give rise to, (3) alternative splicing, and (4) the presence of multiple domains (Karlsson et al., 2012; Roth et al., 2005; Smith & Kelleher 2013). This section only focuses on protein domains which are also known as the structural and functional units of proteins that diversify their function.

Biologically, protein domains are defined as highly conserved protein regions that fold independently and are self-stabilising (Murzin et al., 1995; Basu et al., 2009). An independent folding unit is a distinct and self-contained region within a protein's three-dimensional structure that can adopt its native conformation by a pathway that is not influenced by folding events occurring elsewhere in the protein (Batey et al., 2008). The folding pathway of an independent folding unit is said to be a cooperative process, meaning that the folding of one region in the protein stimulates the rest of the protein to fold as well (Batey et al., 2008). Therefore, the folding pathway of an independent folding unit is best described by the two-state model which is characterised by a sigmoidal curve with a single smooth transition from the unfolded state to the native state (Batey et al., 2008). This means that if a domain from a multidomain protein is
expressed independently of the rest of the protein, it should cooperatively adopt its native structure (Murzin et al., 1995). Protein domains are also defined as self-stabilising units (Murzin et al., 1995; Basu et al., 2009). A self-stabilizing domain maintains its structural integrity despite variations in environmental conditions like temperature, pH, and salt concentration. Importantly, its stability is independent of the rest of the protein's capacity to withstand similar structural changes. The native conformation of a domain from a multidomain protein should, therefore, demonstrate adequate stability when studied in isolation (Batey et al., 2008; Murzin et al., 1995; Basu et al., 2009). Now that it is known what a protein domain is, the next question is “What does it do?”. Since the structure of a domain is highly specific to its function, protein domains serve as the functional units of proteins whether it is in isolation (single-domain proteins) or in concert with other domains in a multidomain protein (Vogel et al., 2004). An example of a single-domain protein is haemoglobin. Haemoglobin comprises a heme-binding domain which is critical for its function in the transportation of oxygen and carbon dioxide (Marengo-Rowe, 2006). Receptor tyrosine kinase is an example of a multidomain protein as it possesses a ligand-binding domain which is key to initiating a signalling transduction cascade and a kinase domain that phosphorylates specific intracellular proteins in response to ligand binding (Hubbard & Till, 2000). Of course, Rotavirus VP6 is also an example of a protein with more than one domain (Mathieu et al., 2001). As mentioned previously in section 1.4.2, the coordination of the VP6 beta-sheet and alpha-helical domains allows the virus to become transcriptionally active when the contacts between the beta-sheet domain and the overlaying VP7 are lost (Aoki et al., 2011; Desselberger, 2014). The combination of domains in a multidomain protein therefore allows a single protein to have multiple functions (Marcotte et al., 1999; Tordai et al., 2005; Basu et al., 2009). Protein domains have also been observed to evolve independently (Basu et al., 2009). This means that, in a multidomain protein, evolutionary changes to the structure of one of the domains do not affect the other domains in the protein. However, said evolutionary changes may modify the overall architecture of the multidomain protein which would allow it to adapt in response to a changing environment and gain (or lose) functions (Basu et al., 2009). Some protein domains have regulatory functions that control the activity of the protein (Sommese et al., 2017). An example of this is the EF-hand domain which is found in calmodulin. Calmodulin is a small protein known to facilitate the regulation of divalent calcium cations in several physiological pathways (Kawasaki & Kretsinger, 2017; Walsh, 1983).
The EF-hand domain of calmodulin is highly sensitive to the divalent calcium cation concentration in its local environment and exerts its regulatory effects by activating (or deactivating) calmodulin activity in response to increasing (or decreasing) calcium concentrations (Kawasaki & Kretsinger, 2017; Walsh, 1983). Finally, protein domains may also serve the purpose of allowing a protein to form larger, more complex structures with other proteins (Affranchino & Gonzalez, 1997). The Rotavirus VP6 domains, though not confined to this single function, are examples of domains that drive protein-protein interactions. Specifically, the beta-sheet domain drives the association of VP6 monomers into the quaternary trimeric conformation which in turn allows the proteins of the outermost layer to bind and form the TLP (Aoki et al., 2011; Desselberger, 2014). On the other hand, the alpha-helical domain allows for interactions with the proteins of the viral core to form the DLP (Affranchino & Gonzalez, 1997).

1.6.2 VP6 and its Domains

1.6.2.1 Structure of VP6

The structural protein VP6 is a trimeric protein made up of identical subunits and each subunit comprises a beta-sheet domain and an alpha-helical domain (Figure 1.3 A). In the TLP, VP6 interacts with the outer (VP7 and VP4) and inner (VP2) capsid proteins through its beta-sheet and alpha-helical domains, respectively (Figure 1.3 B; Desselberger, 2014).

1.6.2.2 The Beta-Sheet Domain

The beta-sheet domain is the upper region of VP6 (Figure 1.3 A) and is formed by amino acid residues 151-334 (Mathieu et al., 2001; Gasteiger et al., 2003). This domain has a molecular weight of 20.79 kDa and demonstrates a Swiss-roll supersecondary structure, which describes a structural motif wherein eight beta strands are arranged in two sheets comprising four strands each (Richardson, 1981; Mathieu et al., 2001; Gasteiger et al., 2003). Deletion studies have shown that this domain is key for the formation of the VP6 trimer (Affranchino & Gonzalez, 1997). This is due to the high number of hydrophobic residues within the beta-sheet domain that make it less likely to dissociate in a polar environment (Mathieu et al., 2001). Amino acid residues 246-314 are key to the formation of trimer, and without these residues trimerization and formation of the DLP do not occur (Affranchino & Gonzalez, 1997).
Figure 1.3: VP6 Monomer, Trimer, and Organisation in the TLP.

(A) Front view of the *Rotavirus* VP6 monomer and homotrimer. The alpha-helical domain (VP6\(\alpha\)) (red) is the base of the subunit and is made up of two segments and is formed by residues 1-150 and 335-397. The beta-sheet domain (VP6\(\beta\)) (purple) is the upper region of the subunit and is formed by residues 151-334. (B) Cross-sectional view of the *Rotavirus* TLP. VP4 (blue), VP7 (dark blue), VP6 (purple and red), and VP2 (orange). The images of the ribbon structures were generated using PyMOL Version 2.5.0 (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). Image A was generated from PDB code 1QHD (Mathieu *et al.*, 2001) and Image B was generated from PDB code 4V7Q (Settembre *et al.*, 2011).
The beta-sheet domain is not only necessary for the formation of the DLP but is also a key component in the reformation of the TLP during the assembly stage of the replication cycle (Aoki et al., 2011; Desselberger, 2014). It is known that VP6 mediates the entry of the DLP into the rER through the interaction of the beta-sheet domain with the NSP4 component of the NSP4/VP4 complex (Aoki et al., 2011; Desselberger, 2014). Once the beta-sheet domain binds to NSP4, the DLP enters the rough ER by receptor-mediated endocytosis. The DLP is then enclosed in a vesicle containing VP7, the second component of the outermost layer of the capsid (Aoki et al., 2011; Desselberger, 2014). The VP7 protein latches onto the underlying beta-sheet domain and acts as a molecular switch that either activates or represses uncoating and transcription (Aoki et al., 2011; Desselberger, 2014). When VP7 latches onto the beta-sheet domain VP6, uncoating and transcription are repressed (Aoki et al., 2011; Desselberger, 2014). When the virus enters the cell, the low concentration of calcium ions in the vesicle causes the VP7 to detach from the beta-sheet domain (Aoki et al., 2011; Desselberger, 2014). This detachment has two effects; firstly, it stimulates VP4 to permeabilise the vesicle thereby allowing the DLP to be released and, secondly, it causes a conformational change in VP6 that activates the transcriptional and mRNA capping activities of the VP1/VP3 complex (Aoki et al., 2011; Desselberger, 2014). There is also evidence suggesting that the body produces antibodies against VP6, and these antibodies specifically recognise and bind to a region of the beta-sheet domain (Aiyegbo et al., 2013; Ward & McNeal, 2010). When said region of the beta-sheet is bound by antibodies, transcription and the release of viral mRNA are inhibited (Aiyegbo et al., 2013; Ward & McNeal, 2010).

1.6.2.3 The Alpha-Helical Domain

The alpha-helical domain is formed by residues 1-150 and 335-397. These residues fold into eight α-helices that form the base of the trimer as shown in Figure 1.3 A. In the TLP (Figure 1.3 B), the alpha-helical domain interacts with the innermost layer of the viral capsid (Desselberger, 2014). This innermost layer comprises twelve VP2 decamers associated with eleven VP1/VP3 complexes (McClain et al., 2010). When the outer layer of the viral particle is lost, VP6 undergoes a conformational change that activates the transcriptional activity of the VP1/VP3 complex and it has been noted that transcription is not activated in the absence of VP6 (Charpilienne et al., 2002; Desselberger, 2014). This is due to the important role of VP6 in maintaining the structural integrity of the DLP (Charpilienne et al., 2002).
In the study by Charpilienne et al. (2002), it was noted that the hydrophobic interactions of the alpha-helical domain with the underlying VP2 were responsible for stabilising the DLP. In the same study, it was also observed that certain residues found in the alpha-helical domain were necessary for activating the transcriptional and mRNA capping activities of the VP1/VP3 complex. These residues are also necessary for the release of the newly synthesised viral mRNA and when said residues were deleted or mutated, a transcriptionally inactive DLP resulted (Charpilienne et al., 2002). This domain is not key in the formation of the VP6 trimer. In deletion studies done by Affranchino & Gonzalez (1997), it was noted that deletions in the alpha-helical domain do not prevent trimerization likely due to the fact that the alpha-helical domain is rich in hydrophilic residues that would easily separate in polar environments. The study also presented a previously unidentified assembly domain formed by residues 122 – 147. An assembly domain is a region of a polypeptide chain that directs the association of the protein with other proteins. This assembly domain is different from the trimerization domain identified in the beta-sheet domain in that the assembly domain is needed for successful interactions between VP6 and VP2 (Affranchino & Gonzalez, 1997). It was reported that mutant VP6 proteins with deletions in the assembly domain could trimerize, but their ability to assemble into DLPs was significantly diminished (Affranchino & Gonzalez, 1997).

1.7 Bacterial Protein Expression and the Formation of Inclusion Bodies

*Escherichia coli* (*E. coli*) is the most widely used bacterial expression system because it can easily and inexpensively be maintained in culture (Lederberg, 1952; Ratzkin & Carbon, 1977). These organisms have short cultivation times and produce recombinant proteins in large quantities (Lederberg, 1952; Ratzkin & Carbon, 1977). The principle of bacterial expression is that the bacteria can be stimulated to take up extracellular DNA (known as plasmids) after being heat shocked and treated with concentrated calcium chloride, as these conditions disrupt the bacterial cell wall to facilitate the uptake of said extracellular DNA. A plasmid is a molecule of circular DNA that replicates independently of host DNA replication (Peña-Miller et al., 2015). Plasmids are often used as vectors in expression studies when biological macromolecules, such as proteins, need to be synthesised in quantities that exceed the amount produced under normal physiological conditions. Expression vectors contain the DNA sequence of the protein of interest, a selectable marker, and an inducible promoter. In the presence of an inducer, the protein of interest is overexpressed in the bacterial cells.
The recombinant protein can be secreted, but it is more common to have the protein accumulate inside the cell since higher product yields are obtained that way (Slouka et al., 2019). In some instances, the accumulation of recombinant protein in the cytoplasm has resulted in the formation of insoluble inclusion bodies (Ratzkin & Carbon, 1977). Inclusion bodies commonly form when bacterial systems are used for the overexpression of proteins. Inclusion bodies are insoluble protein aggregates that form within bacterial cells when the production rate of the recombinant protein exceeds the cellular capacity for proper folding (Singh et al., 2015). This typically occurs in high-expression systems, where protein synthesis overwhelms the host cell's chaperone and folding machinery. Consequently, misfolded or partially folded protein molecules aggregate and precipitate, forming inclusion bodies (Singh et al., 2015). The culturing conditions have also been linked to inclusion body formation. Indeed, the high temperatures, high inducer concentrations, lengthy post-induction incubation periods, and the bacterial strain used may drive protein misfolding and inclusion body formation (Van den Berg et al., 1999; Van den Berg et al., 2000; Singh et al., 2015). There are two main types of inclusion bodies: classical and non-classical. These distinctions are based on differences in protein composition and structure. Classical inclusion bodies consist primarily of misfolded protein aggregates (Balachander et al., 2016). They are often dense, highly structured, and rich in β-sheet content. These inclusion bodies are typically resistant to solubilization by conventional means and require harsh denaturants like strong detergents, chaotropic agents, or high concentrations of urea for solubilization (Singh & Panda, 2005). In contrast, non-classical inclusion bodies are less structured and contain a higher proportion of native-like protein conformations (Upadhyay et al., 2012). Non-classical inclusion bodies are generally more amenable to solubilization under mild conditions, such as changes in pH, temperature, or ionic strength (Upadhyay et al., 2012). Before any further studies can be done, the inclusion bodies must first be isolated and the insoluble contents must be solubilised, purified, and refolded.

1.8 Solubilisation of Inclusion Bodies

Solubilization of inclusion bodies is a critical step to recover and purify recombinant proteins. Inclusion bodies must be solubilized because they are initially insoluble, rendering the proteins biologically inactive. The solubilization process aims to unfold and dissociate protein aggregates, restoring the proteins to their native, functional conformation. Classical inclusion bodies are typically solubilized by the addition of denaturing agents (such as urea or guanidine hydrochloride), detergents, and/or by altering the pH (Fischer et al., 1992; Rudolph & Lilie,
Urea and guanidine hydrochloride disrupt protein-protein interactions and unfold the protein thereby allowing it to regain solubility (Yang et al., 2011). Non-classical inclusion bodies can also be solubilized using detergents, which is a milder alternative to denaturing agents (Burgess, 1996; Kudou et al., 2011; Singh et al., 2015). Adjusting pH or ionic strength have also been reported to facilitate the solubilization of inclusion bodies (Singh & Panda, 2005). The Sigma Aldrich website (accessed August 2023) lists the advantages and disadvantages of solubilising recombinant protein from inclusion bodies. The advantages are, firstly, solubilization allows for the recovery of a significant amount of the recombinant protein, maximizing the yield. Secondly, proper solubilization often results in highly pure protein fractions, as contaminants remain in the insoluble fraction. Thirdly, solubilized proteins can be refolded, which is crucial for functional assays and therapeutic protein production. The disadvantages of solubilising proteins from classical inclusion bodies include time-consuming refolding experiments, potential loss of sample due to aggregation when the denaturant or solubilising agents are removed during refolding and loss of protein activity upon refolding.

1.9 Previous Expression and Purification of Rotavirus VP6

Recombinant VP6 has been expressed using a number of techniques that produced variable results in terms of the solubility of the recombinant protein. In separate studies done by Bredell et al. (2016) and Aijaz & Rao (1996), VP6 was expressed in E. coli BL21(DE3). In both studies it was observed that the BL21(DE3) strain produced a large amount of recombinant VP6, however, the protein was entirely insoluble and was expressed as inclusion bodies and it was proposed that the trimeric quaternary structure of VP6 is too complex for the bacterial systems to process which caused protein misfolding and aggregation (Bredell et al., 2016). In the study by Zhao et al. (2011), a protocol for the solubilisation and renaturation of insoluble VP6 was outlined. In their protocol, the authors also expressed VP6 in E. coli BL21(DE3) and the recombinant VP6 was expressed as inclusion bodies. The inclusion bodies were isolated and treated with high concentrations of urea. The urea separates protein aggregates by disrupting the bonds between the misfolded proteins thereby solubilising them (Bennion & Daggett, 2003). The soluble proteins were purified and then renatured by the gradual dialytic removal of the denaturing agents. This protocol has two disadvantages: firstly, the high urea concentration may result in the biological function of the recombinant protein being lost even if the protein is renatured and secondly, there may be instances of sample loss due to
precipitation (Bugli et al., 2014). A protocol for the freeze-thaw solubilisation of inclusion bodies was proposed by Qi et al. (2015). In this protocol, E. coli BL21(DE3) cells were induced to overexpress the epidermal growth factor protein (EGFP) and the catalytic domain of human macrophage metalloelastase (MMP-12_CAT). Both proteins were reported to have been expressed as inclusion bodies which were subsequently isolated and frozen in phosphate buffered saline containing different molar concentrations of urea (0 M – 8 M) and then thawed at room temperature. The study showed that freezing and thawing inclusion bodies in a buffer with at least 2 M urea was as effective as conventional solubilisation in 8 M urea, but less damaging to the native secondary structures of the proteins. The authors also reported that, following the removal of urea either by dialysis or rapid dilution, the biological activity of both proteins were recovered. Russell and Gildenhuys (2018) have shown that this protocol can be applied to viral structural proteins as they successfully solubilized Bluetongue Virus (BTV) VP7 inclusion bodies and reported that the purified protein had native-like secondary and tertiary structural features.

Chromatography is a common technique used for the physical separation of a mixture by the selective distribution of mixture components between a mobile and stationary phase (Coskun, 2016). In previous studies, VP6 had been purified using size exclusion chromatography (SEC), ion-exchange chromatography (IEC), hydrophobic interaction chromatography (HIC), and His-tag affinity chromatography (Plascencia-Villa et al., 2011; Li et al., 2014; Badillo-Godinez et al., 2015). Plascencia-Villa et al. (2011) purified recombinant VP6 using SEC and IEC. In SEC, the column is filled with beads of different pore sizes. The beads are designed to exclude proteins that exceed the size limit of their upper fractionation range (Porath, 1997). This means that proteins are retained in the matrix for different periods of time. Large proteins spend the least time in the column as these proteins are likely to be excluded from all the beads in the column (Porath, 1997). Small proteins on the other hand are retained for the longest time because these proteins are not excluded from any of the different beads in the column (Porath, 1997). Ion-exchange chromatography is used to separate proteins based on the charge of the protein at a specific pH. The net charge on a protein under a set given condition is influenced by its isoelectric point (pI), a parameter that refers to the pH at which a given protein has a net charge of zero (Schuurmans-Stekhoven et al., 2008). When the pH of the buffer exceeds its pI, then the protein will be negatively charged and when the pI of the protein exceeds the pH of the buffer, then the protein will be positively charged (Rabilloud & Lelong, 2011; Coskun,
The charged protein binds to an oppositely charged resin, while neutral proteins and proteins with the same charge as the resin are washed away (Himmelhoch, 1971). The resins can either be cation exchangers (negatively charged) or anion exchangers (positively charged). Soluble VP6, which has a pI between 5.25 – 5.8, was purified by anion exchange chromatography using a Q-Sepharose resin and a buffer with a pH of 6.16 (Emslie et al., 2000; Plascencia-Villa et al., 2011). In studies by Li et al. (2014), recombinant VP6 was purified using HIC. Hydrophobic interaction chromatography is a chromatographic technique used to separate proteins based on their hydrophobicity and phenyl-sepharose is the most commonly used adsorbent for this application (Prescott et al., 1993). In the study by Badillo-Godinez et al. (2015), recombinant VP6 was purified using His-tag affinity chromatography. In this study, the gene sequence of VP6 was cloned into plasmids that add a 6 × His-tag to the N-terminus of the recombinant proteins. A column containing a nickel resin was used as the stationary phase because the imidazole side chain of histidine interacts and readily forms bonds with the nickel ions on the resin. This immobilises the tagged protein on the column while the rest of the crude mobile phase is washed away. Affinity chromatography is an excellent means of purifying a protein of interest from a complex mixture, such as bacterial cell lysate, as the affinity of the his-tag for the nickel resin reduces non-specific binding (Adamíková et al., 2019). Purification by IMAC results in higher yields of pure protein compared to SEC and IEC in a single purification. SEC and IEC may have to be performed in concert with other purification techniques to achieve yields that are comparable to those obtained by IMAC alone (Adamíková et al., 2019).

1.10 Probing the Structure and Conformational Stability of Proteins
Spectroscopic techniques such as far-UV circular dichroism and intrinsic tryptophan fluorescence spectroscopy are powerful tools for probing protein structure and stability. Far-UV circular dichroism (CD) measures the difference in the absorption of left and right circularly polarised light by molecules demonstrating the property of chirality (Greenfield, 2006). Chirality is a geometric property seen in molecules and refers to structures that cannot be superimposed onto their mirror image (Brooks et al., 2011). Proteins are chiral macromolecules therefore their structure and stability can be characterised using CD. Far-UV CD spectroscopy is based on the interaction of proteins with circularly polarised light. Circularly polarised light comprises a pair of electromagnetic waves of equivalent amplitude that oscillate perpendicular to each other and to the plane of the direction of the wave
These electromagnetic waves are of equal amplitude and rotate either clockwise or anti-clockwise (Greenfield, 2006). When circularly polarised light passes through a protein sample, the protein differentially absorbs left and right circularly polarised light and this differential absorption is known as circular dichroism (Hammes, 2005; Berova et al., 2000; Greenfield, 2006; Ranjbar & Gill, 2009). Far-UV CD is used to characterise the secondary structures of proteins. It has been observed that alpha helices, beta sheets, and random coils have distinct spectra (Greenfield, 2006; Ranjbar & Gill, 2009). Alpha helices have characteristic minima at 222 nm and 208 nm and a peak around 190 nm (Greenfield, 2006; Ranjbar & Gill, 2009). The signals at 208 nm and 190 nm result from the carbonyl π→π∗ transition. The π→π∗ transition occurs when an electron in a bonding π-orbital of the carbonyl group is excited to a non-bonding π∗-orbital (Ranjbar & Gill, 2009; Toniolo et al., 2012; Zsila, 2022). This transition is significantly impacted by the overall conformation of the polypeptide backbone. The signal at 222 nm originates from the n→π∗ transition. The n→π∗ transition occurs when one of the electrons in an amide lone pair is excited to a non-bonding π∗ orbital. The signal at 222 nm is highly sensitive to the alpha-helical content of a protein (Kelly et al., 2005; Ranjbar & Gill, 2009; Woody, 2006; Toniolo et al., 2012; Zsila, 2022). Beta sheets have a general negative band between 210 nm – 220 nm with a trough near 218 nm and a positive band of high magnitude around 195 nm (Kelly et al., 2005; Greenfield, 2006; Ranjbar & Gill, 2009). The trough at 218 nm results from the π→π∗ transition while the peak at 195 nm is caused by the n→π∗ transition (Kelly et al., 2005; Greenfield, 2006; Ranjbar & Gill, 2009). Random coils have a negative band of great amplitude around 195 nm - 200 nm and a positive band at 212 nm (Greenfield, 2006; Ranjbar & Gill, 2009; Woody, 2006; Toniolo et al., 2012). The negative band is caused by the π→π∗ transition while the positive band at 212 nm is caused by the n→π∗ transition (Greenfield, 2006; Ranjbar & Gill, 2009; Woody, 2006; Toniolo et al., 2012). The Rotavirus capsid protein VP6 was analysed by far-UV circular dichroism in a study by Zhao et al., 2011. The spectrum presented by the authors had two negative bands at 208 nm and 222 nm. The 208 nm band was of greater magnitude to that of the 222 nm and this was reportedly due to the presence of beta-sheet structures that countered the signal derived from the alpha-helices. Circular dichroism is sensitive to changes in secondary structures and is therefore ideal for monitoring the changes that occur when a sample is exposed to changing buffer conditions (such as changes in pH and ionic strength) as well as increasing temperature and denaturant concentrations (Greenfield, 2006). In the study done by Russell & Gildenhuys (2018) far-UV circular dichroism was used to monitor changes in the secondary structure of
the *Bluetongue Virus* VP7 (which is similar in structure to *Rotavirus* VP6) in response to increasing temperature. There were clear shifts in the CD spectra as the structure of VP7 was altered. The authors also selected key wavelengths, specifically 218 nm and 222 nm, to analyse the effect of the increasing temperature on the beta-sheet and alpha-helical content of VP7. Circular dichroism can similarly be used to monitor changes in secondary structure in response to increasing or decreasing concentrations of denaturants (such as urea and guanidine hydrochloride) as well as changes in pH, buffer composition, and ionic strength.

Intrinsic tryptophan fluorescence is an excellent way to probe the changes in the tertiary structure of proteins (Vivian & Callis, 2001). Tryptophan is an aromatic amino acid with a non-polar side chain. This means that tryptophan is hydrophobic and commonly located in globular protein's core. The electrons in the aromatic ring of tryptophan residues can be selectively excited to a higher energy level at a wavelength of 295 nm (Mollar & Denicola, 2006; Hellman & Schnider, 2019). When these excited electrons return to the ground state, they emit a low-energy photon and this emission is known as fluorescence (Lakowicz, 1999). The fluorescence spectrum of tryptophan differs depending on where the residue is located for example, free tryptophan in solution has a maximum emission peak near 350 nm and tryptophan residues that are buried in the core of a protein peak between 330 nm – 340 nm (Mollar & Denicola, 2006). This makes tryptophan a suitable probe for monitoring changes in protein conformation, specifically protein unfolding as changes in the local environment of tryptophan will result in a shift of the maximum emission peak to longer or shorter wavelengths. When the emission peak shifts to longer wavelengths, it is called a red shift or a bathochromic shift and this indicates that the local environment of the tryptophan residues has become more polar (Vivian & Callis, 2001). In proteins, it means that tryptophan residues are exposed to the solvent and in the case of globular proteins where tryptophan is usually buried, this would indicate that the protein has unfolded. The emission peak can also shift to shorter wavelengths and this is known as a blue shift or hypsochromic shift. In this case, the local environment of tryptophan residues becomes less polar which indicates refolding or increased viscosity of a protein sample (Vivian & Callis, 2001). As shown in Figure 1.4, VP6 has a total of five tryptophan residues: two in the beta-sheet domain and three in the alpha-helical domain (Zhao *et al*., 2011). The tertiary structure of VP6 was also characterised by fluorescence spectroscopy in the study by Zhao *et al*., 2011.
The fluorescence emission spectrum presented by the authors had a maximum emission wavelength at 341 nm indicating tryptophan residues that are mainly buried. Tryptophan fluorescence was also used in the study by Russell & Gildenhuys (2018) to monitor changes in the tertiary structure of the Bluetongue Virus VP7 in response to increasing temperature. As the case with their reported CD data, the fluorescence spectra were noted to have red shifted in response to the increasing temperature which indicated that the local environment of the tryptophan residues had become more solvent-exposed. The authors also selected the maximum emission wavelength (341 nm) to monitor changes in the fluorescence intensity at this wavelength in response to increasing temperature. The unfolding curve obtained from the fluorescence data can be used to further analyse the stability of a protein. This means that tryptophan fluorescence is also a suitable technique for monitoring protein structure under changing conditions. Both far-UV CD and intrinsic tryptophan fluorescence are powerful tools that are used to study the conformational stability of proteins. By selecting key wavelengths from each data set, it is possible to visualise the pathways by which specific protein structures are altered or denatured. These pathways would reveal (1) regions where structural changes are first observed, (2) the cooperativity of the unfolding/refolding event, (3) whether the unfolding of a protein occurs in a single step or through a multi-state process involving the formation of intermediates, and (4) whether or not an unfolding event is reversible. More detailed information about the stability of a protein can be extracted from CD and fluorescence data when the data are fit to models, such as the two-state or multi-state models. By fitting the data to an appropriate model, it is possible to determine thermodynamic parameters, such as Gibbs free energy, entropy, and enthalpy, associated with the unfolding of a protein. It is however only possible to fit the data to the two-state model if an unfolding event occurs in a single cooperative step and is reversible (Freire & Murphy, 1991; Zwanzig, 1997; Bakk et al., 2000). The multistate model can be used to fit data where intermediates were detected (Harder et al., 2004). In order to fit data to this model, however, the protein unfolding must be reversible, the number of intermediates that form must be known, and the intermediates must be confirmed by multiple probes (Harder et al., 2004).
**Figure 1.4: Location of Tryptophan in VP6**

Space-fill model of the 1QHD PDB structure of VP6 (Mathieu et al., 2001) generated using PyMOL Version 2.5.0 (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). VP6$_β$ is shown in purple and VP6$_α$ is shown in red. Tryptophan residues are shown as blue spheres. Image A is the front view of the VP6 monomer and Image B is the view when Image A is rotated 180° around the central vertical axis.
1.11 Differential Scanning Calorimetry

Differential scanning calorimetry (DSC) is a non-spectroscopic technique that is used to study the thermal stability of biological and non-biological polymers. In DSC, the difference in heat capacity of a protein sample and a reference sample (usually the buffer the protein was dialysed into) is measured (Gill et al., 2010). The melting temperature (Tm) of the protein corresponds to the peak of the thermogram. Thermodynamic parameters such as the enthalpy (ΔH) and heat capacity associated with the unfolding can also be determined by DSC provided that the thermal reaction is reversible (Gill et al., 2010). The melting temperature of VP6 was determined by DSC to be 78°C (Li et al., 2014).

1.12 Aim and Objectives.

The aim of this project is to establish expression and purification protocols that would allow the structure and stability of the VP6 alpha-helical and beta-sheet domains to be characterised. This study comprises five objectives. The first is to optimise the bacterial expression of the domains by expressing the domains in two *E. coli* strains (BL21 (DE3) and NiCo21 (DE3)) by varying the inducer concentration, the post-induction temperature, and post-induction incubation time. The second is to solubilise and then purify the bacterially expressed domains by immobilised metal chromatography (IMAC). The third objective is to characterise the structures of the purified domains by mass spectroscopy, far-UV circular dichroism, intrinsic fluorescence spectroscopy, and native-PAGE. The fourth objective is to probe the secondary and tertiary structural changes that occur as the domains are exposed to increasing denaturant (urea and guanidine hydrochloride) concentrations and temperature using far-UV circular dichroism and intrinsic fluorescence spectroscopy. The final objective is to measure the melting temperatures of the domains using differential scanning calorimetry.
Chapter 2 Methodology

2.1 Materials
The pET15a plasmids containing the codon optimised consensus sequence (Parbhoo et al., 2016) of either the RV VP6 α-helical domain (VP6α) or the RV VP6 β-sheet domain (VP6β), were purchased from GenScript (Hong Kong, China). The codon optimisation of the consensus sequences was done by GenScript (Hong Kong, China). Escherichia coli BL21 (DE3) and NiCo21 (DE3) strains were purchased from New England Biolabs (Massachusetts, USA). Isopropyl β-D-1-thiogalactopyranoside (IPTG), ampicillin, and high purity urea were purchased from Melford (Ipswich, United Kingdom). Prepacked Nickel cross-linked agarose resin columns for chromatography were purchased from Cytiva (Upssala, Sweden). Ultra-pure water was used for all buffer preparations. All other reagents used in this study were of analytical grade.

2.2 Bacterial Transformation
Competent BL21 (DE3) and NiCo21(DE3) E. coli cells were transformed as per the manufacturers’ instructions with the pET15a plasmids containing the selectable marker (AmpR), the IPTG-inducible T7 promoter sequence, and the cDNA consensus sequence (that is, the series of nucleotides that appear most frequently in RV VP6 genes) of either VP6α or VP6β (Sternke et al., 2020). The transformed cells were plated on Luria-Bertani (1.0% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) sodium chloride) agar containing 0.1 mg/ml of the ampicillin and incubated at 37°C overnight. A single colony was selected from the respective overnight culture plates and used to inoculate freshly prepared LB broth with 0.1 mg/ml of ampicillin. The culture was incubated overnight at 37°C with 200 rpm shaking. A 1 in 50 dilution of these overnight cultures were made using freshly prepared LB broth supplemented with 0.1 mg/ml ampicillin and incubated at 37°C with 200 rpm shaking. To monitor bacterial growth, the OD_{600} of the cultures were measured every 20 minutes for 5 hours using the Implen benchtop spectrophotometer.

2.3 Glycerol Stock Preparation
Glycerol stocks were prepared by combining equal volumes of bacterial culture (OD_{600} 0.6 – 0.8) and sterile 60% glycerol. The stocks were stored at -80°C until needed.
2.4 Protein Expression Studies

Growth curves were constructed using the data obtained from measuring the OD_{600} of the 50 times diluted bacterial cultures every 20 minutes for 5 hours. This was done to determine (1) the OD_{600} reading that corresponds to the early to mid-log phase of bacterial growth and (2) the time taken to reach the mid-log phase of bacterial growth. All cultures were grown to mid-log prior to induction with IPTG. The plasmid that was used to transform the bacterial cells contains the lacI gene which encodes a tetrameric protein that acts as a repressor of the lac operon. The binding of this repressor prevents the expression of the gene encoding the RNA polymerase (RNAP) that binds to the T7 promoter sequence. The silencing of the RNAP gene prevents the expression of all genes that are regulated by the lac operon. An inducer, such as IPTG, is used to displace the repressor of the T7 promoter and this allows for the expression of the RNAP and subsequently, the gene of interest (Reznikoff et al., 1974). For the expression of the alpha-helical and beta-sheet domains, BL21(DE3) cultures were supplemented with either 0.1 mM or 1.0 mM of IPTG while the NiCo21(DE3) cultures were supplemented with 0.04 mM or 0.4 mM of IPTG (Tarar et al., 2020). The different IPTG concentrations used to induce expression in the BL21 (DE3) and NiCo21 (DE3) strains were based on the optimum inducer concentrations given for each strain by the manufacturer. The cells were then grown at 20°C for 16 h and at 37°C for 2 h– 7 h and 16 h, respectively. Cultures were incubated with 200 rpm shaking. Post-induction cells were collected by centrifugation (4400 rpm for 1 hour) and resuspended in ultrapure water. The cells were placed on ice and sonicated using a Qsonica sonicator (Newtown, United States of America). Sonication was done in five two-minute intervals of 36 amperes, two seconds on one second off. The samples were kept on ice for two minutes between intervals. The lysate was centrifuged at 13400 rpm for 25 minutes to separate the soluble and insoluble fractions. The expression and solubility of the domains were assessed by analysing the soluble, and insoluble fractions by SDS-PAGE.

2.5 Isolation and Solubilisation of Inclusion Bodies

Post-induction bacterial cells were collected by centrifugation and lysed as described in section 2.4. The lysate was centrifuged and aliquots of the supernatants and pellets were reserved for SDS-PAGE. Once the insolubility of the proteins was confirmed by SDS-PAGE, the remaining supernatants were discarded, and the pellets were kept for further processing. The pellets containing VP6β were resuspended in Wash Buffer A (100 mM Tris-HCl pH 7.00 with 1% Triton X-100) and incubated at room temperature for 10 minutes before centrifugation (15000
rpm for 15 minutes). The supernatant was discarded and the pellet was washed twice with Wash Buffer B (100 mM Tris-HCl pH 7.00). The inclusion bodies were then resuspended in a range of 100 mM Tris-HCl buffers with pH values between 7.00 – 10.00 and urea concentrations of 2 M or 5 M. For every 0.1 g (wet weight) of pellet, 10 ml of buffer was used. The samples were incubated at room temperature for 90 minutes before being frozen at -20°C overnight. Pellets containing VP6α were resuspended in deionised water and the sample was centrifuged at 15000 rpm for 25 minutes. The supernatant was discarded and the pellets were resuspended in a range of 100 mM Tris-HCl buffers with pH values between 7.00 – 10.00 and urea concentrations of 2 M or 5 M, incubated at room temperature for 90 minutes, and frozen at -20°C overnight. The samples were thawed on ice and centrifuged at 15000 rpm for 15 minutes and protein solubility was assessed by running the supernatants and pellets on an SDS-PAGE gel. Supernatant and pellet samples were collected between all wash steps and these samples were also run on an SDS-PAGE gel. All buffers were filtered using a 0.45 micron filter before use.

2.6 Purification

All protein purification work was done using the Biorad NGC System (Hercules, United States of America). Unless otherwise stated, the flow rate used during purification was 2 ml/min. Soluble VP6β was loaded onto a 5 ml His-trap column that had been pre-equilibrated with 15 column volumes of Equilibration buffer A (100 mM Tris-HCl pH 9.00, 2 M urea, 150 mM NaCl, 40 mM imidazole, and 0.02% (w/v) sodium azide). The flow rate used for the sample application was 0.5 ml/min. The column was washed with 15 column volumes of Column wash buffer A (100 mM Tris-HCl pH 9.00, 2 M urea, 300 mM NaCl, 40 mM imidazole, and 0.02% (w/v) sodium azide). The protein was then eluted on a gradient (40 mM – 600 mM imidazole) using column wash buffer A and elution buffer A (100 mM Tris-HCl pH 9.00, 2 M urea, 300 mM NaCl, 600 mM imidazole, and 0.02% (w/v) sodium azide) over 10 column volumes. The column was finally washed with five column volumes of elution buffer A. An aliquot of each of the fractions, as well as the flow-through collected during sample application and the column wash steps were prepared for SDS-PAGE. The fractions containing pure VP6β were pooled and concentrated using an Amicon Stirred Cell and Ultracel membrane (Merck, Darmstadt Germany). The concentrated sample was dialysed at room temperature using 10 kDa molecular weight cutoff tubing. The first and second buffer exchanges were done over eight hours (the buffer was replaced after four hours) in Dialysis Buffer 1 (20 mM Sodium phosphate (dibasic-monobasic) pH 7.40 with 2 M urea).
The third exchange was done overnight in Dialysis buffer 2 (20 mM Sodium phosphate (dibasic-monobasic) pH 7.40) while the final exchange was done over four hours into dialysis buffer 2. If the sample was to remain in 2 M urea, then dialysis was done similarly using only dialysis buffer 1.

The VP6α sample was solubilised as described in section 2.5 and the soluble protein was applied to a pre-equilibrated His-tag column at a flow rate of 1 ml/min. The column was equilibrated with 10 column volumes of equilibration buffer B (100 mM Tris-HCl pH 7.40, 2 M urea, 150 mM NaCl, 40 mM imidazole, and 0.02% (w/v) sodium azide). The column was then washed with 15 column volumes of column wash buffer B (100 mM Tris-HCl pH 7.40, 2 M urea, 300 mM NaCl, 40 mM imidazole, and 0.02% (w/v) sodium azide) before being eluted on an imidazole gradient (40 mM – 500 mM). The elution gradient was run using 15 column volumes of elution buffer B (100 mM Tris-HCl pH 7.40, 2 M urea, 300 mM NaCl, 500 mM imidazole, and 0.02% (w/v) sodium azide) and wash buffer A. The column was finally washed with five column volumes of elution buffer B. Aliquots of the collected fractions and flow-through were run on an SDS-PAGE gel. The fractions containing pure VP6α were pooled, concentrated, and dialysed. VP6α was dialysed as previously described using dialysis buffers 1 and 2, depending on whether the sample was to remain in urea or not. The proteins were usually used up immediately after the final buffer exchange; however, in exceptional cases the protein was stored at –80°C until needed. The structures of protein samples that have been frozen were assessed by circular dichroism and intrinsic fluorescence spectroscopy after thawing. All buffers were filtered using a 0.45 micron filter before use.

2.7 Protein Gel Electrophoresis
2.7.1 SDS-PAGE

The principle of SDS-PAGE is that the anionic detergent, sodium dodecyl sulphate (SDS), linearises and provides proteins with a net negative charge thereby allowing for size-based separation on a polyacrylamide gel (Laemmli, 1970; Hjerten et al., 1988). This technique was used at different stages in the study for the assessment of 1) recombinant protein expression, 2) the solubility of the recombinant proteins, 3) the solubility (and optimum solubilising conditions) of the recombinant proteins following solubilisation studies, 4) the purity of the fractions collected after IMAC purification. The samples were prepared for SDS-PAGE by mixing samples with sample buffer (62.5 mM Tris-HCl pH 6.8, 10 % (v/v) glycerol, 2 % (v/v)
SDS, 5 % (v/v) β-mercaptoethanol, and 0.05 % (w/v) bromophenol blue) in a 1:4 ratio and boiling at 94°C for 5 minutes. The samples were electrophoresed (160 V for 1 hour) using SDS-PAGE (12% resolving gel, 5% stacking gel) in a discontinuous buffer system using a BioRad Mini-PROTEIN Electrophoresis Cell (Hercules, USA). The Unstained Protein MW Marker from Thermo Scientific product code 26610 (Massachusetts United States of America) was used in all SDS-PAGE gels. The gels were stained in InstantBlue (Abcam, United Kingdom) for 1 hour and were viewed using a BioRad Gel Dock Universal Hood (Hercules, United States of America).

2.7.2 Native-PAGE

The proteins were also assessed by native-PAGE. In native-PAGE, proteins are prepared and run in non-denaturing and non-reducing conditions. This allows for separation based on the size, shape, and pI of the protein of interest. The absence of denaturing and reducing agents in the native-PAGE also means that all interactions between different subunits in multimeric proteins are retained. This makes native-PAGE a suitable technique for assessing the quaternary structure of proteins. VP6β has a pI of 7.14 (Gasteiger et al., 2003) and was prepared for native-PAGE by mixing equal volumes of the protein and native sample buffer (62.5 mM Tris-HCl pH 6.8, 40% (v/v) glycerol, 0.05% Coomassie Brilliant Blue R250, and 0.05% (w/v) bromophenol blue). The addition of Coomassie imparts a negative charge on proteins without denaturing them and was therefore needed to allow the domain to enter the gel as it would theoretically have a positive charge in the pH 6.8 buffer (Schägger & von Jagow, 1991).

VP6α has a pI of 5.6 (Gasteiger et al., 2003) and was prepared for native-PAGE by mixing equal volumes of protein and native sample buffer (62.5 mM Tris-HCl pH 6.8, 40% (v/v) glycerol, and 0.05% (w/v) bromophenol blue). The acidic pI of VP6α makes it suitable for use in the pH 6.8 native sample buffer without Coomassie Brilliant Blue as it would have a net negative charge at pH 6.8. The samples were loaded onto a Pre-cast BioRad Mini Protein TGX Protein Gels (Hercules, USA) and run at 200 V for 2 hours using a BioRad Mini-PROTEIN Electrophoresis Cell (Hercules, USA). The gels were stained overnight in InstantBlue (Abcam, United Kingdom) and were viewed using a BioRad Gel Dock Universal Hood (Hercules, United States of America).
2.8 Spectrophotometry

2.8.1 Spectrophotometric Determination of Protein Concentration

2.8.1.1 Absorbance Spectrometry

The absorbance spectra of purified VP6α and VP6β were recorded in triplicate in a 10 mm quartz cuvette using the Photophysics Chirascan and in a 10 mm plastic cuvette using the Implen P330 Nanophotometer. The A_{280} reading was corrected for scatter (and buffer contributions) and used to calculate protein concentration using the Beer-Lambert law:

\[ A = εl[C] \]

Where A is the absorbance at 280 nm, l is the pathlength (cm), C is the concentration (M), and ε is the extinction coefficient (M^{-1}. cm^{-1}) of the protein in question. The extinction coefficients at 280 nm of VP6β and VP6α were calculated using the following equation described by Pace et al. 1995:

\[ ε_{280} M^{-1}. cm^{-1} = (5500 \times \sum W) + (1490 \times \sum Y) + (125 \times \sum C) \]

The extinction coefficients of VP6β and VP6α were calculated as 15720 M^{-1}. cm^{-1} and 27055 M^{-1}. cm^{-1}, respectively. The absorbance spectra of VP6β and VP6α were recorded in Dialysis Buffer 1 which was filtered before use.

2.8.1.2 Bradford Assay

The concentrations of the purified VP6β and VP6α were also determined by the Bradford assay using bovine serum albumin (BSA) as a standard. A BSA 2 mg/ml stock was prepared using filtered Dialysis Buffer 2 as a solvent. The rest of the BSA standards (0.25 mg/ml – 1.5 mg/ml) were prepared by direct dilution of the 2 mg/ml stock using the sterile filtered dialysis buffer 2 as the diluent. The standards and the domain samples were incubated in Bradford reagent (BioRad, Hercules, United States of America) for 30 minutes then the absorbance at 595 nm was monitored using the Implen P330 Nanophotometer (München, Germany). Absorbance readings were done in triplicate and the average absorbance of the standards were plotted against the respective concentrations. A trendline was fitted and the concentrations of VP6β and VP6α were calculated using the equation of the trendline.
2.8.2 Mass Spectroscopy

The sequences of the purified proteins were analysed by mass spectroscopy. A single band from the respective purification SDS-PAGE gels was excised and sent to the Council of Scientific and Industrial Research (CSIR) (Pretoria, South Africa) where the domains were extracted from the gel and cleaved using trypsin as described by Schevchenko et al., (2006). The peptides thus obtained were analysed by mass spectroscopy (LSMS) using a Dionex Ultimate 3000 RSLC system coupled to an AB Sciex 5600 TripleTOF mass spectrometer.

2.8.3 Circular Dichroism

Far-UV circular dichroism was used to assess the secondary structures of VP6β and VP6α. The CD spectra were recorded using the Applied Photophysics Chirascan Plus (Leatherhead, United Kingdom). Spectra were recorded between 190 nm – 250 nm in a 1 mm quartz cuvette. The bandwidth was set to 1.5 nm and the scan rate was 2 nm/s. The spectra were recorded in units of millidegrees (mdeg) which were converted to mean residue ellipticity (MRE) using the following formula:

\[ [\theta] = \frac{100 \text{ (signal)}}{Cn\ell} \]

where C is the concentration of protein in mM, n is the number of amino acid residues and l is the path length in cm. The spectra were recorded in triplicate at a temperature of 20°C in both dialysis buffers. All spectra had buffer contributions subtracted and both dialysis buffers were filtered before use.

2.8.4 Intrinsic Tryptophan Fluorescence Spectroscopy

The tertiary structure of VP6β and VP6α in dialysis buffers 1 and 2 were assessed by intrinsic tryptophan fluorescence spectroscopy in a 1 cm quartz cuvette (bandwidth 2.5 nm and step-size of 1 nm) using the Applied Photophysics Chirascan Plus (Leatherhead, United Kingdom). Tryptophan residues were selectively excited at a wavelength of 295 nm and emission was recorded between 280 nm – 400 nm. All spectra were recorded in triplicate at a temperature of 20°C and had buffer contributions subtracted.
2.8.5 Thermal Stability Studies

Far-UV CD and intrinsic tryptophan fluorescence spectroscopy were used to monitor the effects of increasing temperature on the structure of VP6β and VP6α in dialysis buffers 1 and 2. The spectra were recorded using the Photophysics Chirascan Plus (Leatherhead, United Kingdom) and the temperature was regulated by the PCS.3 Single Cell Peltier Temperature Controller with its accompanying chiller unit (Leatherhead, United Kingdom). The temperature of the samples was increased in increments of 2°C from 20°C to 90°C. Far-UV CD spectra were recorded between 250 nm - 190 nm and intrinsic tryptophan fluorescence spectra were recorded between 280 nm - 500 nm.

2.8.6 Chemical Stability Studies

The Applied Photophysics Chirascan Plus was used to record the far-UV CD and intrinsic tryptophan fluorescence spectra of VP6β and VP6α in varying concentrations of urea (0 M – 8 M) and guanidine hydrochloride (0 M – 4 M). The samples were prepared using VP6β and VP6α stocks in dialysis buffer 2, a 10 M urea stock, and a 6 M guanidine hydrochloride stock. The denaturant stocks were prepared as described by Pace (1986) using Dialysis Buffer 2 as a solvent and were filtered before use using a 0.45 micron filter. The samples were incubated for three hours at room temperature and spectra were recorded in triplicate as described in sections 2.8.2 and 2.8.3, respectively. To test the reversibility of the chemical denaturation of the domains, another set of samples was prepared where the stock solutions were made up of VP6β or VP6α in the maximum denaturant concentration (that is, 8 M urea and 4 M guanidine hydrochloride). Following a three-hour incubation, direct dilutions of these stocks were made to lower the denaturant concentrations. The protein-urea stock was diluted down to 6 M, 5 M, 4 M, 3 M, and 2 M while the protein-guanidine hydrochloride stock was diluted down to 3 M, 2.5 M, 2 M, 1.5 M and 1 M. The diluted samples were incubated at room temperature for three hours and spectra were recorded in triplicate as described previously.

2.8.6.1 Normalising Fluorescence Spectra

The fluorescence spectra were normalised in a way that represented the maximum fluorescence emission as a value of 1, while all other data points were expressed as fractions of 1. This was achieved by dividing each data point by the maximum emission value. This was done in order to clearly show any shifts in the emission spectra as the denaturant concentration was increased.
2.8.6.2 Conversion to Fraction Unfolded

The chemical unfolding data were converted to fraction unfolded as follows. A least squares regression line was fit to the pre- and post-transition regions of the unfolding curve, respectively. The y value obtained by extrapolation of the least squares regression line fit to the pre-transition represents the native state of the protein ($y_N$). Similarly, the y value obtained by extrapolation of the least squares regression line fit to the post-transition region represents the denatured protein ($y_D$). The fraction of unfolded protein was calculated using the following formula:

$$\text{Fraction Unfolded} = \frac{(y-y_N)/(y_D-y_N)}$$

Where $y$ is the signal obtained from either far-UV CD or intrinsic fluorescence spectroscopy.

2.9 Differential Scanning Calorimetry

Differential scanning calorimetry was done using the TA Instruments NanoDSC with DSCRun software v4.2.13 (Delaware, United States of America). Dialysis buffer 2 was run first to generate a baseline by loading the buffer into the sample and reference cells heating from 20°C to 90°C at a rate of 1°C/min following a 600 s equilibration. Following the generation of the baseline, the protein samples were loaded into the sample cells and Dialysis Buffer 2 was loaded into the reference cell and the experiment was done as described previously. The thermograms thus obtained were processed and analysed using TA Instruments’ NanoAnalyze software v3.12.0 (Delaware, United States of America).

2.10 AlphaFold

Alpha fold is a useful tool that employs artificial intelligence to predict the three-dimensional structure of proteins based on the proteins primary structure (Jumper et al., 2021). In this study, AlphaFold2 was used to predict the tertiary and quaternary structures of VP6R and VP6α. The prediction was done using open-source Google colab templates developed by Mirdita et al., 2022. The amino acid sequences of the respective domains were entered into the appropriate cell and the program was run according to the developers’ instructions. The monomeric structures were predicted using this template and the trimeric structures were predicted using this template.
2.11 Molecular Dynamics Simulation

The protein data bank (PDB) files of the monomeric and trimeric structures of VP6β were obtained from AlphaFold2 as described in section 2.10. The VP6β structures were prepared for the molecular dynamic simulation using the “prepare protein” workflow in Discovery Studio Client (DS) (v20.1.0.19). The default settings in DS were used to apply a CHARMM36 and to solvate the proteins. The standard dynamic cascade (SDC) conducted included a 20 ps equilibration, a 200 ps production step observation. With the exception of the steps mentioned previously, default DS settings and parameters were used to set-up the SDC. Root Mean Squared Fluctuation (RMSF) values were calculated for each amino acid based on the trajectories of the different conformations that resulted from the SDC. These data were visualised using SciDAVis (version 2.7.1).

2.12 Software and Online Tools

SciDAVis version 2.7.1 ([https://scidavis.sourceforge.net/index.html](https://scidavis.sourceforge.net/index.html)) was used to plot all spectral data. The ProtParam tool on ExPASY was used to determine the theoretical molecular weight and isoelectric point of the domains (Gasteiger et al., 2003). The K2D3 online CD analysis tool (Louis-Jeune et al., 2012) was used to calculate the secondary structure composition of the domains based on the far-UV CD data. The raw CD data was first converted to MRE as described in section 2.8.2 then the converted values were submitted to K2D3.
Chapter 3 Results

3.1 The Beta-Sheet Domain

3.1.1 Bacterial Growth Curves and Recombinant Protein Expression

3.1.1.1 Bacterial Growth Curves

Bacterial growth curves were used to monitor the growth of the BL21 (DE3) and NiCo21 (DE3) E. coli strains following their transformation with a pET15a plasmid containing the codon-optimized consensus DNA sequence of the Rotavirus VP6 beta-sheet domain (VP6β). A bacterial growth curve consists of three main phases: the lag phase, the exponential (or log) phase, and the stationary phase (Zwietering et al., 1990). These phases together form a sigmoidal curve representing the typical growth pattern of bacteria (Zwietering et al., 1990). The lag phase is associated with low replication as the bacteria are adjusting to the new environment (Rolfe et al., 2012). The exponential or log phase is associated with rapid bacterial replication and the stationary phase indicates that the number of new cells produced is equal to the number of cells that are dying and there is no further increase in cell bacterial growth due to limited resources (Navarro et al., 2010). Since the early-to-mid log phase of bacterial growth corresponds to optimum bacterial growth and replication, this is usually the stage of bacterial growth where protein expression is induced (Langlais & Korn, 2005). The growth curves of the BL21 (DE3) and NiCo21 (DE3) E. coli strains containing the codon-optimized consensus DNA sequence of VP6β are shown in Figure 3.1 (A and B), with the induction points represented by filled circles. It was concluded that, in both strains, a 1:50 diluted culture takes approximately 90 minutes to reach the early-log phase, and the OD$_{600}$ value at this early-log phase is 0.6.

3.1.1.2 Recombinant Protein Expression

Recombinant protein expression was assessed in the BL21 (DE3) and NiCo21 (DE3) E. coli strains by varying IPTG concentration and post-induction incubation time. The samples collected from the different expression conditions were prepared as described in section 2.4 and subjected to SDS-PAGE for visualization. The SDS-PAGE gels of the induced and uninduced cultures are shown in Figure 3.2 and Figure 3.3, respectively.
Figure 3.1: Bacterial Growth Curves.

Growth curves of the BL21 (DE3) strain (A) and NiCo21 (DE3) strain (B). The filled dots indicate the induction points.
Figure 3.2: Bacterial Expression of VP6β.

SDS-PAGE of VP6β expressed in the BL21 (DE3) and NiCo21 (DE3) *E. coli* strains. Pellet (P) and supernatant (S) samples from cells grown for 2 h, 3 h, 4 h, 5 h, 6 h, 7 h, and 16 h post-induction are shown. A and B represent BL21 (DE3) cultures induced with 1 mM and 0.1 mM IPTG; C and D represent NiCo21 (DE3) cultures induced with 0.04 mM and 0.4 IPTG, respectively. Overnight expression at 20°C is shown in panel E. Panel F is the labelled molecular weight marker (M). The arrows indicate VP6β.
Figure 3.3: Uninduced Cultures

(A) SDS-PAGE gels of the supernatant (S) and pellets (P) of uninduced NiCo21 (DE3) and BL21 (DE3) E. coli cultures with an arrow indicating VP6β. (B) Labelled molecular weight marker (M).
It can be noted that high intensity bands are visible in the insoluble fractions of the induced BL21 (DE3) and NiCo21 (DE3) cultures. These bands were thought to represent the recombinant VP6β and to verify this, a standard curve was constructed using the migration distance of the proteins in the molecular weight marker (Figure 3.2 F and Figure 3.3 B). The approximate molecular weight of the band thought to contain VP6β was calculated as 20.72 kDa. This extrapolated weight is consistent with the theoretical weight of VP6β (20.79 kDa) calculated by the ProtParam tool on ExPASY (Gasteiger et al., 2003). It was therefore concluded that both strains had successfully expressed VP6β. The findings of the expression study are as follows, firstly, there is an increase in recombinant protein production up to 5 h of post-induction incubation in both strains (Figure 3.2 A – D). Secondly, the different IPTG concentrations used did not result in differences in the expression of VP6β in either strain. Indeed, comparable amounts of protein were produced in BL21 (DE3) (Figure 3.2 A and B) and NiCo21 (DE3) (Figure 3.2 C and D), despite the tenfold difference in the IPTG concentrations tested in each strain. Thirdly, no increase in VP6β expression was observed in either strain when the temperature was lowered to 20°C (Figure 3.2 E). The high intensity bands seen in Figure 3.2 E is due to the samples not having been equalised for cell number prior to loading. Finally, all conditions tested (strain, IPTG concentration, post-induction time, and post-induction temperature) produced insoluble VP6β, as the band known to contain VP6β is in the insoluble fraction (pellet). When comparing the expression of VP6β between the BL21 (DE3) and NiCo21 (DE3) strains, it was observed that comparable amounts of recombinant VP6β were produced under the conditions tested however, the NiCo21 (DE3) cells may be marginally better than the BL21 (DE3) cells since very low concentrations of IPTG is needed to induce overexpression of VP6β. Therefore, the strain and expression conditions that corresponded to the optimum expression of VP6β were the NiCo21(DE3) cells induced with 0.04 mM IPTG grown for 7 h post-induction. These conditions were used for subsequent VP6β expression. An uninduced culture was used as a negative control and to determine whether there is leakage expression of VP6β (Figure 3.3). Leakage expression occurs when the protein encoded by the plasmid is synthesised in the absence of an inducer (Ellefson et al., 2014; Kato, 2020). This is a common occurrence in BL21 (DE3) strains and the strains derived from it due to basal expression of the T7 RNA polymerase. These basal activities of the T7 RNA polymerase result in the synthesis of mRNA from the plasmid which is subsequently translated (Du et al., 2021). It is important to test for potential leakage as some recombinant proteins may
be toxic to the bacteria and negatively affect bacterial growth rate and viability (Ellefson et al., 2014; Kato, 2020). The calculated molecular weight of VP6β is 20.72 kDa and a faint band of this size is visible in the insoluble fraction of the uninduced BL21 (DE3) culture but not in NiCo21 (DE3) culture indicating that VP6β is expressed at a low level in BL21 (DE3) in the absence of IPTG (Gasteiger et al., 2003).

3.1.2 Solubilisation Study

Due to the insoluble expression of VP6β (Figure 3.2), a solubilisation step was required before the recombinant protein could be purified. The solubilisation study was done by resuspending the insoluble VP6β in various buffers (100 mM Tris-HCl pH 7 – 10, with 0 M, 2 M, or 5 M urea) and freezing at – 20°C overnight (Qi et al., 2015). The SDS-PAGE gels of the solubilisation study along with the labelled molecular weight marker are shown in Figure 3.4 (B – F). It was noted that VP6β could not be solubilized without urea at any pH level, indicating that pH alone is insufficient to bring the domain into solution. As shown in Figure 3.4 (B and C), minimal solubilisation of VP6β was achieved using the pH 7 and pH 8 buffers in the presence of urea. This is likely due to the pI of VP6β (7.35) being too close to the pH of these buffers (Gasteiger et al., 2003). It is well known that proteins are charged, and therefore readily solubilised in conditions where the pH of the buffer is at least one unit above or below the proteins’ pI (Shaw et al., 2001). This was seen in the pH 9 and 10 buffers where the solubility of the domain had significantly increased. The degree of solubility achieved between the pH 9 and pH 10 (Figure 3.4 D and E) buffers are comparable however the pellet of the sample in pH 9 buffer with 2 M urea has the lowest amount of insoluble VP6β remaining after solubilisation. The 100 mM Tris-HCl, pH 9 buffer with 2 M urea was thus identified as the solution that effectively solubilised VP6β.

3.1.3 Purification

Soluble VP6β was purified by affinity chromatography on a prepacked column containing nickel resin using the Biorad NGC system. The column was equilibrated before the sample was loaded and was washed to remove all unbound material after the sample was loaded. An imidazole gradient was used to elute bound VP6β from the column (Figure 3.5 A). The protein was eluted in ten 2 ml fractions between imidazole concentrations of 150 mM – 270 mM. The solubilised VP6β, column washes and eluent were analysed by SDS-PAGE (Figure 3.5 B) and the molecular marker used is shown in Figure 3.5 C. It was noted that fractions 2 – 10 contained highly pure VP6β. All fractions were therefore pooled and concentrated before being dialysed.
Figure 3.4: Solubilisation of VP6β.

SDS-PAGE of VP6β frozen in 100 mM Tris-HCl buffers of varying pH (7 – 10) and urea concentrations (0 M, 2 M, and 5 M). (A) Supernatants (S) and pellets (P) obtained after washing the pellet with wash buffers A (wash 1) and B (wash 2). (B) Solubilisation at pH 7 (C) Solubilisation at pH 8. (D) Solubilisation at pH 9. (E) Solubilisation at pH 10. M is the molecular weight marker and VP6β is indicated by the arrow. (F) Labelled molecular weight marker.
Figure 3.5: Purification of VP6β.

(A) Elution profile of VP6β. The domain was eluted by an imidazole gradient (red) using Wash Buffer A and Elution Buffer A. (B) SDS-PAGE of the VP6β purification flow through and eluent. S is the solubilised VP6β sample. A is the sample application flow-through. W is the column wash flow-through. F1 – F10 are the fractions collected during elution. (C) Labelled molecular weight marker (M). The arrows indicate VP6β.
The concentration of the dialysed sample was determined through protein absorbance measurements using the Applied Photophysics Chirascan Plus and the Implen P330 Nanophotometer. A Bradford Assay was also done to validate the concentration determined by protein UV absorbance. The concentration of VP6β, as determined by protein UV absorbance (Figure 3.6 A), was found to be 0.95 mg/ml (approximately 46 µM). The concentration determined by a Bradford assay using bovine serum albumin (BSA) as a standard (Figure 3.6 B) was 0.93 mg/ml (approximately 45 µM) which is comparable to the findings of the benchtop spectrophotometer and Applied Photophysics Chirascan Plus. This concentration corresponds to a yield of 13.32 mg of pure VP6β from 1 L of culture.

3.1.4 Characterising the Structure of VP6β

3.1.4.1 Primary Structure
To confirm the identity of the recombinant protein domain, a sample of purified VP6β was sent to the CSIR for analysis by mass spectroscopy as described in section 2.6. The peptides detected by mass spectroscopy were matched to the known sequence of VP6β as shown in Figure 3.7. Overall, 52% of VP6β peptides were detected thereby confirming the purified sample was indeed VP6β. The mass spectroscopy data also showed that one asparagine residue was deamidated (Figure 3.7). Deamidation is a chemical process whereby the side-chains of asparagine or glutamine are altered or converted to a different functional group which affects the stability and ability of a protein to carry out its function (Kato et al., 2020). In the study done by Midelfort and Mehler (1972), the authors reported that deamidation of asparagine results in protein conformational changes that are directly related to degradation. When considering the sequence of VP6β, there are a total of 26 asparagine and glutamine residues and only one residue out of the 26 was deamidated suggesting that the domain was not in the process of degradation. This is supported by the spectral characterisation data (3.1.4.2 and 3.1.4.3) which indicated that the domain retained native-like conformations in the conditions tested. It is important to note that asparagine residues are more vulnerable to deamidation when located next to a residue with a small side-chain group on its C-terminal end, as there is no steric hindrance that shields it from deamidation (Robinson & Robinson, 2001). As shown in the sequence, the deamidated asparagine residue has a alanine residue on its C-terminal end and since alanine has a small side-chain group, it is possible that deamidation occurred due to the exposure of that asparagine residue (Robinson & Robinson, 2001).
Figure 3.6: Quantification of Purified VP6β.

(A) The absorbance spectrum of purified VP6β was recorded between 220 nm and 380 nm in a 0.1 cm quartz cuvette. (B) The standard curve \( y = 0.067x + 0.0122; r^2 = 0.9490 \) that was used to calculate the concentration of VP6β in the Bradford Assay. The orange dot indicates the coordinates (protein concentration; absorbance 595 nm) of VP6β on the standard curve.
Figure 3.7: VP6β Peptides Detected by Mass Spectroscopy

The sequence of VP6β with the peptides identified by mass spectroscopy is shown in purple. The detected peptides account for 52% of the VP6β sequence. The deamidated asparagine residue is indicated by the arrow.
3.1.4.2 Secondary Structure
Far-UV circular dichroism (CD) was used to characterise the secondary structure of the purified VP6β sample in the presence and absence of 2 M urea. The purified protein was dialysed into a 20 mM sodium phosphate solution pH 7 and spectra were recorded between 200 nm and 250 nm. VP6β was expected to have a general negative band between 210 nm and 220 nm and a peak around 200 nm, as spectra with these features are representative of proteins with mainly beta-sheet secondary structures (Corrêa & Ramos, 2009). The CD spectrum that was obtained (Figure 3.8; black spectrum) is consistent with that of a protein containing mainly beta-sheet secondary structures. The structure of VP6β was similarly assessed in the presence of 2 M urea (Figure 3.8; blue spectrum). In 2 M urea, there was a decrease in ellipticity between 210 nm – 200 nm. The CD data were converted to mean residue ellipticity as described in section 2.8.2 and loaded onto K2D3 (Louis-Jeune et al., 2012) to determine the percentage structural composition of VP6β. The K2D3 analysis revealed that VP6β comprises 46.03% and 43.64% beta-sheet structures in 0 M and 2 M urea, respectively. Figure 3.8 also includes the spectrum of VP6β in 8 M urea (purple spectrum). Due to the sample containing a high urea concentration, the spectrum could only be recorded up to 218 nm as the signal-to-noise ratio beyond 218 nm was too low. It is clear, however, that the spectrum of VP6β in 8 M urea is significantly different to that of the 0 M and 2 M urea samples. There is an increase in the mean residue ellipticity between 218 nm – 250 nm, which is consistent with the formation of random coils (Greenfield, 2006; Ranjbar & Gill, 2009). It can be concluded that VP6β retained native-like secondary structural features in 0 M – 2 M urea, but was denatured in 8 M urea.

3.1.4.3 Tertiary Structure
3.1.4.3.1 Intrinsic Tryptophan Fluorescence Spectroscopy
Intrinsic tryptophan fluorescence spectroscopy is often used to probe the tertiary structure of proteins. Tryptophan residues are known to be sensitive to the polarity of their local environment in a protein. In a tryptophan fluorescence spectrum, one would observe a maximum emission at a wavelength that corresponds to the local environment of tryptophan residues. Emission maxima between 300 nm – 330 nm indicate buried tryptophan residues while maxima between 345 nm – 355 nm indicate fully exposed tryptophan residues (Vivian & Callis, 2001). In the 0 M urea sample, it is natural to assume that the sample would adopt a native quaternary structure or otherwise form large oligomers or aggregates.
Figure 3.8: Far-UV CD Spectra of VP6β.

The far-UV CD spectra of VP6β recorded in 0 M (black), 2 M (blue), and 8 M urea (purple).
VP6β has two tryptophan residues namely Trp180 and Trp247 and, as shown in the space-fill model in Figure 3.9 A, Trp180 appears to be partially exposed while Trp247 is more buried. In this study, the tertiary structure of VP6β was probed using intrinsic tryptophan fluorescence spectroscopy. The spectra were recorded in 0 M, 2 M, and 8 M urea as described in section 2.8.3. The normalised tryptophan emission spectra are shown in Figure 3.9 B and it can be noted that there is an emission maximum at 338 nm in the 0 M and 2 M urea samples and at 351 nm in 8 M urea. This indicates that the tryptophan residues are mostly buried in 0 M – 2 M urea which is consistent with the location of the residues shown in Figure 3.9 A, and fully exposed in 8 M urea (Vivian & Callis, 2001). This indicates that a native-like tertiary conformation observed in 0 M urea is maintained in the presence of 2 M urea but lost following treatment with 8 M urea.

3.1.4.3.2 Molecular Dynamics Simulations

Though the spectra of the 0 M and 2 M urea samples indicate a native-like tertiary conformation of VP6β, there was a notable increase in fluorescence intensity in the spectrum of the domain in 2 M urea (Figure 3.9 B). This was investigated by molecular dynamics simulation as shown in Figure 3.10. An opensource Google co-lab template was used to access AlphaFold2 (Jumper et al., 2021; Mirdita et al., 2022) to generate Protein Data Bank (.pdb) files of the monomeric and trimeric forms of VP6β as described in section 2.10 (Figure 3.10 A). Molecular dynamic simulations were done on both the monomeric and trimeric conformations and the RMSF values for each amino acid were obtained (Figure 3.10 B). Both tryptophan residues (Trp34 and Trp100) are located in (or in the proximity of) the high flexibility region between residues Glu83 – Trp100 (Figure 3.10 B). There is a global decrease in residue flexibility when VP6β adopts a trimeric quaternary structure (Figure 3.10 B). The reduced flexibility seen in higher-order structures could potentially explain the observed fluctuations in fluorescence intensity. Several amino acid residues, such as phenylalanine, histidine, asparagine, aspartic acid, glutamine, and tyrosine, have been noted to quench tryptophan fluorescence by various mechanisms (Chen & Barkley, 1998) and as shown in (Figure 3.10 C), Trp100 is surrounded by these residues. The reduced flexibility of residues in the trimeric (or oligomerised) conformation of VP6β may result in energy transfers from the excited (or oligomerised) tryptophan residues to the non-fluorescing residues (that is, the intrinsic quenchers) mentioned previously in this section. Conversely, the effects of the intrinsic quenchers on tryptophan fluorescence may be reduced in the monomer as the flexibility in the region of Trp100 and Trp34 increases.
Figure 3.9: VP6β Tertiary Structure.

(A) Location of tryptophan residues in VP6β based on information from Zhao et al., 2011. Image created using the AlphaFold2 structure of the VP6β monomer (Jumper et al., 2021; Mirdita et al., 2022) and PyMOL Version 2.5.0 (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). (B) Intrinsic tryptophan fluorescence emission spectra recorded in 0 M (black), 2 M (blue), and 8 M (purple) urea.
Figure 3.10: Molecular Dynamics and Scatter.

(A) Structures of the VP6β monomer (side view) and trimer (viewed down its three-fold axis) predicted using AlphaFold2 (Jumper et al., 2021; Mirdita et al., 2022) and visualised using PyMOL Version 2.5.0 (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). (B) RMSF of the residues in the monomer (red) and trimer (black). (C) Intrinsic quenchers (red sticks) surrounding tryptophan residues (blue sticks) in VP6β. (D) Fluorescence emission at 295 nm of VP6β in 0 M, 2 M, and 8 M urea. This emission is caused by large light-scattering particles in the sample.
This may allow the side chains of tryptophan and quenching residues to adopt various orientations that minimise the energy transfer to non-fluorescing species, thereby resulting in an increase in fluorescence intensity. Aggregation is also known to significantly affect protein fluorescence. When proteins aggregate, a number of proteins are brought into close proximity to each other. This increases the probability of energy transfers between the neighbouring molecules which ultimately results in a decrease in fluorescence intensity; this is known as self-quenching (Chaari et al., 2015). Aggregation can be monitored in many ways including the evaluation of the emission intensity at the wavelength corresponding to the excitation wavelength. This is known as Rayleigh scattering and indicates whether the excitation light was scattered by large molecules in the sample (Tan et al., 2020). High emission intensities at the excitation wavelength indicate the presence of large particles that scatter the light. Notably, the 0 M urea sample was more aggregated than the 2 M urea sample as indicated by the fluorescence scatter data shown in Figure 3.10 D. The addition of 2 M urea to the VP6β sample reduced the number of large oligomers (or aggregates) in the solution without unfolding the protein (Figures 3.10 D and 3.9 B). This likely reduced the effects of self-quenching which then resulted in an increase in the fluorescence intensity. Though the scatter in the 8 M urea sample is more extensively reduced, there is no increase in fluorescence intensity as seen in the 2 M urea sample. This is because in addition to preventing oligomerisation and aggregation, 8 M urea also unfolds the protein as indicated by the redshift from 338 nm to 351 nm, and redshifts are usually accompanied by decreased fluorescence intensity due to increased exposure of the fluorophores to the solvent (Duy & Fitter, 2006).

### 3.1.4.4 Quaternary Structure

The quaternary structure of VP6β was evaluated in the 0 M – 8 M urea by native-PAGE in a Tris-Glycine buffer system (Figure 3.11). As shown in the native-PAGE gel, there was a single band observed in all urea conditions tested. Since the migration distance between the native sample (0 M urea) and the denatured sample (8 M urea) was comparable, it was concluded that the band observed in the gel most likely represents the monomeric form of the protein. However, based on the light scattering fluorescence data (Figure 3.10 D), it is possible that there are larger structures (or aggregates) present in the samples with low to no urea that have not entered the gel.
Figure 3.11: Native-PAGE of VP6β.

Native-PAGE gel of VP6β in different urea concentrations.
3.1.5 Characterising the Conformational Stability of VP6β.

3.1.5.1 Chemical Conformational Stability

3.1.5.1.1 Conformational Stability in Urea

The conformational stability of VP6β was assessed by far-UV CD and intrinsic tryptophan fluorescence spectroscopy in 0 M – 8 M urea. The far-UV CD and fluorescence spectra are shown in Figure 3.12 (A and B). There were clear changes in both the CD and fluorescence spectra as the urea concentration increased. The CD spectra show an increase in mean residue ellipticity as the urea concentration increased which as explained previously in section 3.3.4, is consistent with the loss of beta-sheet secondary structures and the formation of random coils (Greenfield, 2006; Ranjbar & Gill, 2009) (Figure 3.12 A). There was a red shift in the maximum emission wavelength from 338 nm to 353 nm in the fluorescence spectra as the urea concentration increased which indicates the loss of VP6β tertiary structure (Figure 3.12 B). To determine whether the urea-induced denaturation of VP6β is reversible, a sample of VP6β in 8 M urea was prepared. This sample was diluted to 2 M urea in 1 M increments and CD and fluorescence spectra were recorded (Figure 3.12 A and B). The same key wavelengths were analysed in the reversibility study and the pathways obtained following conversion to fraction unfolded are shown in Figure 3.12 C. Both the CD and fluorescence data show that VP6β does not re-acquire its native structural conformations by the same pathways they were lost thereby further illustrating the complexity of this unfolding pathway. Since the unfolding and refolding pathways do not overlay, it can be concluded that the chemical unfolding of VP6β in urea is a complex and irreversible process. Because of this irreversibility, fitting the data to a two-state or multi-state model for gaining deeper insights into the thermodynamics of this unfolding process is not possible. The turbidity data obtained from CD and fluorescence are shown in Figure 3.12 D. The HV readings provide insights into the voltage needed by the detector to effectively measure the incident light passing through a sample (Rodger and Marshall, 2021; Miles et al., 2021). This measurement is directly related to the absorption of the sample. A highly absorbing sample would significantly reduce the light reaching the detector which will in turn increase the voltage needed to pick up a signal (Rodger and Marshall, 2021; Miles et al., 2021). The signal at 295 nm indicates the presence of large light-scattering particles in a solution. It can be noted that the HV signal from the CD analysis increases with urea concentration while there is an overall decrease observed in the 295 nm fluorescence scatter signal (Figure 3.12 D).
Figure 3.12: VP6\textsubscript{β} Unfolding with Urea.

Far-UV CD spectra (A) and normalised intrinsic tryptophan fluorescence emission spectra (B) of VP6\textsubscript{β} recorded before denaturing (black), in 2 M urea (green) after denaturing with 8 M urea (red), and after refolding (blue). (C) Unfolding and refolding curves of VP6\textsubscript{β} as based on the changes in the ellipticity at 218 nm and the ratio of the fluorescence at 338 nm and 353 nm in response to increasing and decreasing urea concentrations. (D) HV at 218 nm and signal at 295 nm plotted against urea concentration.
This indicates that the addition of urea does not induce the formation of larger structures but rather drives the sample to a more uniform state. The reason the HV signal increases despite the simplification of the sample is due to the effect of the denaturant concentration on the signal-to-noise ratio. As the urea concentration increases, the signal-to-noise ratio decreases and this is reflected by an increase in HV. It is worth noting that a straightforward linear trend is not observed in the HV plot. There is a steep increase in the HV signal following the addition of 5 M urea. There are many possible explanations for this observation including conformational changes in the protein that increase its absorption. Urea is a chaotropic agent that, at high enough concentrations, destabilises protein structure thereby affecting the shape of the protein. The conformation of a protein will directly affect its interaction with the incident light and consequently, its absorption (Fohely & Suardi, 2018). In the context of this study, this explanation is plausible as the data clearly shows significant changes in the secondary and tertiary structures as the urea concentration reached 5 M and higher. Since the HV is measured alongside CD, it is worth noting the changes in the secondary structure of VP6β in this region. The loss of VP6β secondary structures increases from ~ 45% in 4.5 M urea to ~ 70% in 5 M urea (Figure 3.12 C). The significant loss of VP6β secondary structure sustained following treatment with 5 M urea may have caused the spike in HV at this concentration.

3.1.5.1.2 Conformational Stability in Guanidine Hydrochloride

The conformational stability of VP6β was assessed by far-UV CD and intrinsic tryptophan fluorescence spectroscopy in 0 M – 5 M guanidine hydrochloride (GdCl). The far-UV CD and normalised fluorescence emission spectra are shown in Figure 3.13 (A and B). Due to the high concentrations of GdCl that were used, far-UV spectra were only recorded between 218 nm – 250 nm, as this region had the best signal-to-noise ratio. The CD spectra show an increase in mean residue ellipticity as the GdCl concentration increased (Figure 3.13 A). This is consistent with the loss of beta-sheet secondary structures and the formation of random coils (Greenfield, 2006; Ranjbar & Gill, 2009). There was a red shift in the maximum emission wavelength from 338 nm to 353 nm in the fluorescence spectra as the GdCl concentration increased (Figure 3.13 B). This indicated that the tryptophan residues were fully exposed to the polar surroundings (Vivian & Callis, 2001). Also shown in Figure 3.13 (A and B) are the CD and fluorescence spectra of VP6β that had been unfolded in 5 M GdCl and refolded by diluting the sample to a final concentration of 1 M GdCl.
Figure 3.13: VP6β Unfolding with Guanidine Hydrochloride.

Far-UV CD spectra (A) and normalised intrinsic tryptophan fluorescence emission spectra (B) of VP6β recorded before denaturing (black), after denaturing with 5 M GdCl (red), and after refolding (green). (C) Unfolding and refolding curves of VP6β as based on the changes in the ellipticity at 222 nm and the ratio of the fluorescence at 338 nm and 353 nm in response to increasing and decreasing GdCl concentrations. (D) HV at 218 nm and signal at 295 nm plotted against GdCl concentration.
The CD spectrum of the refolded sample did not perfectly align with the spectrum of the native protein (Figure 3.13 A) but it appears that, based on the decrease in the MRE, native structures were recovered. The fluorescence spectrum of the refolded sample aligns well with that of the native protein, indicating that the tertiary structure of VP6β was also recoverable following chemical denaturation (Figure 3.13 B). As explained in section 3.3.5.1.1, key wavelengths were selected from the far-UV CD and fluorescence data and plotted against the GdCl concentration to inspect the pathway by which VP6β unfolds and refolds in GdCl. The CD mean residue ellipticity (MRE) at 218 nm and fluorescence intensity at 338 nm and 353 nm were selected to monitor structural changes in VP6β in the presence of GdCl. The CD and fluorescence data show similar unfolding pathways, however, the secondary structure of VP6β is destabilised at a lower GdCl concentration (1 M) than the tertiary structure (1.5 M) (Figure 3.13 C). A similar discrepancy in regions of minimal structural changes in VP6β was detected by the different spectral probes in GdCl as was seen in urea. There were minimal changes in the secondary structure of VP6β between 1.25 M – 2 M and 2.5 M – 3 M GdCl as detected by CD (Figure 3.13 C). The fluorescence data indicates that there were minimal changes in the tertiary structure between 2.25 M – 2.50 M and 2.5 M – 2.75 M GdCl (Figure 3.13 C). As explained in section 3.3.5.1.1, this is likely due to (1) secondary structural changes occurring in a region distal to the local tryptophan residues and (2) tertiary structural changes not affecting the secondary structure content of VP6β. The unfolding of VP6β using GdCl as a denaturant reveals a complex pathway that includes the formation of intermediates. To determine whether the GdCl-induced denaturation of VP6β is reversible, VP6β was left to equilibrate in 5 M GdCl before being incrementally diluted to 1 M GdCl. The same key wavelengths (218 nm (CD) and 338 nm and 353 nm (fluorescence)) were analysed in the reversibility study and the pathways from these points are shown in Figure 3.13 C. The CD and fluorescence data indicate that the refolding of VP6β in GdCl is likely a reversible process as the refolding pathways align more closely to the unfolding pathway. Though this process is reversible, fitting the data is not possible because the probes detected intermediate structures at different GdCl concentrations. One can therefore not be sure how many intermediates form along this pathway as the regions where the intermediates were detected are different between the CD and fluorescence probes. The turbidity data obtained from CD and fluorescence are shown in Figure 3.13 D. An increase in the HV signal and a decrease in the fluorescence intensity at 295 nm was also observed in the study with GdCl.
The opposite trends observed by these probes indicate that the addition of GdCl did not drive oligomerisation or aggregation of VP6β and the increase in the HV signal is due to the negative effects of the increasing GdCl concentration on the signal-to-noise ratio.

3.1.5.2 Thermal Conformational Stability

The thermal conformational stability of VP6β was assessed by far-UV circular dichroism and intrinsic tryptophan fluorescence spectroscopy while heating the sample to 90°C in a 20 mM sodium phosphate buffer with and without urea. This was done to compare the behaviour of the protein in different conditions where native-like conformation is maintained. In 0 M urea, VP6β was shown to have native-like secondary and tertiary structural features but contained large particles (likely, oligomers or aggregates) that were scattering the light. A decrease in the intensity at 295 nm was observed following the addition of 2 M urea which indicated that the denaturant had reduced oligomerisation or aggregation without altering the structure of VP6β.

3.1.5.2.1 Thermal Conformational Stability in 0 M Urea

The thermal conformational stability of VP6β was assessed by far-UV circular dichroism and intrinsic tryptophan fluorescence spectroscopy while heating the sample from 20°C to 90°C in a 20 mM sodium phosphate buffer without urea (Figure 3.14 A and B). Fluorescence and CD spectra were also recorded when the sample was cooled back to 20°C immediately after heating and after overnight incubation at 4°C. The CD spectra of the heated and cooled sample show that there was a decrease in MRE between 200 nm – 210 nm and an increase in MRE between 210 nm – 220 nm as the sample was heated (Figure 3.14 A). The decrease in MRE indicates the formation of random coils as the spectrum of random coil conformation is characterised by a negative band of large magnitude at 200 nm (Greenfield, 2006; Ranjbar & Gill, 2009; Woody, 2006; Toniolo et al., 2012). The decrease in MRE in this region may also be interpreted as the loss of beta-sheet secondary structures as the standard spectrum for beta-sheets is characterised by a positive band of high magnitude at 200 nm (Kelly et al., 2005; Greenfield, 2006; Ranjbar & Gill, 2009). The increase in MRE seen between 210 nm – 220 nm is minimal but still distinguishes the denatured sample from the native and is representative of a conformational change that was brought about by heating. When the sample was cooled to 20°C immediately after heating, the recorded spectrum overlapped the spectrum of the denatured sample. This indicated that the thermal denaturation of VP6β was not reversible in the timeframe of the experiment.
Figure 3.14: VP6β Thermal Studies in 0 M Urea.

Far-UV CD spectra (A) and intrinsic tryptophan fluorescence emission spectra (B) of VP6β recorded before heating (black), at 90°C (red), after cooling to 20°C (blue), and after cooling at 4°C overnight (green). (C) Unfolding and refolding curves of VP6β as based on the changes in the ellipticity at 218 nm (red) and the fluorescence intensities at 340 nm (green) and 355 nm (blue) in response to increasing temperature. (D) HV at 218 nm (black) and the signal at 295 nm (red) plotted against temperature.
To determine whether cooling the sample over a longer period would restore the native conformation, the heated sample was allowed to cool at 4°C overnight. The spectrum of the 4°C sample shows a further decrease in MRE between 200 nm – 210 nm and an increase between 210 nm – 220 nm. This indicates that the overnight cooling failed to restore the native secondary structure and may have caused further structural changes in VP6β. The fluorescence spectra of the heated and cooled samples are shown in (Figure 3.14 B). It can be noted that there was a shift in the maximum emission wavelength (λ_{max}) from 340 nm (at 20°C) to 355 nm (at 90°C) as well as a decrease in fluorescence intensity at 90°C. The shift in λ_{max} to longer wavelengths indicates that the local environment of the tryptophan residues had become more polar as the sample was heated, that is, the protein had unfolded (Vivian & Callis, 2001). The changing local environment of the tryptophan residues could result in a decrease in fluorescence intensity. As tryptophan is exposed to a polar environment, the residue may experience solvent quenching. Solvent quenching is a phenomenon that occurs when the fluorescence intensity of a fluorophore, such as tryptophan, is reduced in a polar environment because of the interactions that occur between excited fluorophores and surrounding solvent molecules (Duy & Fitter, 2006). Said interactions dissipate the excitation energy without generating fluorescence thereby reducing the intensity (Duy & Fitter, 2006). The increasing temperature may also be the cause of the observed decrease in fluorescence intensity. When a sample is heated, the molecules therein experience accelerated motion which results in more frequent and forceful collisions (Demchenko, 1986). These collisions facilitate the transfer of energy to non-fluorescent species, consequently reducing the amount of energy that is emitted as fluorescence (Demchenko, 1986). The spectrum of the sample that was cooled to 20°C is of equal magnitude to that of the denatured sample, however, there was a shift in λ_{max} to 345 nm (Figure 3.14 B). This is not consistent with the magnitude or λ_{max} of the native sample which means that the native tertiary structure is not recovered in the timeframe of the experiment. As observed in the CD spectra, the spectrum of the sample cooled overnight at 4°C did not result in the restoration of the native tertiary conformation.

To visualise the pathways by which the secondary and tertiary structures of VP6β are thermally denatured, key wavelengths from CD (218 nm) and fluorescence (340 nm and 355 nm) were plotted against temperature (Figure 3.14 C). The CD data shows that VP6β begins undergoing structural changes from 44°C as indicated by the increase in the MRE at 218 nm. Minimal secondary structural changes occur between 60°C – 68°C as the MRE in that region is fairly
constant. This could be an intermediate structure, however, there are not any corresponding regions in the 340 nm or 355 nm curves that show a similar trend. Instead, a decrease in fluorescence intensity with increasing temperature is observed at both wavelengths. As explained previously, this is likely due to solvent quenching or heat-induced dissipation of the excitation energy.

The sample turbidity at 218 nm (HV) and fluorescence intensity at 295 nm (scatter), were plotted against temperature to monitor aggregation (Figure 3.14 D). There were increases in both the HV at 218 nm and fluorescence intensity at 295 nm as the temperature increased which is indicative of thermally induced aggregation. Notably, there is a peak in the 295 nm curve between 60°C – 70°C, indicating that scatter is at a maximum in this temperature range. This is also the temperature range where an intermediate structure was detected by CD. This suggests that the intermediate likely forms as the sample aggregates.

It can be concluded that VP6β is irreversibly denatured when heated in denaturant-free conditions. The domain is prone to aggregation and an intermediate forms along this aggregation pathway. Thermodynamic data analysis is not possible due to the irreversible nature of the thermal unfolding pathway. Only an approximation of the melting temperature (T_m) could be extracted from the thermal data. Since an intermediate forms in this pathway, two T_m values can be described for this thermal study namely T_{m1} (~ 56°C) and T_{m2} (~ 72°C).

### 3.1.5.2.2 Thermal Conformational Stability in 2 M Urea

The thermal conformational stability of VP6β was assessed by far-UV circular dichroism and intrinsic tryptophan fluorescence spectroscopy while heating the sample from 20°C to 90°C in a 20 mM sodium phosphate buffer with 2 M urea (Figure 3.15 A and B). Fluorescence and CD spectra were also recorded when the sample was cooled back to 20°C immediately after heating and after overnight incubation at 4°C.
Figure 3.15: VP6β Thermal Studies in 2 M Urea.

Far-UV CD spectra (A) and intrinsic tryptophan fluorescence emission spectra (B) of VP6β recorded before heating (black), at 90°C (red), after cooling to 20°C (blue), and after cooling at 4°C overnight (green). (C) HV at 218 nm (black) and the signal at 295 nm (red) plotted against temperature. (D) Unfolding and refolding curves of VP6β as based on the changes in the ellipticity at 218 nm (red) and the fluorescence intensities at 340 nm (green) and 354 nm (blue) in response to increasing temperature.
The CD spectra of the heated and cooled sample show an increase in MRE between 210 nm – 220 nm and a change in the shape of the spectrum as the sample was heated (Figure 3.15 A). This is indicative of a loss of beta-sheet secondary structures and conformational changes that were brought about by heating. When the sample was cooled to 20°C immediately after heating, there was a slight decrease in MRE, however, the shape of the cooled curve was similar to that of the denatured sample. This indicated that the thermal denaturation of VP6β was not reversible in the timeframe of the experiment. The heated sample was allowed to cool at 4°C overnight to determine whether more time and a lower temperature would aid the refolding process. The spectrum of the 4°C sample shows a decrease in MRE but is still closer in shape to the spectrum of the denatured protein than to that of the native. This indicates that prolonged cooling is also ineffective in restoring the native conformation. The fluorescence spectra of the heated and cooled samples are shown in (Figure 3.15 B). It can be noted that there was a significant decrease in the fluorescence intensity as the sample was heated to 90°C as the spectrum recorded at this temperature lacks a definitive peak. A possible explanation for this observation could be that the protein underwent conformational changes that oriented the tryptophan residues in such a way that they were exposed to the solvent and therefore, highly likely to be involved in interactions with the surrounding solvent molecules resulting in quenching of the emitted fluorescence signal. Another possible reason for this observation is aggregation. As explained previously in section 3.1.4.3, the clumping together of protein molecules increases the chances of energy transfers from intrinsic fluorophores (tryptophan residues) to non-fluorescing molecules and decreases the emitted fluorescence. When considering the sample turbidity data, it can be noted that the HV and signal at 295 nm curves increased sharply from ~ 56°C and peaked between 62°C – 68°C (Figure 3.15 C). The 295 nm peak in particular was of great magnitude which indicates that the sample was rich in large light scattering aggregates at those temperatures. It is therefore possible that the combination of solvent-exposed tryptophan residues, extensive aggregation, and the dissipation of the excitation energy as with increasing temperature could have caused the observed decreases in the fluorescence intensity as the sample was heated.

The spectrum of the sample that was cooled to 20°C has a slightly higher intensity than the spectrum of the sample at 90°C but is not comparable to the spectrum of the native protein (Figure 3.15 B). This indicates that the native tertiary structure is not recovered in the timeframe of the experiment.
The spectrum of the sample cooled overnight at 4°C increased significantly in magnitude but did not overlay the spectrum of the native protein. This indicates that the overnight cooling step may have resulted in the partial recovery of the tertiary structure. However, considering that this sample was visibly aggregated after heating and the secondary structure analysis on the cooled sample indicates a non-native conformation, it is more prudent to conclude that the native conformation of VP6β cannot be recovered by a prolonged cooling period following thermal denaturation in 2 M urea.

To visualise the pathways by which the secondary and tertiary structures of VP6β are thermally denatured, key wavelengths from CD (218 nm) and fluorescence (340 nm and 354 nm) were plotted against temperature (Figure 3.15 D). The fluorescence data once again shows a decrease in the emission intensity at both wavelengths as the sample was heated. As explained previously in this section, this is likely due to solvent quenching, aggregation, and dissipation of the excitation energy that occurs when the sample is heated. The CD data shows that VP6β undergoes structural changes from 60°C. In the presence of 2 M urea, the unfolding of VP6β appears to be more cooperative, as there were no intermediates detected by either probe. The addition of 2 M urea to the VP6β sample increased the temperature at which structural changes first occur and despite not preventing aggregation, it was able to delay it by increasing the temperature needed to form aggregates. Since the denaturation of VP6β in 2 M urea was not reversible upon cooling, this data could not be analysed beyond the T_m which was found to be ~ 66°C.

### 3.1.5.3 Differential Scanning Calorimetry

Differential scanning calorimetry was done using a Nano DSC (TA Instruments). The run was set up to heat the sample from 30°C to 100°C at a heating rate of 1°C/min. The raw data was processed and analysed using NanoAnalyze software. The thermogram obtained for VP6β had a peak at 67.94°C (Figure 3.16). This is close to the T_m2 of ~ 72°C indicated by the thermal far-UV circular dichroism.
Figure 3.16: DSC Thermogram of VP6β in 0 M Urea.

The thermogram indicates the Tm of VP6β which is 67.94°C. The thermogram was generated by preparing the raw data using NanoAnalyse software.
3.2 The Alpha-Helical Domain

3.2.1 Bacterial Growth Curves and Recombinant Protein Expression

3.2.1.1 Bacterial Growth Curves

The commonly used BL21 (DE3) *E. coli* strain and its modified counterpart the NiCo21 (DE3) strain were both transformed with a pET15a plasmid containing the codon-optimized consensus DNA sequence of the *Rotavirus* VP6 alpha-helical domain (VP6α). To determine the time taken for a culture to reach the early to mid-log phase of bacterial growth, which as explained previously represents the stage of bacterial growth where cells are rapidly dividing, the optical density at 600 nm (OD$_{600}$) was measured every 20 minutes over 5 hours. The growth curves of the BL21 (DE3) and NiCo21 (DE3) *E. coli* strains containing the codon-optimized consensus DNA sequence of VP6α are shown in Figure 3.17 (A and B), with the induction points shown as filled circles. It was concluded that a 1:50 diluted BL21 (DE3) culture takes approximately 100 minutes to reach the early-log phase, and the OD$_{600}$ value at this early-log phase is ~ 0.5 and a 1:50 diluted NiCo21 (DE3) culture takes approximately 90 minutes to reach the early log phase, and in this case the corresponding OD$_{600}$ value is ~ 0.45.

3.2.1.2 Recombinant Protein Expression

Both *E. coli* strains were grown to mid-log before VP6α overexpression was induced by supplementing BL21 (DE3) cultures with either 0.1 mM or 1.0 mM IPTG and the NiCo21 (DE3) cultures with either 0.04 mM or 0.4 mM IPTG and incubating. Samples were collected after 2 h, 3 h, 4 h, 5 h, 6 h, 7 h, and 16 h of post-induction incubation at 37°C. Cultures grown overnight at 20°C and uninduced cultures were also sampled. The inducer concentration, post-induction incubation time, and temperature were varied to monitor the effect of those expression conditions on VP6α expression. The results of the expression study together with the molecular weight marker used for sizing the domain are shown in Figures 3.18, 3.19, and 3.20. It was noted that all SDS-PAGE expression gels contained a band of great intensity which was thought to be VP6α. This band was sized using a standard curve based on the migration of the proteins in the molecular weight marker. Based on the standard curve, the observed band was calculated to be 24.36 kDa in size which is close to the theoretical molecular weight of VP6α (24.16 kDa) calculated by ProtParam tool on ExPASY (Gasteiger et al., 2003). This means that the domain had been successfully expressed in both strains. Though leakage expression of VP6α was observed in both strains, the band containing VP6α from the uninduced BL21 (DE3) culture is of higher intensity than the band seen from the corresponding NiCo21 (DE3) culture.
Figure 3.17: Bacterial Growth Curves.

(A) Growth curve of the BL21 (DE3) strain. (B) Growth curve of the NiCo21 (DE3) strain. The filled dots indicate the induction points.
Figure 3.18: Bacterial Expression of VP6\(\alpha\) in BL21 (DE3).

SDS-PAGE of VP6\(\alpha\) expressed in the BL21 (DE3) \textit{E. coli} strain. Lysate (L), Pellet (P) and supernatant (S) samples from cells grown for 2 h, 3 h, 4 h, 5 h, 6 h, 7 h, and 16 h post-induction. (A) Labelled molecular weight marker (M). (B) Uninduced BL21 (DE3) cultures. (C) Cultures induced with 0.1 mM IPTG. (D) Cultures induced with 1.0 mM IPTG. The arrows indicate VP6\(\alpha\).
Figure 3.19: Bacterial Expression of VP₆α in NiCo21 (DE3).

SDS-PAGE of VP₆α expressed in the NiCo21 (DE3) *E. coli* strain. Lysate (L), Pellet (P) and supernatant (S) samples from cells grown for 2 h, 3 h, 4 h, 5 h, 6 h, 7 h, and 16 h post-induction. (A) Labelled molecular weight marker (M). (B) Uninduced NiCo21 (DE3) Culture. (C) Cultures induced with 0.04 mM IPTG (D) Cultures induced with 0.04 mM IPTG. The arrows indicate VP₆α.
Figure 3.20: Bacterial Expression of VP6α at 20°C.

(A) SDS-PAGE of VP6α expressed in the BL21 (DE3) and NiCo21 (DE3) E. coli strains. Pellet (P) and supernatant (S) samples from cells that were grown overnight at 20°C. VP6α is indicated by the arrow. (B) Labelled molecular weight marker (M).
This indicates that there is more basal or uninduced expression of VP6α in BL21 (DE3) compared to NiCo21 (DE3). The different expression conditions tested yielded the following results. Firstly, in both strains, there was an increase in recombinant protein production up to 5 h of post-induction incubation (Figures 3.18 and 3.19). Thereafter, there were no significant increases in expression. Secondly, the different IPTG concentrations used did not result in significant differences in the expression of VP6α in either strain as comparable amounts of protein were produced by each strain. Thirdly, the solubility of the expressed VP6α did not improve in either strain when the temperature was lowered to 20°C (Figure 3.20). Although the expression levels appear to be higher at 20°C than at 37°C it is important to note that the cultures were not equalised for cell number before SDS-PAGE analysis was done. The background proteins on the SDS-PAGE gel of the 20°C culture are more concentrated than the 16 h culture incubated at 37°C hence it is more prudent to conclude that the overnight expression levels of VP6α at 37°C and 20°C are comparable. Lastly, insoluble VP6α was produced in all conditions tested (strain, IPTG concentration, post-induction time, and post-induction temperature). When the expression of VP6α in the BL21 (DE3) and NiCo21 (DE3) strains was compared, it was observed that comparable amounts of VP6α were produced under the conditions tested, though the NiCo21 (DE3) strain was slightly better at producing VP6α at lower inducer concentrations. Therefore, the strain and expression conditions that corresponded to the optimum expression of VP6α were the NiCo21(DE3) cells induced with 0.04 mM IPTG grown for 16 h post-induction at 37°C or 20°C. In this study, subsequent expression was done at 37°C.

3.2.2 Solubilisation Study
VP6α was expressed as insoluble under all expression conditions tested and had to be solubilised before purification. The washed pellets containing the alpha-helical domain were resuspended in a range of buffers (100 mM Tris-HCl pH 7-10, with 0 M – 5 M urea), frozen at -20°C overnight, and thawed. Figure 3.21 shows the SDS-PAGE gels obtained following a single freeze-thaw cycle as well as the molecular marker used to size the protein. VP6α was successfully solubilised in all buffers that contained urea after freezing and thawing. It was noted that the pH of the buffer alone was not enough to solubilise the domain and effective solubilisation required at least 2 M urea. When comparing the solubilisation of the VP6α in the presence of 2 M and 5 M urea, it was noted that both concentrations adequately solubilised the domain at pH 7, pH 8, and pH 9). At pH 10, however, some of the alpha-helical domain was still insoluble in 5 M urea.
**Figure 3.21: Solubilisation of VP6α.**

SDS-PAGE of VP6α was frozen in buffers of varying pH (7, 8, 9, and 10) and urea concentration (0 M, 2 M, and 5 M). Supernatants (S) and pellets (P) were collected at various stages of the solubilisation protocol. (A) Ultra-pure water wash. (B) pH 7 Solubilisation. (C) pH 8 solubilisation. (D) pH 9 solubilisation. (E) pH 10 solubilisation. (F) Labelled molecular weight marker (M). VP6α is indicated by the arrows.
Since the sample could be solubilised at any pH between 7 and 10 with at least 2 M urea, the final solubilisation conditions selected were 100 mM Tris-HCl, pH 7.4 with 2 M urea. The pH of 7.4 was chosen as this was the manufacturer’s recommended pH for equilibrating the column that was used for purification.

3.2.3 Purification

The BioRad NGC system was used for the purification of VP6α. The domain eluted in 38 fractions and all fractions with high absorbance readings were analysed by SDS-PAGE. Figure 3.22 shows the elution profile obtained following purification, the SDS-PAGE gels of the collected fractions, and the molecular weight marker used to size the domain. Most of the protein that eluted at the first peak contained low molecular weight contaminants (Figure 3.22). Fractions 1 – 8 and 26 – 38 were pooled, concentrated and dialysed for further analysis. The concentration of purified VP6α was determined by absorbance spectroscopy (Figure 3.23 A). After correcting the A_{280} reading for scatter, the concentration was found to be 4.3 mg/ml (approximately 178 µM) which amounts to about 25 mg of pure protein from 1 L of culture. The Bradford assay with BSA standard (Figure 3.23 B) was used to confirm the result from the absorbance spectrum. The concentration of the purified alpha-helical domain extrapolated from the standard curve was 4.4 mg/ml (182 µM), which is comparable to the result obtained from the Applied Photophysics Chirascan Plus.

3.2.4 Characterising the Structure of VP6α.

3.2.4.1 Primary Structure

To determine the sequence of the recombinant protein, a sample of purified VP6α was sent to the CSIR for analysis by mass spectroscopy as described in section 2.6. The peptides detected by mass spectroscopy were matched to the known sequence of VP6α and approximately 74% of VP6α peptides were detected, confirming the purified sample was indeed VP6α (Figure 3.24). Three asparagine residues were noted to have been deamidated. The first deamidated asparagine residue (Figure 3.24; red) precedes a glycine residue. Glycine is known to have a small flexible side-chain which would not provide the necessary steric hindrance to reduce the rate of deamidation of asparagine (Robinson & Robinson, 2001). The second and third deamidated asparagine residues are flanked by residues with larger side-chain groups which should reduce the rate of asparagine deamidation (Robinson & Robinson, 2001). As mentioned previously, deamidation is a chemical event that precedes protein degradation however there is no evidence of degradation as only three of 30 VP6α deamidation prone residues were deamidated.
Figure 3.22: Purification of VP6α.

(A) Elution profile of VP6α. The domain was eluted by an imidazole gradient (red) using Wash Buffer A and Elution Buffer A. (B) Labelled molecular weight marker. (C) SDS-PAGE of the VP6α purification flow through and eluent. Lane 1: Solubilised VP6α. Lane 2: Sample application flow-through. Lane 3: Column wash flow-through. Lanes 1 – 38 Eluent fractions. The arrow indicates VP6α.
Figure 3.23: Quantification of Purified VP6α.

(A) The absorbance spectrum of purified VP6α was recorded between 220 nm and 380 nm. A 1 mm cuvette was used to record the absorbance of this sample. (B) The standard curve that was used to calculate the concentration of VP6α. A range of standards were prepared from a bovine serum albumin stock. The equation of the standard curve is 0.067x + 0.0122 and R-squared value is 0.9490. The orange dot indicates the coordinates (protein concentration; absorbance 595 nm) of VP6α on the standard curve.
Figure 3.24: VP6α Peptides Detected by Mass Spectroscopy.

The sequence of VP6α with the peptides identified by mass spectroscopy is shown in purple. The detected peptides account for 74% of the VP6α sequence. The first deamidated asparagine residue is shown in a red rectangle, the second in green, and the third in blue.
Additionally, the structural conformation studies done on VP6α (section 3.2.4.2 – 3.2.4.3) show that the domain had native-like conformation in the conditions tested. It is worth noting that deamidation is a process that occurs spontaneously in physiological conditions therefore this observation could be a consequence of the spontaneity of the deamidation reaction in the conditions tested (Kato et al., 2020).

3.2.4.2 Secondary Structure

The secondary structure of the purified VP6α sample was characterised by far-UV CD in 0 M and 2 M urea. The purified protein was dialysed into a 20 mM sodium phosphate solution pH 7 without and with 2 M urea and spectra were recorded between 200 nm and 250 nm. The spectrum of VP6α was expected to have negative bands of equal magnitude at 222 nm and 208 nm, and a positive band around 190 nm as spectra with these features are representative of proteins that are rich in alpha-helical secondary structures (Corrêa & Ramos, 2009; Kelly et al., 2005; Ranjbar & Gill, 2009; Woody, 2006; Toniolo et al., 2012; Zsila, 2022). The CD spectrum that was obtained (Figure 3.25; black spectrum) is consistent with the spectrum of an all-alpha-helical protein. The structure of VP6α was also assessed in the presence of 2 M urea (Figure 3.25; blue spectrum). In 2 M urea, there was a slight increase in ellipticity at 222 nm but the 208 nm signal is consistent with the native spectrum. The signal-to-noise ratio decreased at wavelengths below 208 nm, therefore the spectrum of VP6α in 2 M could not be recorded beyond this point. The CD data were analysed using the K2D3 tool on Dichroweb (Louis-Jeune et al., 2012) and the sample was found to be 86.38% alpha-helical with 0.89% beta-sheet structures in the urea-free sample and 86.30% alpha-helical with 1.24% beta-sheets in the 2 M urea sample. The spectrum of VP6α in 8 M urea is shown in Figure 3.25 (purple spectrum). Due to the sample containing a high urea concentration, the spectrum could only be recorded up to 222 nm as the signal-to-noise ratio below this wavelength was too low. The spectrum of VP6α in 8 M urea is significantly different to that of the 0 M and 2 M urea spectra. There was an increase in the mean residue ellipticity between 222 nm – 250 nm, which is consistent with the formation of random coils (Corrêa & Ramos, 2009; Kelly et al., 2005; Ranjbar & Gill, 2009; Woody, 2006; Toniolo et al., 2012; Zsila, 2022). It can be concluded that VP6α retained native-like secondary structural features in 0 M – 2 M urea, but was denatured in 8 M urea.
Figure 3.25: Far-UV CD Spectra of VP6α.

The CD spectra of VP6α recorded in 0 M (black), 2 M (blue), and 8 M urea (purple).
3.2.4.3 Tertiary Structure

Fluorescence spectroscopy was used to probe the tertiary structure of VP6α in the region of the intrinsic tryptophan residues. Tryptophan is known to be sensitive to the polarity of its local environment. Recall that the emission maxima between 300 nm – 330 nm indicate buried tryptophan residues while maxima between 345 nm – 355 nm indicate fully exposed tryptophan residues (Vivian & Callis, 2001). VP6α has three tryptophan residues namely Trp59, Trp139, and Trp182 as shown in the space-fill models in Figure 3.26 A. Trp59 and Trp139 appear to be mostly buried while Trp182 is partially exposed. Fluorescence spectra of VP6α were recorded in 0 M, 2 M, and 8 M urea as described in section 2.8.3. The tryptophan emission spectra are shown in Figure 3.26 B. There are emission maxima at 337 nm, 340 nm, and 354 nm in 0 M, 2 M, and 8 M urea, respectively. This indicates that the tryptophan residues are mostly buried in 0 M – 2 M urea which is consistent with the location of the residues shown in Figure 3.26 A, and fully exposed in 8 M urea (Vivian & Callis, 2001). This indicates that a native-like tertiary conformation observed in 0 M urea is not significantly altered in 2 M urea but is lost following treatment with 8 M urea. The fluorescence emission of VP6α in 2 M urea was significantly higher than its emission in 0 M urea (Figure 3.26 B). A possible explanation for the observed differences in fluorescence intensity is that VP6α was extensively aggregated in the 0 M urea sample as indicated by the 295 nm signals (Figure 3.26 C). In the absence of urea, the sample was rich in large light-scattering particles, such as oligomers or aggregates. The addition of 2 M urea reduced the signal at 295 nm, indicating a reduction in the number of large particles in the sample that scatter the light (Figure 3.26 C). Since the increase in fluorescence intensity in 2 M urea was accompanied by a decrease in the 295 nm signal, it is highly likely that the aggregates or large oligomers that form in the absence of urea resulted in decreased fluorescence intensity. When proteins come together to form larger structures or aggregates, it becomes more probable that energy would be transferred from the excited fluorophores to non-fluorescing residues and molecules thereby decreasing the energy that is emitted as fluorescence (Chaari et al., 2015). Since 2 M urea prevents the protein-protein interactions driving oligomerisation/aggregation, energy transfers to non-fluorescing species were reduced resulting in more of the excitation energy being emitted as fluorescence. It is worth noting that adding 8 M urea to VP6α reduced the 295 nm signal (Figure 3.26 C) and caused the maximum emission wavelength to shift to longer wavelengths (Figure 3.26 B). The emission intensity of VP6α in 8 M urea was lower than the emission intensity observed in 2 M urea (Figure 3.26 B).
Figure 3.26: VP6α Tertiary Structure.

(A) Side view of VP6α showing the location of tryptophan residues based on information from Zhao et al., 2011. Image created using the AlphaFold2 (Jumper et al., 2021; Mirdita et al., 2022) structure of the VP6α monomer in PyMOL Version 2.5.0 (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). (B) Intrinsic tryptophan fluorescence emission spectra recorded in 0 M (black), 2 M (blue), and 8 M (purple) urea. (C) Signal at 295 nm of VP6α in 0 M, 2 M, and 8 M urea.
This is due to the 8 M urea unfolding the protein and exposing the tryptophan residues to the surrounding polar environment where the fluorophores are likely to experience solvent quenching. Notably, the fluorescence emission of VP6α was higher in 8 M than in 0 M despite the spectrum of the 8 M urea being redshifted (Figure 3.26 B). This supports the idea that the tryptophan residues experience some form of self-quenching in the aggregated state as there was an improvement in the fluorescence emission when urea was added at denaturing and non-denaturing concentrations.

3.2.4.4 Quaternary Structure
The quaternary structure of VP6α was evaluated by native-PAGE in 0 M – 8 M urea. The gel indicates that the sample is a mixture of three oligomeric structures. As the urea concentration increases, the bands containing the larger oligomers become less concentrated while the concentration of the lower molecular weight structure increases (Figure 3.27 A). The sample however does not become fully monodispersed even in increasing urea concentrations. The quaternary structure data is supported by the fluorescence data which indicates that the scatter is reduced in 2 M and 8 M urea (Figure 3.26 C) and this reduction is represented in the native-PAGE gel by the reduced smearing seen near the wells. In the study by Affranchino & Gonzalez (1997) it was reported that VP6α, though not essential for trimerization, is key for forming contacts with VP2 to form the DLP. AlphaFold2 was used to predict quaternary conformation VP6α monomers would form and the result indicated that VP6α could potentially associate into a native-like trimeric structure (Figure 3.27 B). The AlphaFold2 prediction, native-PAGE, and the known behaviour of VP6α indicate that the domain could form various quaternary structures in solution.

3.2.5 Characterising the Conformational Stability of VP6α.

3.2.5.1 Chemical Conformational Stability
3.2.5.1.1 Conformational Stability in Urea
The conformational stability of VP6α was assessed by far-UV CD and intrinsic tryptophan fluorescence spectroscopy in 0 M – 8 M urea. The far-UV CD and fluorescence spectra are shown in Figure 3.28 (A and B). Far-UV CD spectra were recorded between 222 nm – 250 nm as this region had the best signal-to-noise ratio (Figure 3.28 A).
Figure 3.27: VP6α Quaternary Structure.

(A) Native-PAGE gel of VP6α in the presence and absence of urea. (B) The quaternary structure of VP6α (viewed down its three-fold axis) predicted using AlphaFold2 (Jumper et al., 2021; Mirdita et al., 2022) and visualised using PyMOL Version 2.5.0 (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).
Figure 3.28: VP6α Unfolding with Urea.

Far-UV CD spectra (A) and normalised intrinsic tryptophan fluorescence emission spectra (B) of VP6α recorded before denaturing (black), after denaturing with 8 M urea (red), and after refolding (blue). (C) Unfolding and refolding curves of VP6α as based on the changes in the ellipticity at 222 nm (black circles and green triangles) and the ratio of the fluorescence at 337 nm and 354 nm in response to increasing and decreasing urea concentrations (red diamonds and orange triangles). (D) HV at 222 nm and signal at 295 nm plotted against urea concentration.
There was an increase in the ellipticity as the urea concentration increased indicating the loss of secondary structures alpha helices and the formation of random coils (Kelly et al., 2005; Ranjbar & Gill, 2009; Woody, 2006; Toniolo et al., 2012; Zsila, 2022). The normalised intrinsic fluorescence spectra indicate that there is a shift in the emission maximum wavelength from 337 nm to 354 nm meaning that the tertiary structure of the domain had been denatured (Figure 3.28 B). Also shown in Figure 3.28 A and B are the CD and fluorescence spectra of VP6α that had been unfolded in 8 M urea and refolded by diluting the sample to a final concentration of 2 M urea. The fluorescence spectrum of the refolded sample overlayed the spectrum of the native sample. There was a decrease in the MRE when urea was removed by dilution indicating the recovery of alpha-helical secondary structures. To inspect the pathway by which VP6α unfolds and refolds, key wavelengths from the far-UV CD and tryptophan fluorescence were extracted from the data and converted to fraction unfolded as described in section 2.13. and plotted against the urea concentration (Figure 3.28 C). In circular dichroism, the ellipticity at 222 nm is caused by the electronic transitions within peptide bond (Kelly et al., 2005; Ranjbar & Gill, 2009; Woody, 2006; Toniolo et al., 2012; Zsila, 2022). The signal at this wavelength is highly impacted by changes in the alpha-helical content of a protein, which makes it a suitable wavelength to monitor changes in alpha-helical secondary structures in response to increasing urea concentrations (Kelly et al., 2005; Ranjbar & Gill, 2009; Woody, 2006; Toniolo et al., 2012; Zsila, 2022). Wavelengths 337 nm and 354 nm were identified as key for monitoring tertiary structural changes as these are the wavelengths that correspond to buried and exposed tryptophan residues, respectively (Vivian & Callis, 2001).

The CD and fluorescence data show similar unfolding pathways, however, the tertiary structure of VP6α appears to be destabilised in a more cooperative manner than the secondary structure. The complexity of this pathway lies in the minimal overlap in the pathways revealed by the different spectral probes. To determine whether the urea-induced denaturation of VP6α is reversible, the domain was denatured in 8 M urea and then refolded by incrementally diluting the sample to a final urea concentration of 2 M. The spectra of the final dilution are shown in (Figure 3.28 A and B) and the refolding pathways probed at key CD and fluorescence wavelengths are shown in Figure 3.28 C. The fluorescence data indicate that the pathway by which VP6α tertiary structures were recovered was similar (though not identical) to the unfolding pathway. The CD data, however, indicate that the refolding pathway was significantly different to the unfolding pathway.
Since the unfolding and refolding pathways do not overlay, it can be concluded that the chemical unfolding of VP6α in urea is a complex and irreversible process. Because of this irreversibility, as well as the complex composition of the starting sample (Figure 3.27 A), fitting the data to a two-state or multi-state model for gaining deeper insights into the thermodynamics of this unfolding process is not possible. The turbidity data obtained from CD and fluorescence are shown in Figure 3.28 D. The HV readings provide insights into the voltage needed by the detector to effectively measure the incident light passing through a sample (Rodger & Marshall, 2021; Miles et al., 2021). The fluorescence signal at 295 nm indicates the presence of large light-scattering particles in a solution. There was an increase in the HV signal as the urea concentration increased while the opposite trend was observed in the 295 nm fluorescence signal (Figure 3.28 D). This indicates that the addition of urea does not induce the formation of larger structures but rather drives the sample to a more uniform state. This observation is also supported by the native-PAGE analysis (Figure 3.27 A). The reason the HV signal increases despite the dissociation of large oligomers and aggregates is due to the signal-to-noise ratio decreasing as a result of the increasing urea concentration.

3.2.5.1.2 Conformational Stability in Guanidine Hydrochloride

The conformational stability of VP6α was assessed by far-UV CD and intrinsic tryptophan fluorescence spectroscopy in 0 M – 5 M guanidine hydrochloride (GdCl). The far-UV CD and normalised fluorescence emission spectra are shown in Figure 3.29 (A and B). Due to the high concentrations of GdCl that were used, far-UV spectra were recorded between 222 nm – 250 nm, as this region had the best signal-to-noise ratio. There was a global increase in the MRE as the GdCl concentration increased indicating the loss of native-like secondary structures and the formation of random coils (Figure 3.29 A). The fluorescence spectra show a shift in the maximum emission wavelength from 337 nm to 354 nm as the GdCl concentration increased indicating complete exposure of the tryptophan residues to the polar solvent (Figure 3.29 B). Also shown in Figure 3.29 (A and B) are the CD and fluorescence spectra of VP6α that had been unfolded in 5 M GdCl and refolded by incrementally diluting the sample to a final concentration of 1 M GdCl. The CD spectrum of the refolded VP6α had a slightly lower ellipticity at 222 nm (Figure 3.29 A), though not in perfect alignment with the native spectrum, the data does indicate the recovery of alpha-helices upon the removal of GdCl.
Figure 3.29: VP6α Unfolding with Guanidine Hydrochloride.

Far-UV CD spectra (A) and normalised intrinsic tryptophan fluorescence emission spectra (B) of VP6α recorded before denaturing (black), after denaturing with 5 M GdCl (red), and after refolding (green). (C) Unfolding and refolding curves of VP6α as based on the changes in the ellipticity at 222 nm (red circles and green triangles) and the ratio of the fluorescence at 337 nm and 354 nm in response to increasing and decreasing GdCl concentrations (black diamonds and purple triangles). (D) HV at 222 nm (red) and signal at 295 nm (black) plotted against GdCl concentration.
The fluorescence spectrum of the refolded sample aligns well with that of the native protein, indicating that the tertiary structure of VP6α was also recoverable following chemical denaturation (Figure 3.29 B). As explained in section 4.1.5.1.1, key wavelengths were selected from the far-UV CD (222 nm) and fluorescence data (337 nm and 354 nm) and plotted against the GdCl concentration to inspect the pathway by which VP6α unfolds and refolds in GdCl. The pathways revealed by CD and fluorescence are distinct in that the unfolding of the secondary structure is a cooperative process and while the unfolding of the tertiary structure involves the formation of an intermediate between 1.5 M – 2.5 M GdCl (Figure 3.29 C). The tertiary structure is also destabilised at a lower GdCl concentration (1 M) than the secondary structure (1.75 M). The unfolding of VP6α using GdCl as a denaturant reveals a complex pathway which includes the formation of a single intermediate. To determine whether the GdCl-induced denaturation of VP6α is reversible, the domain was denatured in 5 M GdCl before being incrementally diluted to 1 M GdCl. The same key wavelengths (222 nm (CD) and 337 nm and 354 nm (fluorescence)) were analysed in the reversibility study and the pathways from these points are shown in Figure 3.29 C. The CD data indicate that the refolding of VP6α in GdCl is a reversible process as the refolding pathways align closely to the unfolding pathway. The fluorescence data indicates that the intermediate detected in the unfolding pathway is not reformed as VP6α refolds. In fact, the refolding probed by fluorescence aligns with the CD unfolding/refolding trends in that it appears to be a cooperative process. Since there were multiple quaternary conformations of VP6α in the denaturant-free starting sample, it is not likely that the denaturation of the domain occurred through a single pathway which explains the discrepancies in the CD and fluorescence unfolding data. It is therefore not possible to fit the data to thermodynamic models. The turbidity data obtained from CD and fluorescence are shown in Figure 3.29 D. An increase in the HV signal and a decrease in the fluorescence intensity at 295 nm was also observed in the study with GdCl. The opposite trends observed by these probes indicate that the addition of GdCl reduced aggregation/oligomerisation and the increase in the HV signal is due to the effects of the increasing GdCl concentration on the signal-to-noise ratio.

3.2.5.2 Thermal Conformational Stability

The thermal conformational stability of VP6α was assessed by far-UV circular dichroism and intrinsic tryptophan fluorescence spectroscopy while heating the sample to 90°C in a 20 mM sodium phosphate buffer with and without urea.
This was done to compare the behaviour of VP6α in different conditions where a native-like conformation was maintained. In 0 M urea, VP6α was shown to have native-like secondary and tertiary structural features but contained large particles that were scattering the light. A decrease in the intensity at 295 nm was observed following the addition of 2 M urea which indicated that the denaturant had reduced oligomerisation or aggregation without significantly altering the structure of VP6α.

3.2.5.2.1 Thermal Conformational Stability in 0 M Urea

The thermal conformational stability of VP6α was assessed by far-UV circular dichroism and intrinsic tryptophan fluorescence spectroscopy while heating the sample from 20°C to 90°C in a Dialysis Buffer 2 (Figure 3.30 A and B).

Fluorescence and CD spectra were also recorded when the sample was cooled back to 20°C immediately after heating and after overnight incubation at 4°C. The CD spectra show that there was a decrease in MRE at 200 nm and an increase in MRE between 200 nm – 240 nm as the sample was heated (Figure 3.30 A). The decrease in MRE indicates the formation of random coils while the increase in MRE represents a loss of the native secondary structures (Greenfield, 2006; Ranjbar & Gill, 2009; Woody, 2006; Toniolo et al., 2012). When the sample was cooled to 20°C immediately after heating, the recorded spectrum overlayed the spectrum of the denatured sample. This indicated that the thermal denaturation of VP6α was not reversible in the timeframe of the experiment. To determine whether cooling the sample over a longer period would restore the native conformation, the heated sample was allowed to cool at 4°C overnight. The spectrum of the 4°C sample shows an increase in MRE between 200 nm – 206 nm and a decrease between 206 nm – 240 nm. This indicates that the overnight cooling allowed for some recovery of the alpha-helical secondary structures, but the spectrum does not resemble that of the native protein. It can therefore be concluded that the native secondary structure of VP6α cannot be recovered even when cooled at a lower temperature for a longer time.

The fluorescence spectra indicate a shift in the maximum emission wavelength ($\lambda_{\text{max}}$) from 340 nm (native) to 347 nm (at 90°C). This red shift indicates that the domain had unfolded as it was heated (Figure 3.30 B).
Figure 3.30: VP6α Thermal Studies in 0 M Urea.

Far-UV CD spectra (A) and intrinsic tryptophan fluorescence emission spectra (B) of VP6α recorded before heating (black), at 90°C (red), after cooling to 20°C (blue), and after cooling at 4°C overnight (green). (C) Unfolding and refolding curves of VP6α as based on the changes in the ellipticity at 208 nm (black) and 222 nm (red) and the fluorescence intensities at 340 nm (green) and 347 nm (blue) in response to increasing temperature. (D) HV at 208 nm (black) and at 222 nm (red) and the signal at 295 nm (green) plotted against temperature.
There is also a decrease in fluorescence intensity at 90°C which is a result of (1) the local environment of tryptophan residues becoming more polar and therefore more susceptible to solvent quenching and (2) the dissipation of the excitation energy as a result of increasing temperature (Duy & Fitter, 2006; Demchenko, 1986). The fluorescence spectrum of the sample that was cooled back to 20°C has \( \lambda_{\text{max}} \) at 342 nm and an intensity comparable to that of the 90°C sample. This indicates that the native tertiary structure was not recoverable upon cooling therefore the thermal unfolding of VP6\( \alpha \) can be concluded to be an irreversible process. The fluorescence spectrum of the sample cooled at 4°C overnight shows a blue shift in \( \lambda_{\text{max}} \) to 340 nm and an increase in intensity. This aligns well with the spectrum of the native sample, but its \( \lambda_{\text{max}} \) is slightly red-shifted relative to the native sample. The spectrum of the 4°C sample indicates that the local environment of the tryptophan residues had become less polar. To visualise the pathways by which the secondary and tertiary structures of VP6\( \alpha \) were thermally denatured, key wavelengths from CD (222 nm and 208 nm) and fluorescence (340 nm and 355 nm) were plotted against temperature (Figure 3.30 C). As mentioned previously, the ellipticity at 222 nm is used to monitor changes in the alpha helices. The ellipticity at 208 nm is another wavelength that can be used to monitor structural changes in proteins. The signal obtained at this wavelength is sensitive to backbone conformational changes occurring in both alpha-helical and non-alpha-helical regions of a protein (Ranjar & Gill, 2009; Toniolo et al., 2012; Zsila, 2022). The fluorescence wavelengths, that is, 340 nm and 347 nm, were selected to monitor tertiary structural changes because they correspond to the native and denatured conformations of VP6\( \alpha \), respectively. The CD data shows that the 208 nm ellipticity increases at a lower temperature than the 222 nm ellipticity. This means that conformational changes occurred in non-helical regions of the protein before the alpha-helices were disrupted (Figure 3.30 C). The trend in the ellipticity at 208 nm has a steeper transition from native to denatured compared to the 222 nm trend. This indicates that the destabilisation of non-alpha-helical regions of the domain is more cooperative than the destabilisation of the alpha-helices. The fluorescence data show a decrease in fluorescence intensity with increasing temperature, which is likely a result of the dissipation of the excitation energy due to the increasing temperature as well as the solvent quenching experienced by tryptophan residues as they become more exposed to the surrounding polar environment. The sample turbidity at 208 nm and 222 nm (HV) and fluorescence intensity at 295 nm (scatter), were plotted against temperature to monitor aggregation (Figure 3.30 D).
There were increases in both the HV (at 208 nm and 222 nm) and fluorescence intensity (at 295 nm) as the temperature increased which is indicative of thermally induced aggregation. This aggregation can be linked back to the fluorescence spectrum of the VP6α sample that was cooled overnight (Figure 3.30 B). There was a clear increase in fluorescence intensity and a blue shift in the maximum emission wavelength, which could be interpreted as the recovery of the native tertiary structure. However, due to the clear and extensive aggregation that occurred when the sample was heated (Figure 3.30 D) and the CD spectrum of the same sample (Figure 3.30 A), it is more prudent to conclude that the fluorescence spectrum of the overnight sample does not represent the native tertiary conformation of VP6α. Overall, it can be concluded that VP6α is irreversibly denatured when heated in the absence of urea. Thermodynamic data analysis is not possible due to the irreversible nature of the thermal unfolding pathway. Only an approximation of the melting temperature (T_m) could be extracted from the thermal data. The approximate T_m values corresponding to the global destabilisation of the VP6α polypeptide backbone (208 nm curve) and VP6α alpha-helical structures (222 nm curve) are 60°C and 70°C, respectively.

3.2.5.2.2 Thermal Conformational Stability in 2 M Urea

The thermal conformational stability of VP6α was assessed by far-UV circular dichroism and intrinsic tryptophan fluorescence spectroscopy while heating the sample from 20°C to 90°C in a 20 mM sodium phosphate buffer with 2 M urea (Figure 3.31 A and B). Fluorescence and CD spectra were also recorded when the sample was cooled back to 20°C immediately after heating and after overnight incubation at 4°C. The CD spectra were recorded from 211 nm to 250 nm, as the signal-to-noise ratio was ideal in this region.

The CD spectra show that there was a global increase in the MRE as the sample was heated (Figure 3.31 A). Notably, there was a slight increase in the ellipticity at 222 nm and a change in the shape of the spectrum. As mentioned in the section above, the 222 nm signal is sensitive to alpha-helical structures. Since there was not a significant increase in the 222 nm ellipticity at 90°C, it can be deduced that there were still alpha-helical structures present at this temperature. However, when considering the shape of the spectrum from 221 nm – 211 nm, there is a clear difference between the unheated and heated samples. The CD spectra of VP6α heated in the presence of 2 M urea, therefore, indicate that VP6α undergoes conformational rearrangements that do not significantly impact the alpha-helical content of the protein.
Figure 3.31: VP6α Thermal Studies in 2 M Urea.

Far-UV CD spectra (A) and intrinsic tryptophan fluorescence emission spectra (B) of VP6α recorded before heating (black), at 90°C (red), after cooling to 20°C (blue), and after cooling at 4°C overnight (green). (C) Unfolding and refolding curves of VP6α as based on the changes in the ellipticity at 222 nm (red) and the fluorescence intensities at 340 nm (green) and 348 nm (blue) in response to increasing temperature. (D) HV at 222 nm (black) and the signal at 295 nm (green) plotted against temperature.
The spectrum of the sample cooled back to 20°C immediately after heating, overlayed the spectrum of the denatured sample, indicating that the thermal denaturation of VP6α was not reversible in the timeframe of the experiment. There was a global increase in MRE in the spectrum of the sample cooled overnight at 4°C indicating a decrease in the alpha-helical content of VP6α. The fluorescence spectra indicate a shift in the maximum emission wavelength ($\lambda_{max}$) from 340 nm (native) to 348 nm (at 90°C). This red shift indicates that the domain had unfolded as it was heated (Figure 3.31 B). There is also a decrease in fluorescence intensity at 90°C which was explained in the previous section. The fluorescence spectrum of the sample that was cooled back to 20°C is comparable to that of the 90°C spectrum. This indicates that thermal unfolding in 2 M urea is not reversible in the time frame of the experiment. The fluorescence spectrum of the sample cooled at 4°C overnight shows a blue shift in $\lambda_{max}$ to 342 nm and an increase in intensity which indicates that the local environment of the tryptophan residues had become less polar. To visualise the pathways by which the secondary and tertiary structures of VP6α were thermally altered, key wavelengths from CD (222 nm) and fluorescence (340 nm and 348 nm) were plotted against temperature (Figure 3.31 C). The CD data show that (1) structural changes occurred from 50°C and (2) the ellipticity at 222 nm increased in a non-cooperative fashion. This means that, in these conditions, the alpha helices of VP6α were destabilised in a non-cooperative manner and partially denatured as the protein underwent conformational changes. There was a region between 54°C – 60°C where minimum fluctuations in MRE were observed which could be indicative of a thermally stable non-native structure that formed as the conformation of VP6α changed. This idea is supported by the fluorescence data as a similar region of minimal fluctuation in the fluorescence signal was detected between 52°C – 64°C. The sample HV at 222 nm and fluorescence intensity at 295 nm, were plotted against temperature to monitor aggregation (Figure 3.31 D). Both the HV at 222 nm and fluorescence intensity at 295 nm increased with the increase in temperature. The increases in the HV signal were gradual in response to the increasing temperature while the 295 nm signal increased sharply around 46°C, indicating aggregation. Since the protein started aggregating from 46°C and stable structures were detected by CD and fluorescence within the 52°C – 64°C range, it can be concluded that the stable structures detected by the respective probes formed as VP6α aggregated. When linking the aggregation data back to the fluorescence spectrum of the sample cooled overnight (Figure 3.31 B), it can be deduced that said fluorescence spectrum does not likely represent the native tertiary structure of VP6α as aggregation is known to prevent refolding. Overall, it can be concluded that VP6α is irreversibly
denatured when heated in 2 M urea. Thermodynamic data analysis is not possible due to the irreversible nature of the thermal unfolding pathway.

3.2.5.3 Differential Scanning Calorimetry
Differential scanning calorimetry was done using a Nano DSC (TA Instruments). The run was set up to heat the sample from 40°C to 94°C at a heating rate of 1°C/min. The raw data was processed and analysed using NanoAnalyze software. The thermogram obtained for VP6α had a peak at 68.55°C (Figure 3.32). This is consistent with the Tm of ~70°C determined by thermal far-UV CD spectroscopy.
Figure 3.32: DSC Thermogram of VP6α in 0 M Urea.

The thermogram indicates the $T_m$ of VP6α which is 68.55°C. The thermogram was generated by preparing the raw data using NanoAnalyse software.
Chapter 4 Discussion

In this study, the beta-sheet (VP6β) and alpha-helical (VP6α) domains of the Rotavirus capsid protein VP6 were expressed in a bacterial system following the transformation of two *E. coli* strains (BL21 (DE3) and NiCo21 (DE3)) with pET15a plasmids containing the codon optimised consensus DNA sequence of the respective domains. In a previous study by Aiyegbo *et al.*, (2013), VP6β had been successfully expressed in the BL21 (DE3) *E. coli* strain. The authors reported that, following an overnight expression at 20°C, they extracted soluble VP6β from the BL21 (DE3) cells. When replicated in this study, however, these conditions (as well as the others tested) did not result in soluble VP6β. The bacterial expression of VP6α was described in this study for the first time and it was observed that VP6α was expressed in an insoluble form in all conditions tested. The insolubility of the domains observed in this study was not surprising as the domains were derived from VP6 which is known to be expressed in the form of insoluble inclusion bodies (Aijaz & Rao 1996; Bugli *et al.*, 2014; Teng *et al.*, 2014; Kumar *et al.*, 2016; Afchangi *et al.*, 2019). The insolubility of the bacterially expressed VP6β and VP6α proved to be a surmountable challenge as the domains were easily, effectively, and inexpensively solubilized in a single freeze-thaw step in mild solubilisation conditions. Urea was noted to be paramount to the solubilisation of both VP6β and VP6α. In this study, 2 M urea was sufficient for the solubilisation of both VP6β and VP6α and as indicated by the secondary and tertiary characterisation data, this condition did not denature the protein. Urea preferentially interacts with hydrophobic regions on proteins (Bennion & Daggett 2003). In doing so, the denaturant prevents protein-protein interactions that drive oligomerisation and aggregation. When considering the domains, it is known that VP6β and VP6α are parts of a protein that (1) has the innate propensity to self-assemble (Lepault *et al.*, 2001) and (2) forms larger and more complex oligomeric structures (such as tubes and spheres) depending on buffering conditions such as pH and divalent cation concentration (Lepault *et al.*, 2001). When considering the natural tendencies of the domains, the interaction of urea with proteins, and the basic spectral characterisation data, it can be deduced that the conditions in the bacterial cell likely favoured the formation of a number of large oligomers that were made up of native-like VP6β and VP6α subunits, that is, the domains formed non-classical inclusion bodies in the cytoplasm of the *E. coli* strains. Said non-classical inclusion bodies could be dissociated chemically by altering the buffer conditions, specifically the pH and urea concentrations, and mechanically by freezing.
The conditions used to solubilise the domains are different to those described for solubilising the full-length VP6 (VP6FL) as VP6FL accumulates in classical inclusion bodies that require harsher solubilisation conditions. As expected based on the behaviour of VP6FL, the domains were expressed in an insoluble form in *E. Coli* but were unexpectedly not expressed in classical inclusion bodies. Since the domains and the full-length protein form different accumulations in the cytoplasm of host cells, it is worth comparing the yields obtained following the solubilisation and purification of these insoluble proteins. In this study pure VP6β and VP6α were obtained in yields of 13.32 mg/L and 25 mg/L, respectively while a previous study by Fingas *et al.*, 2018 reported a yield of 90 mg VP6FL from 1 L bacterial culture. The difference in yields obtained is likely linked to the type of inclusion bodies formed in the host by the domains and the full-length protein. Though classical inclusion bodies pose a challenge in solubilisation, they have the advantage of being resistant to host cell proteases (De Bernardez-Clark & Georgiou, 1991). Conversely, non-classical bodies are easy to solubilise but since they are simply accumulations of insoluble recombinant proteins in the cytoplasm, they are not protected from degradation by host cell proteases. Therefore, the low yields of the domains are a consequence of their susceptibility to degradation by *E. coli* proteases. This should be considered when the domains are expressed in future studies. Though the addition of protease inhibitors may improve the yield, studies have shown that using expression systems that have had protease-coding genes knocked out is the best way to prevent the degradation of recombinant proteins in bacterial cells (Westers *et al.*, 2008).

The results of the expression, solubilisation, and basic spectral characterisation studies did not implicate aggregation as a result of misfolding as the reason for the insoluble expression as the native conformations of VP6β and VP6α were maintained under solubilisation conditions. This is also supported by the 295 nm signal in the chemical studies with urea, which indicated a decrease in large light-scattering particles (such as oligomers or aggregates) with increases in urea concentration. However, when the domains were heated, extensive aggregation occurred rendering the thermally induced conformational changes in VP6β and VP6α irreversible upon cooling in 0 M and 2 M urea. When VP6β was heated without urea in the buffer, secondary structural changes were observed from 44°C and an intermediate was detected by far-UV CD between 60°C – 68°C (Figure 3.13). The region where the intermediate was detected aligned with the peak in the 295 nm plot which indicates the intermediate structure was formed along the aggregation pathway.
Though the addition of 2 M urea eliminated the intermediate and maintained the native structure up to a higher temperature, it did not prevent the sample from irreversibly aggregating (Figure 3.14). In the case of VP6\(_{\alpha}\), the domain demonstrated adequate thermal stability by resisting secondary structural changes up to 50°C and 44°C in 0 M and 2 M urea, respectively (Figures 3.29 and 3.30). It was clear that VP6\(_{\alpha}\) behaved differently when heated in the two conditions tested. In the absence of urea, the native alpha-helices were denatured in a single step from 50°C. Though structural changes occurred at a lower temperature in 2 M, these rearrangements appeared to have only slightly reduced the alpha-helical content of VP6\(_{\alpha}\) as the sample was heated which suggested that the addition of 2 M urea to the buffer had stabilising effects on VP6\(_{\alpha}\). In a study done on ferrocytochrome c, a small alpha-helical protein, it was reported that low urea concentrations had stabilised the protein and denaturation was only observed when the urea concentration was pushed beyond a specific threshold (Bhuyan, 2002). It is possible that the addition of 2 M urea to the VP6\(_{\alpha}\) sample created an environment wherein the alpha-helices were more resistant (but not entirely immune) to thermal denaturation. As the conformation of VP6\(_{\alpha}\) was changing in response to the increasing temperature, VP6\(_{\alpha}\) adopted a thermally stable conformation in 2 M urea but not in 0 M urea. The formation of this structure indicated that the loss of alpha-helical structures in one region of the protein did not destabilise, in a cooperative manner, the alpha-helices in other regions of VP6\(_{\alpha}\). The melting temperatures (T\(_m\)) of VP6\(_{\beta}\) (67.94°C) and VP6\(_{\alpha}\) (68.55°C) obtained in this study were not significantly lower than the reported T\(_m\) (78°C) of the full-length VP6 (Li et al., 2014). This highlights the self-stabilising nature of VP6\(_{\beta}\) and VP6\(_{\alpha}\) as protein domains and suggests that the overall thermal stability of full-length VP6 may arise from a synergistic interplay between these domains. Despite the slight decrease in the T\(_m\) of the domains, which is a natural consequence of splitting a larger protein into its smaller constituents, VP6\(_{\beta}\) and VP6\(_{\alpha}\) are thermally stable enough to maintain their native structure should they be used independently of each other. By definition, a protein domain is a highly conserved, independently folding, and self-stabilising unit of a protein (Batey & Clarke, 2008). The spectral data and thermal studies have indicated that the purified VP6\(_{\beta}\) and VP6\(_{\alpha}\) had native-like secondary and tertiary structures and were thermally stable. Though the stability of the VP6\(_{\beta}\) and VP6\(_{\alpha}\) was further demonstrated by their ability to resist major structural changes up to 2 M urea and 1 M GdCl, respectively, different unfolding pathways were revealed by the different probes.
In the case of VP6β, a loss of cooperativity was detected in both denaturants and by the different spectral probes which indicates that VP6β does not act as an independent folding unit and in this regard, VP6β does not fit the definition of a domain (Batey & Clarke, 2008). Though intermediate structures were detected in both the urea and GdCl pathways, there were differences in the (1) overall unfolding trends revealed, (2) stability of the intermediates, and (3) reversibility of unfolding in the respective denaturants. Regarding unfolding trends, it was observed that denaturation with urea required a higher concentration (4.5 M) compared to GdCl (1.25 M) to denature 50% of VP6β native structure. This was not an alarming observation because GdCl is known to be a more potent denaturant compared to urea. Though both urea and GdCl are chaotropic agents of similar structure, they interact with and destabilise proteins by different mechanisms. Urea and GdCl are both efficacious in the disruption of hydrogen bonds but these chaotropic agents otherwise have different preferences when it comes to interacting directly with the protein. Urea has been noted to preferentially bind to exposed hydrophobic patches and residues while GdCl preferentially interacts with hydrophilic residues and masks electrostatic interactions (Qasim & Taha, 2013). Therefore, the efficiency with which urea and GdCl denature a protein depends on how important certain interactions are in stabilising the native conformation. This would explain the observed differences in the unfolding pathways revealed when VP6β was treated with these denaturants. This indicates that the loss of cooperativity could be a result of changes in the interactions VP6β can make in the absence of VP6α. Regarding the stability of the detected intermediate structures, it was noted that the intermediate detected by far-UV CD in the GdCl study persists over a greater concentration range than the intermediate detected in urea. Beta-sheets are typically more hydrophobic than alpha helices and as mentioned previously, GdCl has limited interactions with hydrophobic residues (Qasim & Taha, 2013). The characteristic properties of beta-sheets and the mechanism of GdCl denaturation may explain why there were limited changes in the secondary structure over a broad range of GdCl concentrations. Regarding the irreversibility of the chemical unfolding of VP6β, it was noted that, in urea, the native VP6β conformation was not recovered by the same pathway it was lost and the intermediates detected in the unfolding did not appear when the sample was refolded. The unfolding and refolding of VP6β in GdCl was, however, closer to a reversible process. In the case of VP6α, secondary and tertiary structural probes revealed different unfolding pathways in urea and GdCl. In the study with urea, no intermediate structures were detected but the unfolding and refolding pathways detected by CD and fluorescence were non-cooperative and did not overlay.
This loss of cooperativity in urea could be attributed to the complex nature of the denaturant-free sample which was shown by native-PAGE analysis to comprise aggregates and oligomers (Figure 3.26 A). As mentioned previously, VP6α was insoluble in 0 M urea which was likely the cause of the observed aggregation. Though the addition of urea reduced aggregation, there were still structures of various sizes present in the different urea conditions. It has been previously stated that, though VP6α does not drive trimerization, it has an association domain that allows VP6FL to form contacts with the underlying VP2 layer (Affranchino & Gonzalez, 1997). Since the association domain essentially drives protein-protein interactions in the DLP, it is likely responsible for VP6α – VP6α associations that result in the formation of the observed oligomers. The absence of VP6β is a potential cause of the lack of specificity of the contacts that are formed between VP6α monomers which is why three conformations were seen in the native-PAGE analysis (Figure 3.26 A). In GdCl the denaturation of the secondary structures was a cooperative and reversible process while an intermediate was detected as the tertiary structure was denatured. Notably, the stable structure formed in GdCl had lost over 70% of the native conformation indicating that GdCl was stabilising a significantly denatured VP6α structure which was not seen again in the refolding pathway. The fact that the intermediate was not detected in the refolding pathway is what highlighted the complexity of this unfolding study. Since the starting sample had various conformations of VP6α (due to the natural tendencies of VP6α to aggregate and associate) it is unlikely that the denatured state was reached through a single pathway and that the stable region seen in the fluorescence unfolding could be representative of multiple structures that formed along multiple unfolding pathways. With GdCl being the more potent denaturant, as seen by the lower concentration (2.25 M) needed to denature 50% of VP6α native structure compared to urea (3.5 M), it was likely more effective than urea in breaking apart the aggregates and oligomers present in the denaturant-free sample that caused the variation in the pathways seen in the unfolding study. It is possible that VP6α was uniformly denatured in the 5 M GdCl that was used to prepare the starting sample of the refolding study and this uniformity likely enabled the cooperative refolding of VP6α. Finally, the chemical stability of VP6FL has not been reported, but work done on a structurally similar protein, Bluetongue Virus VP7, showed that VP7 had retained its native structure in 8 M urea (Calvo-Pinilla et al., 2014). This decreased chemical stability is likely a consequence of splitting VP6 and losing the necessary interdomain contacts that contribute to the overall stability of the domain.
Protein stability is key in the development and evaluation of protein-based therapeutics such as vaccines. As mentioned in the introduction, the VP6$\text{FL}$ is an attractive candidate for the development of novel rotavirus vaccines due to the beta-sheet domain bearing epitopes that are recognized by human antibodies (Afchangi et al., 2019). Since this study was done to gain basic insights into the structure and behaviour of VP6$\beta$ when expressed separately from VP6$\alpha$, no work was done with VP6$\text{FL}$ specific antibodies to determine whether the purified domain presents epitopes. However, this work can now be done knowing that VP6$\beta$ is a viable candidate for the development of a subunit vaccine as it is thermally stable and can maintain its secondary and tertiary structures in low denaturant concentrations. Additional vaccine studies, such as monitoring VP6$\beta$ conformation in different buffers containing the reagents that are commonly used in vaccine formulations (such as salts, stabilizers, adjuvants, preservatives, and antibiotics) can also be done. The study by Lepault et al., 2001 demonstrated that VP6$\text{FL}$ trimers can associate to form higher order structures such as aggregates, spheres, and tubes depending on the buffer conditions. These higher order structures of VP6$\text{FL}$ have applications in vaccine development and drug delivery (Shoja et al., 2022). In the case of vaccine development, it was noted that spheres formed by VP6$\text{FL}$ trimers induce the synthesis of high binding and long-lasting Rotavirus VP6$\text{FL}$ specific antibodies that are highly effective in preventing Rotavirus infections. Since a sphere is more stable than a trimer, the sphere would survive in the body for a longer time which gives the immune system more time to mount a response. Therefore, further studies can be done to determine whether VP6$\beta$ trimers also associate into higher order structures and to identify the buffer conditions that allow these structures to form. A study could also be done to determine the type of higher order structures that are formed by VP6$\alpha$. Though the higher order quaternary structures of VP6$\alpha$ would not have applications in vaccine development as the domain lacks epitopes, those structures could potentially be used as drug delivery systems. It has been noted that nanotubes formed by VP6$\text{FL}$ trimers could maintain their structure in the gastrointestinal tract making them suitable for targeted delivery of therapeutics to the gut (Shoja et al., 2022). It may therefore be worth determining whether the higher order structures formed by the individual domains could be used in a similar context.
Chapter 5 Conclusion

Approximately 13 mg and 25 mg of pure VP6β and VP6α were respectively obtained from 1 L of NiCo21 (DE3) E. coli culture following a single freeze-thaw solubilisation in non-denaturing conditions and affinity purification. The secondary and tertiary structures of pure VP6β and VP6α were native-like and thermally stable up to 44°C and 50°C, respectively. Beyond these temperatures the domains irreversibly aggregate. In the case of VP6β an intermediate forms along the thermal aggregation pathway in the absence of urea but is not seen following the addition of 2 M urea. Conversely, VP6α is thermally denatured in a single step in denaturant-free conditions but undergoes conformational changes with minimal loss of the native secondary structures when heated in 2 M urea and an intermediate structure forms as the protein aggregates. Both domains were concluded to have successfully acquired thermally stable native-like secondary and tertiary structural features independently of each other but VP6α adopted multiple quaternary conformations in the absence of VP6β. Both VP6β and VP6α also resisted structural changes up to 2 M urea and 1 M GdCl, before being denatured. VP6β was denatured by non-cooperative pathways involving the formation of intermediate structures. Native-like secondary and tertiary structures are recovered following denaturant removal in a reversible manner in GdCl, but not in urea. The distinct trends revealed by the chemical unfolding indicate that VP6β is not an independent folding unit and likely depends on key interactions with VP6α to achieve its native conformation in a cooperative manner. This highlights the importance of interdomain contacts in allowing the full-length VP6 to adjust its conformation during various stages of the Rotavirus replication cycle. The unfolding of VP6α appeared to be more cooperative and reversible in GdCl compared to urea. A stable intermediate was detected in the GdCl study by fluorescence spectroscopy however this intermediate did not appear in the refolding pathway. The complexity of this unfolding pathway was rooted in the various quaternary conformations adopted by VP6α in the absence of VP6β. The distinct unfolding trends revealed by the respective denaturants show that VP6α likely relies, to some extent, on interactions with the beta-sheet domain to adopt its native structure cooperatively. Overall, this study has shown that native-like VP6β and VP6α could be easily and cost-effectively expressed, solubilized, and purified in high enough concentrations to allow for further analysis. The domains demonstrated adequate thermal and chemical stability which highlights their suitability for use in isolation and with other reagents.
Chapter 6 References


Teng, Y, B Zhao, X Pan, Y Wen, and Y Chen. 2014. "A new Rotavirus VP6-based foreign epitope presenting vector and immunoreactivity of VP4 epitope chimeric proteins." Viral Immunology 27 (3): 96-104.


