Therapeutic and Diagnostic Potential of Rotavirus A Capsid Proteins

A Thesis Submitted in Partial Fulfilment for the award of the degree of

Doctor of Philosophy

by

Pooja Rani Kuri

DEPARTMENT OF BIOSCIENCES AND BIOENGINEERING

Indian Institute of Technology Guwahati

Guwahati-781039, Assam, India

September, 2024

To my family, supervisor, friends, and mentors—

For their unwavering support, guidance, and invaluable lessons.

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
Department of Biosciences and Bioengineering
Guwahati: 781039

STATEMENT

I do hereby declare that the matter embodied in this thesis entitled "**Therapeutic and Diagnostic Potential of Rotavirus A Capsid Proteins**", is the result of investigations carried out by me in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Assam, India, under the guidance of Prof. Pranab Goswami.

In keeping with the general practice of reporting scientific observations, due acknowledgements have been made wherever the work described is based on the findings of other investigators.

September, 2024

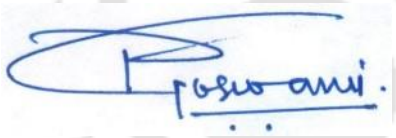
Pooja Rani Kuri

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
Department of Biosciences and Bioengineering
Guwahati: 781039

CERTIFICATE

It is certified that the work described in this thesis, entitled "**Therapeutic and Diagnostic Potential of Rotavirus A Capsid Proteins"**, done by Pooja Rani Kuri (Roll no. 176106028) for the award of degree of Doctor of Philosophy is an authentic record of the results obtained from the research work carried out under my supervision in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, India.

The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree.



Dr. Pranab Goswami

Professor (Higher Administrative Grade)

(Thesis Supervisor)

ACKNOWLEDGEMENT

"We must find time to stop and thank the people who make a difference in our lives."

- *John F. Kennedy*

The work presented herein was carried out in the Biosensors and Biofuel Cell Lab at the Indian Institute of Technology (IIT) Guwahati. The challenging yet the exciting PhD journey would not have been feasible without the intellectual, technical, and emotional support from numerous individuals who I am indebted to. Therefore, here is the chance to shower my gratitude to all of them.

First and foremost, I express my profound gratitude to my supervisor, Prof. Pranab Goswami, for providing me the opportunity to work under his guidance. His ceaseless thirst to explore the unconventional domains has greatly inspired me and shaped my perceptivity of scientific research. Besides, his robust technical expertise and ability to visualize the pros and cons of an experimental plan have assisted me during my research. I sincerely appreciate his constant support, time, ideas, and above all, providing a state-of-the-art laboratory to carry out my research work. My heartfelt gratitude goes to him for supporting me emotionally and teaching me the art of patience and perseverance. Thank you, Sir, for everything from teaching me to think like a researcher to many vital life lessons in general that have moulded me into a better person.

Next, I would like to acknowledge my doctoral committee members Dr. Biplab Bose (Chairman), Prof. Manish Kumar, and Prof. A. S. Achalkumar, for their critical comments, insightful suggestions, and encouragement that lifted me towards the accomplishment of my research objectives. I offer my sincere thanks to all the staff members of Department of Biosciences and Bioengineering, IIT Guwahati, for their limitless help and technical assistance in handling the various instruments. I would also like to acknowledge the Central Instruments Facility (CIF), Centre of Excellence (CoE) of the Department of Biosciences and Bioengineering, and Department of Chemistry, IIT Guwahati, for providing access to various instrumental facilities without which the

work would not have been complete. Further, I am tremendously grateful to IIT Guwahati, Ministry of Human Resource and Development, India and Department of Biotechnology, Government of India, for their financial assistance and funding, respectively, towards my research work.

I greatly believe that a positive work environment always expedites one's personal growth. Hence, I am thankful to all the past and present lab members of Biosensors and Biofuel Cell lab for their valuable support.

I extend my heartfelt thanks to my loved ones Dr. Kiran Vilas Dhobale, Mrs. Caraline Ann Jacob, Ms. Aradhana Priyadarsini, Mr. Rahul Pegu, Dr. Kaustubh Khaire, Mrs. Sunanda Chettri Sharma, Mr. Sibojyoti Baksi, Mr. Rohan Pal, Mr. Amay Sanjay Redkar, and Ms. Neha Mariam Unnoony. Their steadfast support during challenging times at IITG, and the joy they have brought into my life have made my journey truly meaningful.

I extend my deepest gratitude to my parents for their unconditional love, blessings, and constant encouragement in every step of the PhD journey. No amount of words can perhaps substantiate my acknowledgement to them for their endless sacrifices and their teachings to be determined, strong, and to believe in myself.

Last but not least, I humbly thank the Almighty for bestowing me the strength to keep moving forward despite several challenges.

Pooja Rani Kuri

September, 2024

TABLE OF CONTENTS

ABSTRACT

Rotavirus, a primary contributor to severe cases of infantile gastroenteritis on a global scale, causes significant morbidity and mortality in the under-five population, particularly in middle to low-income countries, including India. Effective management of rotavirus disease encompasses both suitable diagnostic strategies and effective therapeutic interventions. Implementation of WHO-approved live-attenuated vaccines in certain regions of the world brought about a significant reduction in rotavirus-led gastroenteritis hospitalizations. However, they are linked to susceptibility to intussusception and exhibit low efficacy, primarily attributed to the high genetic diversity of rotavirus, varying over time and across different geographic regions. Herein, molecular data on Indian rotavirus A (RVA) has been reviewed through phylogenetic analysis, revealing G1P[8] to be the prevalent strain of RVA in India. The conserved sequences of rotavirus capsid proteins VP7, VP4 and VP6 across G1P[8] strain were used to examine helper T lymphocyte, cytotoxic T lymphocyte and linear B cell epitopes. Twenty epitopes were identified after evaluation of factors such as antigenicity, non-allergenicity, non-toxicity, and stability. These epitopes were then interconnected using suitable linkers and an N-terminal beta defensin adjuvant. The in silico designed vaccine exhibited structural stability and interactions with integrins (αvβ3 and αIIbβ3) and toll-like receptors (TLR2 and TLR4) indicated by docking and normal mode analyses. The immune simulation profile of the designed RVA multiepitope vaccine exhibited its potential to trigger humoral as well as cell-mediated immunity, indicating that it is a promising immunogen. These computational findings indicate potential efficacy of the designed vaccine against rotavirus infection.

Among the capsid proteins, the inner capsid protein, VP6, has received much attention as a promising candidate for next-generation vaccines against rotaviruses owing to its abundance in virion particles and high conservation. However, the formation of inclusion bodies during prokaryotic VP6 expression poses a significant hurdle to its research and application. Here, we employed experimental and computational approaches to investigate its inclusion body formation and aggregation-prone regions (APRs). Heterologous recombinant VP6 expression in *E. coli* BL21(DE3) cells resulted

in inclusion body formation, confirmed by transmission electron microscopy revealing amorphous aggregates. Thioflavin T assay demonstrated incubation temperature-dependent aggregation of VP6 inclusion bodies. Computational predictions of APRs in RVA VP6 protein were performed using sequence-based tools (TANGO, AGGRESCAN, Zyggregator, Waltz, FoldAmyloid, ANuPP, Camsol intrinsic) and structure-based tools (SolubiS, CamSol structurally corrected, Aggrescan3D). Out of a total of 24 consensus APRs identified, 21 were found to be surface-exposed in the VP6 structure. All identified APRs exhibit a predominance of hydrophobic amino acids, with proportions ranging from 33% to 100%. Computational identification of these APRs corroborates our experimental observation of VP6 inclusion body or aggregate formation. Characterization of VP6's aggregation propensity facilitates understanding of its behaviour during prokaryotic expression and opens avenues for protein engineering of soluble variants, advancing research on rotavirus VP6 in pathology, therapy, and diagnostics.

The rotavirus-led fatal infantile gastroenteritis globally demands a portable, specific, and low-cost diagnostic tool for its timely detection and effective surveillance in a mass population. Consequently, the design and development of an advanced biosensing technique for its detection is of paramount importance. A highly conserved 23-nucleotide sequence, 5' GCTAGGGATAAGATTGTTGAAGG 3', was identified through analysis of human RVA VP6 gene sequences and designated as the target. A molecular beacon of 33 nucleotides was designed with the sequence 5'[Fluorescein] ATAGTCCTTCAACAATCTTATCCCTAGCACTAT[Dabcyl]3', incorporating stem and loop regions. Secondary and tertiary structure characterization confirmed the desired stem-loop structure without internal secondary structures. The thermal stability of the molecular beacon-target complex was investigated using a temperature vs. Gibbs free energy change plot, melting curve analysis based on absorbance vs. temperature, and an experimental fluorescence resonance energy transfer melting assay. The melting temperature of the molecular beacon-target complex was experimentally determined to be 62°C. Spectral analysis showed fluorescence restoration in the presence of the synthetic VP6 target. Assay conditions were optimized with an excitation wavelength of 470 nm and a 10-minute incubation time. The assay demonstrated a linear correlation between fluorescence intensity restoration and target concentration, with a limit of detection of 18.8 nM. Interference studies with single mismatch, double mismatch, and

scrambled targets revealed that the molecular beacon has strong specificity for the VP6 target, effectively discriminating against non-target sequences. This piece of work demonstrates the molecular beacon's potential as a sensitive and specific tool for detecting RVA VP6 gene, with promising applications in diagnostic assays for the rotavirus disease management.

This research enhances rotavirus disease management by leveraging the potential of capsid proteins for both therapeutic and diagnostic applications. The development of an RVA multiepitope vaccine, the probing of APRs in RVA VP6, and a sensitive molecular beacon-based genosensor for RVA detection offer promising advancements in disease prevention and detection, addressing key gaps in current strategies.

LIST OF ABBREVIATIONS

APRs	Aggregation-Prone Regions
CAI	Codon Adaptation Index
CDC	Centers for Disease Control and Prevention
CPE	Cytopathic Effects
CTL	Cytotoxic T Lymphocytes
DNA	Deoxyribonucleic Acid
DLP	Double Layered Particle
ELISA	Enzyme Linked Immunosorbent Assay
EM	Electron Microscopy
ER	Endoplasmic Reticulum
FRET	Fluorescence Resonance Energy Transfer
GIC	Golgi Intermediate Compartment
GRAVY	Grand Average of Hydropathicity
GRSN	Global Rotavirus Surveillance Network
HTL	Helper T Lymphocytes
IEDB	Immune Epitope Database and Analysis Resource
IVAC	International Vaccine Access Center
LAT	Latex Agglutination Test
LOD	Limit of Detection
MDS	Molecular Dynamics Simulation
MHC	Major Histocompatibility Complexes

MHCI	Major Histocompatibility Complex Class I
MHCII	Major Histocompatibility Complex Class II
µM	Micromolar
NCBI	National Center for Biotechnology Information
NSP	Non-Structural Protein
NPT	Number of particles (N), Pressure (P), and Temperature (T)
NVT	Number of particles (N), Volume (V), and Temperature (T)
ORT	Oral Rehydration Therapy
ORS	Oral Rehydration Salts
PCR	Polymerase Chain Reaction
PFU	Plaque-Forming Units
PAMPs	Pathogen-Associated Molecular Patterns
PME	Particle Mesh Ewald
qPCR	Quantitative Polymerase Chain Reaction
RCWG	Rotavirus Classification Working Group
RMSF	Root Mean Square Fluctuation
RNA	Ribonucleic Acid
RT-PCR	Reverse Transcription Polymerase Chain Reaction
RVA	Rotavirus A
RVAMEV	Rotavirus A Multi-Epitope Vaccine
SDS-PAGE	Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis
SNR	Signal-to-Noise Ratio
SLP	Single Layered Particle
TAP	Transporter Associated with Antigen Processing

TFE	Trifluoroethanol
Tm	Melting Temperature
TLR	Toll-Like Receptor
TLP	Triple Layered Particle
VLPs	Virus-Like Particles
VP	Viral Protein
WHO	World Health Organization
Ps	Picoseconds
Ns	Nanoseconds
Nm	Nanometer
ThT	Thioflavin T

LIST OF SYMBOLS

◦C	Degree Celsius	**mg**	Milligram
%	Percent	**mL**	Millilitre
2D	Two dimensional	**mM**	Millimolar
3D	Three dimensional	**M**	Molar
µM	Micromolar	**nm**	Nanometer
µm	Micrometre	**ns**	Nanosecond
λ	Wavelength	**R^2**	Coefficient of Determination
λ_{ex}	Excitation wavelength		
λ_{em}	Emission wavelength	**s**	Second
σ	Standard deviation	**A**	Adenine
ΔG	Change in free energy	**T**	Thymine
kcal	Kilocalorie	**G**	Guanine
kDa	Kilodalton	**C**	Cytosine

LIST OF FIGURES

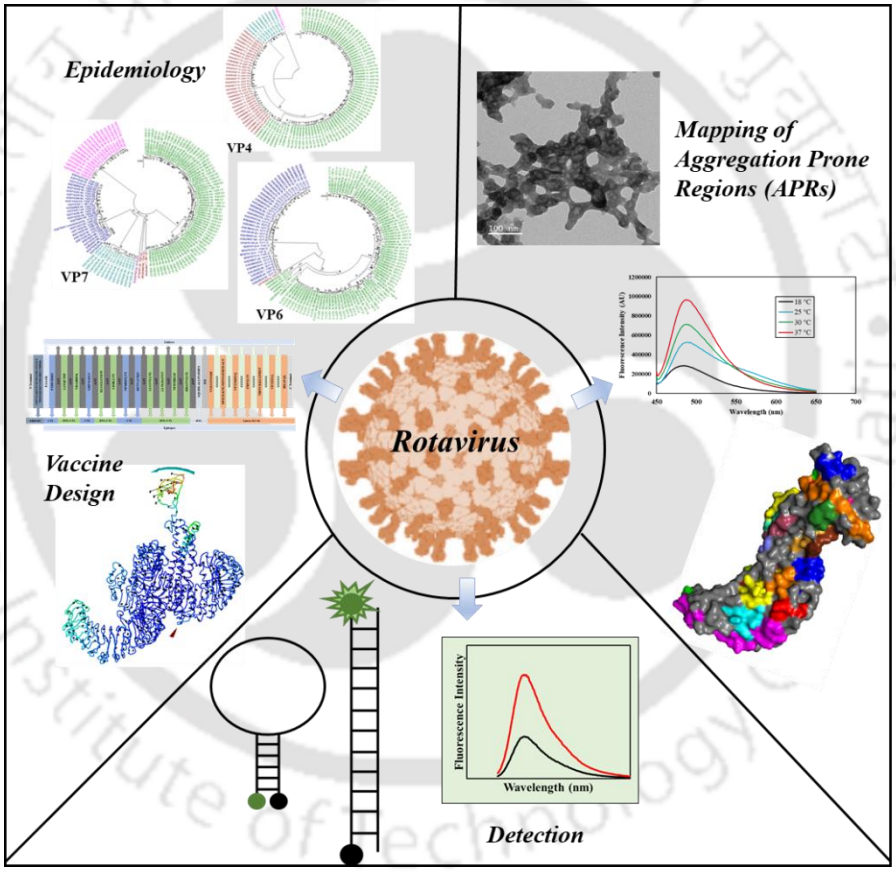
LIST OF TABLES

“Somewhere, something incredible is waiting to be known."

— Carl Sagan

Chapter 1

INTRODUCTION AND LITERATURE REVIEW



1. Introduction

Rotavirus gastroenteritis is one of the leading causes of paediatric morbidity and mortality worldwide in infants and children under five years of age. Rotavirus, as an etiologic agent of infantile gastroenteritis, was identified after a breakthrough in 1973 when this virus was discovered in faecal samples from children representing acute diarrhoea (Flewett et al., 1973) and duodenal biopsy samples representing severe diarrhoea (Bishop et al., 1973). Since then, rotavirus has been confirmed as a major cause of life-threatening infantile gastroenteritis in under-five globally and in the young of many avian and mammalian species (Estes & Greenberg, 2013).

While vaccination programs have reduced the incidence of rotavirus infections, their effectiveness remains suboptimal in low-income countries. This is due to the extensive genetic diversity of circulating RVA strains and complicating factors such as malnutrition and co-infections that limit vaccine efficacy. As a result, there is an ongoing need for improved preventive and diagnostic measures to address this public health challenge effectively.

In addition, current diagnostic methods, such as Enzyme-Linked Immunosorbent Assay (ELISA) and Polymerase Chain Reaction (PCR), are often too time-consuming and resource-intensive for use in low-resource settings where rapid detection is crucial. This gap underscores the need for more efficient and accessible diagnostic tools that can be deployed widely to enable timely detection and intervention.

To address these gaps, this research undertakes a comprehensive approach that aligns with the need for innovative solutions in both prevention and detection of RVA. By leveraging advanced methodologies and developing more effective tools, this work aims to enhance current capabilities, thereby contributing to more effective management and control of rotavirus infections, particularly in low-resource settings where the current strategies are inadequate. Adopting this dual approach could result in meaningful improvements in public health, especially in regions most burdened by rotavirus infections.

1.1. Rotavirus Structure and Classification

The complete rotavirus particle (virion) is non-enveloped and has icosahedral structure, comprising of three concentric protein layers enclosing a genome of 11 segments of dsRNA and can be structurally referred to as triple layered particle (TLP), double layered particle (DLP) or single layered particle (SLP) depending on the intactness of its protein layers. Rotavirus structure has been depicted in **Figure 1.1.** (Jiménez-Zaragoza et al., 2018).

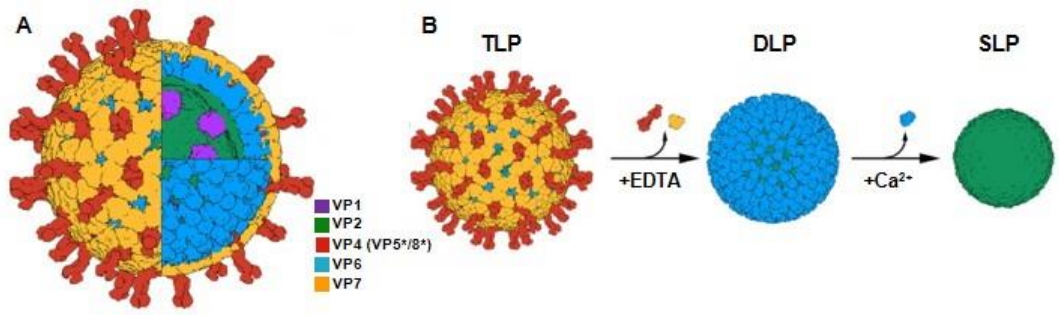


Figure 1.1. Structure of rotavirus. (A) Schematic representation of the mature rotavirus triple layered particle (TLP). (B) Double layered particle (DLP) and single layered particle (SLP) generation from TLP. VP7 and VP5*/VP8* are disassembled from TLP in the presence of EDTA. High concentration of Ca^{2+} ions takes apart VP6 trimers to liberate SLP. Adapted from Jiménez-Zaragoza et al., 2018. Copyright Jiménez-Zaragoza et al., 2018. Under creative common license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

The TLP is the fully infectious rotavirus particle and refers to the DLP covered by 260 trimers of VP7 and 60 trimers of VP4 (spike protein) (Jayaram et al., 2004). The glycoprotein VP7 and protease-sensitive VP4 together form the outer layer of rotavirus particle. Trypsin cleavage of VP4 into VP5* and VP8* enhances rotavirus infectivity (Crawford et al., 2001). The DLP consists of SLP enclosed by a layer of 260 trimers of VP6 forming the middle layer of the rotavirus particle. VP6 has been reported to make intense interactions with both, outer layer (VP4 and VP7) and inner layer (VP2), which is essential for capsid assembly and transcriptase activity (Crawford et al., 2001) (Settembre et al., 2010; Charpilienne et al., 2002). The SLP (core shell) consists of 120 molecules of VP2, which form 60 dimers. 5 of these dimers associate to form a decamer. Therefore, SLP basically consists of 12 decamers of VP2 which encapsulate the dsRNA

segmented genome and replication enzyme complexes, RNA-dependent RNA polymerase (VP1) and the capping enzyme (VP3) (Jayaram et al., 2004). This core shell is essentially uniform except for small pores along the fivefold axis (McClain et al., 2010). The segmented dsRNA genome encodes six structural viral proteins: VP1, VP2, VP3, VP4, VP6 and VP7 and six non-structural proteins: NSP1, NSP2, NSP3, NSP4, NSP5 and NSP6. The structural proteins determine host specificity, cell entry, enzymatic machinery for generation of viral transcripts and contains immunological epitopes. The non-structural proteins are produced only after entering a host cell and during the infection process. They are involved in genome replication and innate immune response antagonism (Estes, 1996). The rotavirus protein attributes have been listed in **Table 1.1** (Angel et al., 2007) for RVA. While the gene-protein assignment and protein functions remain the same for all RVA strains, the length of the coding sequence may vary slightly from one strain to another.

Rotaviruses comprise the genus *Rotavirus* belonging to the family Reoviridae (King et al., 2012). The genus *Rotavirus* has been classified serologically into ten different rotavirus species (Rotavirus groups A - J) based on genetic variability and antigenic specificity of viral capsid protein VP6 (Matthijnssens *et al.*, 2012; Mishalov-Kovács *et al.*, 2015; Bányai *et al.*, 2017). Rotaviruses have been reported to infect vertebrates (birds and mammals) (Estes and Greenberg, 2013). Group A rotavirus is the most common cause of human infantile gastroenteritis and is therefore the focus of the current investigation. According to a binary system of classification, group A rotaviruses have been classified into P genotypes (P for protease-sensitive VP4) and G genotypes (G for glycoprotein VP7) based on sequence differences in RNA segments 4 and 9, encoding VP4 and VP7 respectively (Estes and Greenberg, 2013). 32 G genotypes and 47 P genotypes of group A rotavirus have been reported till date (RCWG, 2017). Worldwide, there are six prevalent G types (G1, G2, G3, G4, G9 and G12) and three P types (P[4], P[6] and P[8]) (Matthijnssens et al., 2009; Gentsch et al., 2005; Santos and Hoshino, 2005). The most common global circulating strains of group A rotavirus that account for more than 90% of rotavirus diarrhoea cases include G1P[8], G2P[4], G3P[8], G4P[8], G9P[8] and G12P[8] (Gentsch et al., 2005; Santos & Hoshino, 2005; Matthijnssens et al., 2010). The most recent rotavirus classification system is based on phylogenetic analysis of the complete genome sequence of 11 dsRNA segments. Specific nucleotide cut-off identity percentage of all the segments VP7-VP4-

VP6-VP1-VP2-VP3-NSP1-NSP2-NSP3-NSP4-NSP5/6 is employed to determine the rotavirus strain and is represented by the constellation Gx-P[x]-Ix-Rx-Cx-Mx-Ax-Nx-Tx-Ex-Hx (Matthijnssens et al., 2008).

Table 1.1. Gene-protein assignments, location and functions of rotavirus A proteins.

Genome segment	Encoded protein	Size (kDa)	Location in virion	Molecules per virion	Functions
1	VP1	125	Core	12	Binds to ssRNA, exhibits RdRp activity, participates in the formation of transcription complex with VP3
2	VP2	94	Core	120	Exhibits RNA-binding activity; involved in replicase activity of VP1.
3	VP3	88	Core	12	Exhibits activity of Guanylyltransferase; methyltransferase, 2,5-phosphodiesterase; binds to ssRNA binding, participates in virion transcription complex with VP1.
4	VP4	86	Outer layer	180	Natively exists as homotrimer; P type neutralization antigen, acts as attachment protein, trypsin cleavage give rise to VP5* and VP8*, virulence, role in fusion with cell membrane
5	NSP1	58	Non-structural	-	Associates with the cytoskeleton, interferon antagonist, E3 ubiquitin ligase; exhibits RNA binding activity.
6	VP6	44	Middle Layer	780	Natively exists as homotrimeric structure; major virion protein; middle capsid structural protein; subgroup antigen; protection (intracellular neutralization); required for transcription.
7	NSP3	34	Non-structural	-	Natively exists as homodimer, specifically binds 3′-end of rotavirus mRNA, binds to elongation factor eIF4G1, participates in translational regulation.
8	NSP2	36	Non-structural	-	Natively exists as octamer, exhibits RNA binding activity, NTPase, NDP kinase, involves in helix destabilizing; involved in viroplasm formation, essential role in dsRNA synthesis.
9	VP7	37	Outer layer	780	Natively exists as homotrimer, located in outer capsid, G type neutralization antigen, RER transmembrane calcium-binding protein.
10	NSP4	20	Mainly non-structural	Very few	Viral enterotoxin (secreted), acts as a receptor for retreival of double-layered particles through the ER membrane, glycoprotein, modulates intracellular calcium levels and RNA replication, virulence
11	NSP5	21	Non-structural	-	Natively exists as dimer, phospho- and O-glycosylated protein, interacts with NSP2 and NSP6; binds ssRNA, component of viroplasm, essential for viral replication.
11	NSP6	12	Non-structural	-	Produced from second ORF of segment 11, exhibits interaction with NSP5, localizes to the viroplasm.

1.2. Mechanisms of Rotavirus Replication

The replication cycle of rotavirus is a complex, multi-step process involving several stages: entry into the host cell, uncoating, transcription, translation, replication, assembly, and release (Desselberger, 2014). The replication cycle of rotavirus has been elucidated in **Figure 1.2** (Papa et al., 2021).

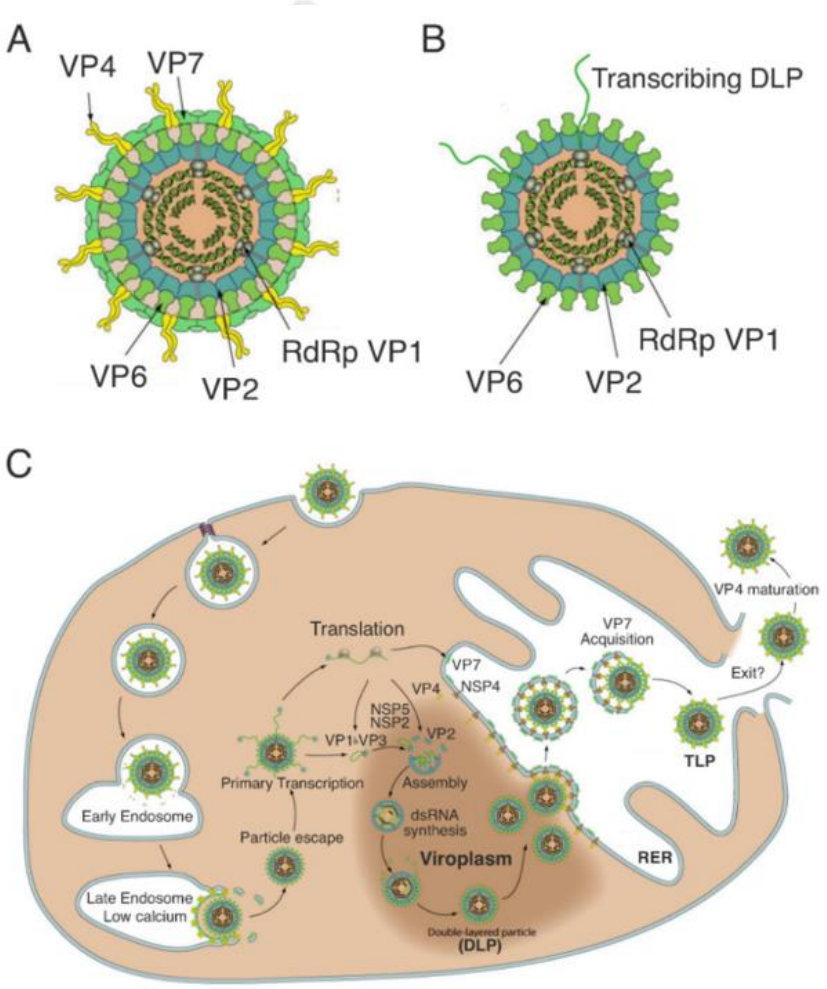


Figure 1.2. Rotavirus structure, replication cycle. (A,B) Schematic representation of triple-layered and double-layered rotavirus particles (TLPs and DLPs, respectively). Outer layer: VP7, VP4 (spikes); middle layer: VP6; inner layer (core): VP2. The core contains 11 segments of dsRNA (diagrammatically shown as concentric rings) and the viral transcription–replication complex, consisting of the RNA-dependent RNA polymerase (RdRp), VP1, and the capping enzyme, VP3. Rotavirus DLPs are transcriptionally active, producing (+)ssRNAs from each genomic RNA segment. (C) Rotavirus replication cycle (diagram) including formation of viroplasms. Copyright

1.2.2. Viral Entry

The rotavirus initiates infection by attaching to the host cell surface. This process is mediated by interactions between viral proteins, particularly VP4 and VP7 and specific cell surface receptors such as integrins and sialic acid-containing glycans. Once attachment occurs, the virus undergoes endocytosis, often via clathrin-mediated pathways. Within the endosomal compartment, the low pH triggers conformational changes in the outer capsid proteins, leading to the release of the viral particle into the cytoplasm.

1.2.3. Uncoating and Transcription

Upon entry into the cytoplasm, the outer layer of the rotavirus capsid, composed primarily of VP7, is removed, exposing the DLP. The DLP serves as the transcriptional machinery within the host cell cytoplasm. The viral RNA-dependent RNA polymerase, VP1, which is located inside the DLP, initiates the transcription of the dsRNA genome segments into positive-sense single-stranded RNA (+ssRNA). These transcripts are then extruded into the cytoplasm through channels in the DLP, where they serve dual roles as mRNA for protein synthesis and as templates for genome replication.

1.2.4. Translation and Replication

The +ssRNA transcripts are translated by the host cell's ribosomes into viral proteins, including structural proteins (VP1-VP4, VP6, and VP7) and non-structural proteins (NSP1-NSP6). The non-structural proteins, particularly NSP2 and NSP5, are crucial in forming viral replication factories known as viroplasms, where genome replication and assembly occur. Within the viroplasms, the +ssRNA serves as a template for the synthesis of complementary negative strands by VP1, resulting in the formation of dsRNA genome segments.

1.2.5. Assembly

The newly synthesized dsRNA segments are then packaged into assembling DLPs within the viroplasms. The DLPs are subsequently coated with VP6, forming a stable inner capsid structure. These DLPs can either remain in the cytoplasm, continuing to

produce +ssRNA, or acquire a transient outer capsid layer composed of VP4 and VP7. The complete triple-layered virion is formed when the DLPs are translocated to the endoplasmic reticulum (ER) where they interact with NSP4, a viral protein that functions as an intracellular receptor for the DLPs. NSP4 facilitates the acquisition of the outer capsid layer, leading to the formation of mature infectious virions.

1.2.6. Release

Mature virions are released from the host cell through a non-lytic pathway involving the ER-Golgi intermediate compartment (GIC). The exact mechanism of rotavirus egress is not fully understood, but it is thought to involve the budding of virions into the ER-GIC, followed by vesicular transport to the cell surface. Some studies suggest that rotavirus may also be released via cell lysis, particularly during later stages of infection when the viral load is high.

The rotavirus replication cycle is a highly orchestrated process that enables the virus to efficiently hijack the host cell's machinery for its own replication. Understanding the intricacies of this cycle is crucial for developing targeted antiviral therapies and improving current vaccination strategies. Despite the availability of vaccines, rotavirus continues to pose a significant public health challenge, particularly in low-resource settings. Ongoing research into the molecular mechanisms of rotavirus replication and pathogenesis will be vital in mitigating its impact on global child health.

1.3. Rotavirus Disease Burden

Rotavirus infection had been associated with 114 million positive cases with 24 million cases requiring outpatient visits, 2.3 million requiring hospitalization for the reported period of 1986 to 2004 and more than 200,000 deaths in under-five globally as estimated in the year 2013 (Tate et al., 2016). In the year 2015, diarrhoea mortality in under-five was estimated to be 500,000 globally, (Troeger et al., 2017; Wang et al., 2016), of which 29.3% of deaths were accounted for by rotavirus infection. In 2016, more than 248 million cases of under-five associated rotavirus infection were reported, requiring 1,537,000 hospitalizations with an estimated 128,500 deaths, which accounted for about 28.8% of total diarrhoea deaths in under-five in that year globally. While rotavirus associated mortality and morbidity rate was found to be highest in sub-Saharan Africa, Southeast Asia, and South Asia, rotavirus associated morbidity was also found to be magnified in high-income countries viz. the United States with 593,000

episodes in under-five (Troeger et al., 2018). Although prevalence of rotavirus infection and associated rotavirus diarrhoea hospitalizations is ubiquitous on a global scale, more than 90% of cases presenting fatal rotavirus infection are from low-income countries, probable contributing factors being limited access to proper sanitation, health care, hydration therapy, and greater prevalence of comorbid conditions and others (Crawford et al., 2018). The introduction of rotavirus vaccination in several countries to alleviate the disease burden of rotavirus diarrhoea has been reported to avert an estimated 28,800 deaths following vaccination of 27.8% of under-five children in 2016. Under-five diarrhoea associated mortality has reduced by more than 45% since 2005 (Naghavi et al., 2017) probably due to improvements in water and sanitation, reduction in comorbid conditions like undernutrition and inclusion of rotavirus vaccination in several countries.

1.4. Rotavirus Therapy

Current therapeutic approaches to managing rotavirus infections focus on supportive care, antiviral agents, and vaccine strategies.

1.4.1. Supportive Care

The main goal of supportive care is to manage dehydration, which is the leading cause of rotavirus-associated mortality. Oral rehydration therapy (ORT) using oral rehydration salts (ORS) is the cornerstone of supportive care, recommended by the WHO for the management of acute diarrhoea. ORT effectively restores fluid and electrolyte balance, reduces the need for intravenous fluids, and lowers the risk of complications. In severe cases where ORT is insufficient or impractical, intravenous fluid replacement may be necessary (Leung & Barankin, 2006). Zinc supplementation is another crucial component of supportive care, particularly in developing countries. Studies have shown that zinc can reduce the duration and severity of diarrhoea and is recommended for children with acute diarrhoea, including rotavirus infections. Additionally, continued feeding during the illness is encouraged to prevent malnutrition and promote recovery (Walker & Black, 2010).

Despite the effectiveness of these supportive measures, they do not directly target the virus, and the management is mainly symptomatic. Therefore, there is an ongoing need to develop specific antiviral therapies and improve vaccination strategies to control rotavirus infections more effectively.

1.4.2. Antiviral Agents

There are currently no specific antiviral drugs approved for treating rotavirus infections, but several agents have shown promise in preclinical and clinical studies. Nitazoxanide is one such agent that has demonstrated activity against rotavirus. It is an antiparasitic drug with broad-spectrum antiviral properties, inhibiting viral replication by interfering with the maturation of viral particles. Clinical trials have shown that nitazoxanide can reduce the duration of rotavirus diarrhoea in children, making it a potential candidate for antiviral therapy. However, more extensive clinical studies are needed to confirm its efficacy and safety (Mahapatro et al., 2017).

Other investigational antiviral therapies include siRNA-based therapies and small-molecule inhibitors targeting viral proteins. These approaches aim to inhibit viral replication and mitigate disease severity. However, challenges such as delivery methods, potential toxicity, and the rapid mutation rate of rotavirus complicate their development (Arias et al., 2004).

1.4.3. Rotavirus Vaccines

Several live-attenuated rotavirus vaccines have been developed and licensed for use in immunization programs over the years. **Table 1.2** summarizes the licensed oral rotavirus vaccines as well as other rotavirus vaccines under development or evaluation (Kirkwood et al., 2019; Burke et al., 2019; Lee & Lee, 2021; Roberto et al., 2021). In 2009, the WHO recommended incorporation of rotavirus vaccine in national immunization programs globally to alleviate the burden of the disease (WHO, 2013). About 114 countries had successfully introduced rotavirus vaccination as a part of their national immunization programmes by the beginning of the year 2022 (Varghese et al., 2022). Rotashield was the first approved rotavirus vaccine introduced in the United States in 1998. However, it was found to be associated with intussusception, resulting in its expeditious withdrawal in 1999 (Centre for Disease Control and Prevention (CDC), 1999) Rotavirus vaccines prequalified by WHO and currently in use include Rotarix (GlaxoSmithKline Biologicals; prequalified in 2009), RotaTeq (Merck & Co., Inc.; prequalified in 2008), Rotavac (Bharat Biotech, Hyderabad, India; prequalified in 2018), and ROTASIIL (Serum Institute of India PVT. LTD., Pune, India; prequalified in 2018) (Burke et al., 2019). Rotavac and ROTASIIL are currently operative in India. As of 2018, Rotarix was included in the national immunization programs by 74

countries, while 14 other counties were using RotaTeq, and 9 other nations included both Rotarix and RotaTeq (IVAC, www.view-hub.org). According to a report utilizing 9 years of data from the WHO coordinated Global Rotavirus Surveillance Network (GRSN), a reduction in acute rotavirus gastroenteritis in WHO regions was observed to range from 26.4% in the Eastern Mediterranean Region to 55.2% in the European Region. 38% of annual acute gastroenteritis cases were found to be rotavirus positive in countries that had not included rotavirus vaccination in their national immunization programs, whereas 23.0% of cases were found to be rotavirus positive for countries that had incorporated rotavirus vaccination in their national immunization programs for under-five, suggesting significant and consistent reduction in hospital admissions for rotavirus-associated gastroenteritis post rotavirus vaccine introduction (Aliabadi et al., 2019). However, the efficacy of licensed vaccines remains suboptimal in low-income countries which still witness more than 200,000 rotavirus associated deaths annually (Tate et al., 2016). For example, the median vaccine effectiveness for corresponding Rotarix and RotaTeq have been reported as 57% and 45% in countries with high rates of rotavirus gastroenteritis, most of which are low to middle-income countries, while it is 84% and 90% for those countries with low rates of rotavirus gastroenteritis (Jonesteller et al., 2017). The cause of such observational discrepancy in vaccine effectiveness is not well understood. However, several contributing factors for the suboptimal vaccine efficacy may be attributed to inhibitory effect of maternal antibodies, co-infections, nutritional deficiencies, gut dysbiosis and genetic factors (Parker et al., 2018; Desselberger, 2017). One of the important underlying reasons concerning variation in vaccine efficacy is rotavirus diversity. Rotaviruses showcase a remarkable degree of antigenic and genetic diversity, attributable to the segmented nature of their genome and interspecies transmission. This facilitates genetic reassortment during co-infection, giving rise to numerous genotype combinations and the emergence of novel genotypes (Martella et al., 2010). Among the globally prevalent rotavirus strains and other less prevalent rotavirus strains, there is a divergence in rotavirus genotypes from one region of the globe to another. As a result, the effectiveness of the existing rotavirus vaccines differs in various geographical regions (Iturriza-Gómara et al., 2011). The challenge of region-specific low efficacy caused by rotavirus diversity, the risk of intussusception, and limitations in vaccinating immunocompromised individuals linked to live-attenuated vaccines underscores the need for ongoing research, design and development of novel rotavirus vaccines. An

account of advantages and disadvantages of different vaccine platforms is given in **Table 1.3**.

Table 1.2. Rotavirus vaccines: licensed and under evaluation.

Vaccine type	Manufacturer/Developer	Composition of strains/antigen	Status
Oral and live-attenuated			
RotaTeq	Merck & Co., Inc.	Human G1, G2, G3, G4, P [8] on bovine rotavirus (WC3) backbone: Human-bovine mono-reassortants, (pentavalent)	WHO pre-qualified
Rotarix	GlaxoSmithKline Biologicals	Human rotavirus G1P[8] (RIX4414) (monovalent)	WHO pre-qualified
Rotavac	Bharat Biotech International Ltd.	Natural human-bovine reassortant: G9P[11] (116E), (monovalent)	WHO pre-qualified
Rotasiil	Serum Institute of India Pvt. Ltd.	Human-bovine mono-reassortants (BRV-PV): human G1, G2, G3, G4, G9 on bovine rotavirus (UK) backbone (pentavalent)	WHO pre-qualified
Rotavin-M1	POLYVAC	Human rotavirus G1P[8] (monovalent)	Nationally licensed (Vietnam)
Lanzhou Lamb Rotavirus Vaccine	Lanzhou Institute of Biological Products	Lamb rotavirus G10P[15] (monovalent)	Nationally licensed (China)
RV3-BB	PT BioFarma	Neonatal human rotavirus G3P[6] (RV3) (monovalent)	Phase 2/3
Non-replicating and parenteral			
P2-VP8*	PATH SK Bioscience	Subunit vaccine: tetanus toxoid P2 fused to P[4] (DS1), P[6] (1076), and P[8] (Wa) VP8* (trivalent)	Phase 3
CDC-9	CDC; Serum Institute of India Pvt. Ltd; Zhifei Lvzhu Biopharmaceutical Co., Ltd.	Inactivated human rotavirus CDC-9 (monovalent)	Preclinical
116E	Bharat Biotech International Ltd.	Inactivated human rotavirus 116E (monovalent)	Preclinical
VLP VP2/6/7	Mitsubishi/Medicago	Virus-like particle	Phase 1
VLP VP2/4/6/7	Baylor College of Medicine	Virus-like particle	Preclinical
VP6 + norovirus	Tampere University, Finland	Nanoparticle: VP6 nanotubes or microspheres and norovirus viral-like protein admixture	Preclinical
VP6 + norovirus	Cincinnati Children's Hospital Medical Center	Sub-viral particles (norovirus S or P) expressing rotavirus VP8* particle	Preclinical
VP8* mRNA	CureVac	VP8* mRNA	Preclinical

Table 1.3. Advantages and disadvantages of different vaccine platforms. Reproduced from Li et al., 2020. Copyright 2020 Li et al., 2020. Under creative common license (http://creativecommons.org/licenses/by/4.0/).

Vaccine type	**Advantages**	**Disadvantages**	**Clinically approved examples**
Whole inactivated virus vaccine	Stronger immune response; Safer than live attenuated virus	Potential epitope alteration by inactivation process	Typhoid, Cholera, Hepatitis A virus, Plague, Rabies, Influenza, Polio (Salk)
Live attenuated virus vaccine	Stronger immune response; Preservation of native antigen; Mimicking natural infection	Risk of residual virulence, especially for immunocompromised people	Measles, Mumps, Polio (Sabin), Rotavirus, Yellow Fever, Bacillus Calmette–Guérin (BCG), Rubella, Varicella
Viral vector vaccine	Stronger immune response; Preservation of native antigen; Mimicking natural infection	More complicated manufacturing process; Risk of genomic integration; Response dampened by preexisting immunity against vector	Ebola virus
Subunit vaccine	Safe and well-tolerated	Lower immunogenicity; Requirement of adjuvant or conjugate to increase immunogenicity	Pertussis, Influenza, *Streptococcus pneumoniae*, Haemophilus influenzae type b
Viral-like particle vaccine	Safe and well-tolerated; mimicking native virus conformation	Lower immunogenicity; More complicated manufacturing process	Hepatitis B virus, Human Papillomavirus
DNA vaccine	Safe and well-tolerated; Stable under room temperature; Highly adaptable to new pathogen; Native antigen expression	Lower immunogenicity; Difficult administration route; Risk of genomic integration	-
RNA vaccine	Safe and well-tolerated; Highly adaptable to new pathogen; Native antigen expression	Lower immunogenicity; Requirement of low temperature storage and transportation; Potential risk of RNA-induced interferon response	-

1.4.3.1. Subunit Vaccine

Subunit vaccines are a sophisticated class of vaccines designed to stimulate an immune response by presenting specific components of a pathogen rather than the entire organism. Unlike traditional whole-cell or live-attenuated vaccines, subunit vaccines utilize purified antigens—typically proteins or polysaccharides—that are integral to the pathogen but devoid of other potentially harmful elements. This approach minimizes the risk of adverse reactions while focusing the immune system on critical targets. By including particularly, the most relevant parts of the pathogen, subunit vaccines can provide robust and specific protection with a favourable safety profile. This methodology is particularly advantageous for pathogens with diverse strains, as it allows for the inclusion of multiple epitopes to enhance cross-protection (Moyle & Toth, 2013).

1.4.3.1.1. Multiepitope Subunit Vaccine

A multiepitope subunit vaccine is an advanced variant of subunit vaccines that incorporates multiple epitopes from various parts of an antigenic protein or across different antigens. This design aims to elicit a comprehensive immune response by targeting several antigenic sites simultaneously. The inclusion of multiple epitopes can improve the vaccine's efficacy and breadth of protection, particularly against pathogens with substantial genetic variability.

For rotavirus (RVA), a multiepitope subunit vaccine could integrate epitopes from both the outer and inner viral proteins, thereby enhancing the immune response against a broader range of rotavirus strains. This approach is advantageous for rotavirus due to its genetic diversity and the presence of multiple serotypes. By combining epitopes from conserved proteins such as VP6 with those from variable proteins like VP4 and VP7, a multiepitope vaccine could potentially offer cross-protection against various rotavirus genotypes and serotypes. Such a vaccine design would address the limitations of current vaccines, which may not cover all prevalent strains due to their limited epitope profiles (Kuri & Goswami, 2023).

1.4.3.1.2. Rotavirus VP6 Protein as a Subunit Vaccine Candidate and Detection Target

Rotavirus VP6 has been acknowledged as a promising non-live next-generation vaccine candidate against rotavirus owing to its attributes of high abundance in the virion particle, high conservancy, and immunogenic nature conferring heterotypic protection (Tamminen et al., 2013; Schwartz-cornil et al., 2002). Diverse vaccine modalities encoding RV VP6 antigen, encompassing DNA vaccines (Afchangi et.al., 2019; Chen et al., 1998; Jalilvand et al., 2015), subunit vaccines incorporating recombinant VP6 protein (Kuri & Goswami, 2023) and self-assembled structures have been documented to elicit immune responses or confer protection in animal models (Afchangi et al., 2017; Feng et al., 2017;Lappalainen et al., 2015; Vega et al., 2012). VP6 self-assembled structures, particularly VP6 nanotubes/Virus-Like Particles (VLPs), demonstrate superior immunogenicity than that of VP6 monomers or trimers, eliciting strong immune responses (Tamminen et al., 2013). Additionally, they exhibit dual functionality as adjuvants and carriers for immunological applications (Blazevic et al., 2011). Advancements in the production and purification of VP6 tubes/VLPs have enabled exploration of their multifaceted potential as immunogens, adjuvants, delivery vehicles, and nano-biomaterials (Shoja et al., 2022). Furthermore, the rotavirus VP6 antigen is utilized as a detection biomarker for the purpose of rotavirus diagnosis (Gautam et al., 2015; Kaplon et al., 2015).

Therefore, recombinant expression of VP6 has been of enormous importance to investigate rotavirus pathogenesis, development of vaccines, and diagnostic potentials. Several expression systems, including prokaryotic *E. coli* (Afchangi et al., 2017; Nanobiotechnol et al., 2019; Zhao et al., 2011) and various eukaryotes such as fungi (Bredell et al., 2016), baculovirus-silkworm systems (Kato et al., 2023), insect cell/baculovirus systems (da Silva Junior et al., 2012; Lappalainen et al., 2016), mammalian cells (da Silva Junior et al., 2012), and plant cells (Zhou et al., 2010; Chung et al., 2000; O'Brien et al., 2000) have been explored for the production of VP6. However, the prokaryotic expression of heterologous proteins remains to be the most convenient and low cost method for laboratory as well as technological purposes (Yin et al., 2007). Ironically, prokaryotic VP6 expression leads to the formation of inclusion bodies (Afchangi et al., 2017; Zhao et al., 2011), which represents a significant obstacle to VP6 protein research and applications.

1.5. Rotavirus Diagnostics

Accurate detection of rotavirus is critical because the symptoms of rotavirus-induced diarrhoea are clinically indistinguishable from those caused by other etiological agents (Glass et al., 2009), necessitating a precise identification for appropriate medical intervention. Early detection is essential for timely treatment, preventing severe dehydration and complications in affected children. Additionally, robust RV detection is vital for surveillance studies (Manouana et al., 2021; Omatola & Odaibo, 2016), enabling the monitoring of virus circulation (Gentsch et al., 2005), vaccine effectiveness (Carvalho & Gill, 2019), and the identification of emerging strains (Rojas et al., 2019) (Malik et al., 2020), all of which are crucial for public health management and the development of targeted prevention strategies. Rotavirus detection techniques encompass a range of methods including electron microscopy (EM), cell culture, immunoassays like ELISA, molecular techniques such as (PCR) and reverse transcription polymerase chain reaction (RT-PCR), and innovative biosensing platforms.

1.5.1. Traditional Rotavirus Detection Techniques

In the field of rotavirus research, EM has been one of the earliest and most effective tools for the direct detection and counting of viral particles since the discovery of rotavirus in 1973. The use of EM has provided crucial insights into the structure and morphology of Rotavirus particles, significantly advancing our understanding of the virus (Ong & Chandran, 2005). However, despite its historical importance, the use of EM has become less common in recent years. This decline is due to its limitations, including low sensitivity, time-consuming processes, high costs of instrumentation, and the requirement for highly trained technicians (Babaei et al., 2022).

Cell culture methods are crucial for rotavirus detection, utilizing specific cell lines for viral propagation and isolation. The process involves infecting cells, observing cytopathic effects (CPE), and quantifying viral presence using plaque-forming units (PFU) and Median Tissue Culture Infectious Dose ($TCID_{50}$). Despite its utility, this method has limitations including time consumption, low sensitivity, and susceptibility to contamination. Many enteric viruses, including Rotavirus, may not display visible CPE, underscoring the need for complementary diagnostic techniques (Babaei et al., 2022).

Immunoassay methods are valuable tools for detecting Rotavirus and other pathogens. These assays utilize antigens and antibodies (Peruski & Peruski, 2003) to identify viral components, with ELISA being the most widely used due to its reliability and ease. However, ELISA has limitations including potential issues with sensitivity, cross-reactivity, time consumption, cost, and the need for skilled personnel. Similarly, the latex agglutination test (LAT) (Xiang et al., 2020) offers a rapid detection method but also has drawbacks, such as lower sensitivity, lack of quantitative data, and labour-intensive procedures (Rovida et al., 2013).

Molecular techniques like PCR, real-time PCR, and RT-PCR, which amplify viral nucleic acids to detectable levels, have become essential for diagnosing enteric viruses due to their high sensitivity and accuracy. These methods, now the gold standard for viral detection, allow for simultaneous identification of multiple viruses in clinical samples. However, they require specialized equipment, involve time-consuming pre-PCR steps, and cannot differentiate between live and dead pathogens. Consequently, there is increasing interest in developing small, smart biosensing platforms that offer rapid, cost-effective alternatives for virus detection (Guliy et al., 2019).

1.5.2. Rotavirus Biosensing Techniques

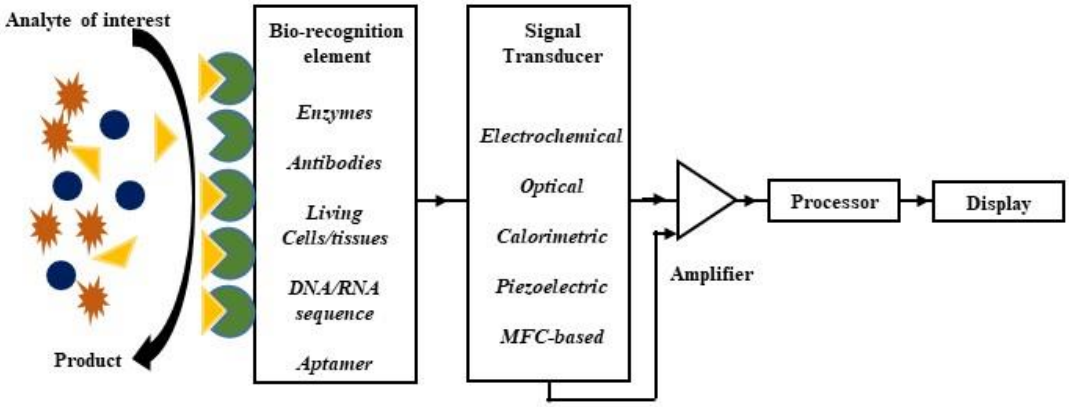


Figure 1.3. Schematic diagram of a biosensor. The diagram depicts the core elements and functionality of a general biosensor. It includes the biosensor's sensing element, which interacts with the target analyte, and the transduction mechanism that converts this interaction into a measurable signal.

A biosensor is a compact analytical device that integrates a biological recognition element, such as an antibody, enzyme, and nucleic acid, with a transducer to detect specific target molecules with high sensitivity and specificity. Schematic diagram of a biosensor is shown in **Figure 1.3**. These devices are characterized by their rapid response time, ability to provide real-time results, and cost-effectiveness, making them ideal for virus detection. Additionally, biosensors are portable, user-friendly, and capable of operating in resource-limited settings, enabling early diagnosis and real-time monitoring of viral infections. As the demand for efficient and accessible diagnostic tools grows, the versatility and adaptability of biosensors make them increasingly popular in both clinical and point-of-care diagnostics (Guliy et al., 2019; Kuri et al., 2020). Despite the significant progress made on the real-world application of biosensors in diverse fields, its development for RVA detection remains limited. Few biosensors with different platforms for rotavirus detection have been reported in the literature, which have been summarized in **Table 1.4**.

Barring a few, rotavirus biosensors reported so far rely on antibodies as bioreceptor molecules (Babaei et al., 2022). While antibodies are specific (Sharma et al., 2016), they encounter several drawbacks such as, low stability, cross-reactivity leading to false positive results, and high cost of their production. Furthermore, the production of antibodies is a complex process, requiring specialized techniques and facilities, and their efficacy generally diminishes over time due to a limited shelf life if not stored under optimal conditions (Mili et al., 2024). In contrast, nucleic acids, present several advantages as bioreceptors in biosensing applications (Elskens & Madder, 2021) (Janssen et al., 2013). They exhibit greater stability at room temperature compared to antibodies, with reduced sensitivity to environmental factors such as temperature and pH. Their specificity is achieved through precise complementary base pairing, enabling accurate detection of target sequences and distinguishing between different mutations (Xi et al., 2021). Additionally, nucleic acids can be utilized in various forms, such as aptamers (Armstrong & Strouse, 2014), complementary probes (Elskens & Madder, 2021), DNAzymes (Ma & Liu, 2020), RNAzymes (Ma & Liu, 2020), or molecular beacons (Bidar et al., 2021), offering versatility in biosensor design. Their production is generally more cost-effective than that of antibodies, and they provide consistent performance with minimal batch-to-batch variation. The ease with which nucleic acids

can be engineered or modified allows for tailored biosensor functionalities, enhancing sensitivity and enabling the detection of low concentrations of targets (Hua et al., 2022).

Table 1.4. Overview of Biosensor Platforms for Rotavirus Detection and Their Performance Metrics.

Biosensor Type	Detection Target	Bioreceptor	Sample	Sensitivity/Detection Limit	Reference
Photonic crystal biosensor	Porcine rotavirus (PoRV)	PoRV-specific antibody	Processed faecal samples from infected piglets	36 FFU 30-minute assay	(Pineda et al., 2009)
F0F1-ATPase molecular motor biosensor	Rotavirus and hepatitis A virus (HAV)	Biotin-labelled DNA probes	Extracted mRNA	RV: 0.005 ng/mL RNA, HAV: 0.01 ng/mL RNA 1 h-assay	(Zhang & Zhao, 2015)
3D photonic crystal biosensor	Rotavirus	Hydrogel-based nanopores with Rotavirus antibody	Rotavirus antigen, strain SA11	6.35 to 1.27 mg/mL rotavirus particles	(Maeng et al., 2016)
Molecular beacon assay	culture-adapted human rotavirus A (HuRVA) strains (RotaTeq vaccine and K8 strains)	Molecular beacon (DNA probe)	HuRVA in cell culture	Varies with MOI: MOI 0.5 (1.25 × 105 PFU/mL) to MOI 1 (2.5 × 105 PFU/mL)	(Bertol & Gatti, 2016)
Fluoroimmunomagnetic microparticles	Rotavirus particles in Rotateq vaccine	Monoclonal antibodies	Rotavirus particles in distilled and drinking water samples	10 rotavirus particles/L 2 h-assay	(Villamizar-Gallardo et al., 2017)
Field-effect transistor (FET) biosensor	Rotavirus	Anti-rotavirus antibodies		Linear range: 10^1 to 10^6 particles/mL	(Pant et al., 2017)
Colloidal gold immunochromatographic strip	Bovine Rotavirus (BRV)	BRV-specific monoclonal and polyclonal antibodies	BRV cell culture samples and clinical samples	1×10^3 $TCID_{50}$/0.1 mL	(Li et al., 2019)
Biolayer interferometry	Rotavirus particles in Rotarix vaccine	Glycoconjugate receptors and monoclonal antibodies	Rotarix vaccine samples from different production batches	Linear range: 2.5×10^7 to 1.0×10^8 particles/mL, detection limit: 2.5×10^6 particles/mL	(Carvalho et al., 2022)
Electrochemical method	Bovine rotavirus (BRV)	BRV-specific affinity peptides	BRV particles in laboratory samples	Detection limit: 5 copies/mL, LOQ: 214 copies/mL	(Cho et al., 2022)
SERS	Rotavirus	None (direct detection)	Rotavirus particles in clinical stool samples	Detection limit: 1%	(Biswas et al., 2023)

FFU: Focus-Forming Units, MOI: Multiplicity of Infection, PFU: Plaque-Forming Units, TCID: Tissue Culture Infectious Dose, LOQ: Limit of Quantification

A few non-antibody based recognition element for developing sensors and biosensors for rotavirus have been reported (Zhang & Zhao, 2015; Bertol & Gatti, 2016). However, these methods have either not been adequately validated for their specificity or suffers from other draw backs such as, complex and cumbersome fabrication process, lengthy processing times, high costs associated with specialized instrumentation, and the need for skilled personnel.

1.5.2.1. Molecular Beacon as Biosensor

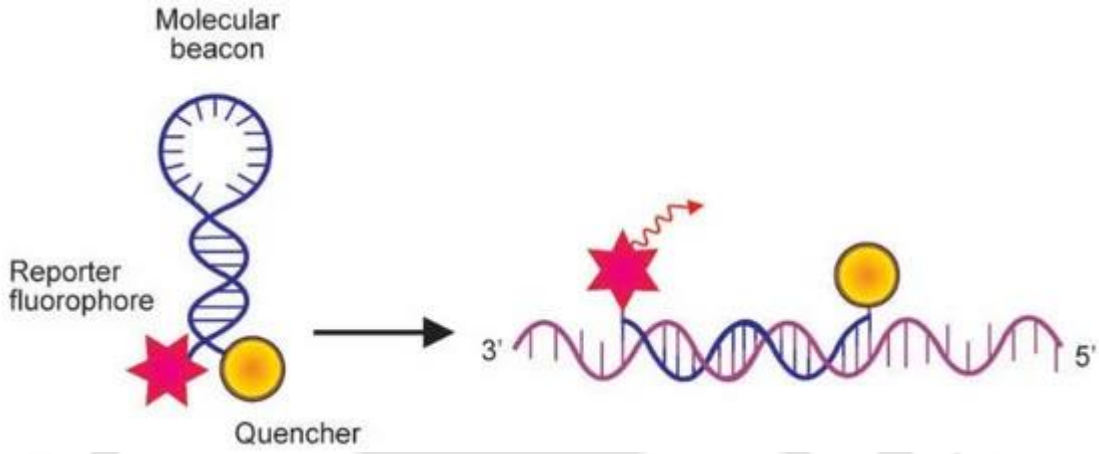


Figure 1.4. Mechanism of action of the molecular beacon. Molecular beacons contain covalently linked fluorescent and quenching dyes at either end of a single-stranded DNA molecule. Whilst free in solution, the probe is maintained in a hairpin conformation by complementary stem sequences at both ends of the probe, which brings the fluorescent dye and the quencher in close proximity. This causes FRET to occur, which suppresses reporter fluorescence. The loop part of the molecule is complementary to the target nucleic acid molecule. In the presence of a target sequence, the loop hybridizes to the complementary target sequence throughout the annealing step, resulting in a conformational alteration that forces the reporter and quencher dyes to separate, and fluorescence is emitted. Copyright Made Artika et al., 2022. Under creative common license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Molecular beacons are advanced nucleic acid-based biosensors designed to detect specific nucleic acid sequences with high precision. Their innovative design leverages fluorescence resonance energy transfer (FRET) to provide real-time detection of target nucleic acids, making them invaluable in diagnostics, research, and environmental monitoring. Each molecular beacon consists of a hairpin-shaped oligonucleotide with a

fluorophore at one end and a quencher at the other. In the absence of the target, the fluorophore and quencher are in close proximity due to the stem-loop structure, leading to quenching of the fluorescence signal. Upon binding of the molecular beacon to its complementary target sequence in the loop region, a conformational change occurs that separates the fluorophore from the quencher. This separation results in an increase in fluorescence, which is directly proportional to the amount of target present (Stobiecka & Chałupa, 2015; Bidar et al., 2021). The mechanism of molecular beacon sensing is delineated in **Figure 1.4** (Artika et al., 2022).

The advantages of molecular beacons are manifold. Their high specificity is achieved through the design of a unique loop sequence complementary to a specific target, which minimizes false positives and ensures accurate detection. The real-time monitoring capability allows for immediate feedback during assays, making molecular beacons particularly valuable in quantitative PCR (qPCR) and other fluorescence-based assays. The sensitivity of molecular beacons enables them to detect even low concentrations of target nucleic acids, which is crucial for applications requiring sensitive measurement. Additionally, molecular beacons offer reduced background signal due to their FRET mechanism, which enhances the signal-to-noise ratio and improves assay accuracy. Their versatility allows them to be adapted for a wide range of nucleic acid sequences, including DNA, RNA, and mutations. Molecular beacons are straightforward to use in various detection platforms, streamlining the detection process and making them accessible for diverse applications (Han et al., 2013).

1.6. Significant Gaps in the Research

Rotavirus research has achieved significant advancements in understanding the virus and developing preventive measures. However, there remain notable gaps in several key areas, particularly in vaccine development, the expression challenges associated with the rotavirus VP6 protein, and the need for more effective diagnostic tools.

1.6.1. Need for Region-Specific Vaccine Development

There is a substantial gap in the development of rotavirus vaccines that are tailored to specific geographical regions. Current vaccines, such as Rotarix and RotaTeq, show reduced efficacy in certain low-income settings. These vaccines were designed to target strains circulating in high-income regions and may not fully protect against region-specific genotypes. There is a lack of vaccines that are specifically formulated to

account for the diverse strain distribution in different parts of the world, particularly in areas with the highest disease burden. This gap underscores the need for a better understanding of regional rotavirus diversity and for the development of vaccines that provide broader protection across various geographical contexts.

1.6.2. Challenges in Prokaryotic Expression of VP6 for Vaccine and Diagnostic Applications

The VP6 protein, a highly conserved inner capsid protein of rotavirus, is considered both a promising vaccine candidate and a valuable detection target in research and diagnostic studies. Its immunogenic properties make it an attractive candidate for vaccine development, while its conservation across different rotavirus strains allows for its use in diagnostic assays. However, a significant gap exists due to the challenges associated with the expression of VP6 in prokaryotic systems. Although prokaryotic expression is convenient, cost-effective, and widely used in research settings, the VP6 protein tends to form insoluble aggregates when expressed in common prokaryotic hosts such as *Escherichia coli*. This solubility issue hinders effective VP6 production and purification, limiting its potential applications in vaccine development, diagnostic tests, and further research.

1.6.3. Insufficient Diagnostics for Rotavirus

Current diagnostic methods for rotavirus are limited by their requirement for laboratory infrastructure, trained personnel, and time-consuming procedures. This presents a significant gap in the availability of rapid, point-of-care diagnostic tools that are easy to use and can provide timely results, especially in resource-limited settings. Additionally, there is a gap in the development of highly sensitive and specific biosensors that could enable quick and accurate detection of rotavirus infections at the point of care, improving both outbreak management and patient care.

1.7. Objectives of the Study

I. Molecular Epidemiology of RVA in India - A Survey Through Sequence Data Mining and Phylogenetic Analysis
II. Immunoinformatics-Mediated Screening of RVA Epitopes for Potential Peptide Vaccine Development

III. Optimization of Recombinant RVA VP6 Protein Expression in *E. coli* System and Probing of Aggregation Hotspots in the VP6 Protein

IV. Development of a Molecular Beacon-Based Genosensor for Detection of Human RVA

1.8. Significance of the Study

The significance of this research is multifaceted, addressing critical needs in rotavirus vaccine design and development, protein expression optimization and detection.

The molecular epidemiology study of RVA in India, involving sequence data mining and phylogenetic analysis, provides essential insights into the genetic diversity and regional distribution of RVA strains. This comprehensive understanding is crucial for tracking the prevalence and evolution of different strains, which supports the development of region-specific vaccines and tailored public health management.

The findings from the phylogenetic study are particularly significant for the second objective, which focuses on immunoinformatics-mediated screening of RVA epitopes for potential peptide vaccine development. By incorporating phylogenetic data, the research can identify epitopes that are not only immunogenic but also relevant to the specific strains circulating in the Indian subcontinent. This targeted approach enhances the potential efficacy of peptide-based vaccines by ensuring they are designed to combat the most prevalent and genetically diverse strains.

In terms of protein expression, optimizing recombinant RVA VP6 protein production in *E. coli* and identifying aggregation hotspots are critical for improving the quality and stability of the protein. Aggregation can significantly impact protein solubility and functionality, which are essential for developing effective vaccines and diagnostic tools. Addressing these challenges through optimization of protein expression and hotspot identification will enhance the reliability and scalability of the VP6 protein production process.

Lastly, the development of a fluorometric genosensor based on a molecular beacon with a fluorescein-dabcyl pair meets a crucial need for rapid and accurate rotavirus detection. This innovative diagnostic tool offers high sensitivity and specificity, making it a valuable asset for clinical and environmental rotavirus surveillance. By providing

timely and precise detection, the genosensor will improve disease management and control measures.

Based on the above gaps and challenges, the current work is categorized into three experimental chapters in addition to the preceding "Introduction and Literature Review" chapter. At the end, "Conclusion and Future Outlook" section summarizes our findings, views and future prospects relating to the current investigation.

Chapter 1: Introduction and Literature Review

Chapter 1 provides a comprehensive overview of rotavirus, covering essential aspects of its biology and health impact. It begins with a detailed description of rotavirus structure and classification, including its classification into different serotypes and genotypes based on surface proteins. The chapter then delves into the mechanisms of rotavirus replication, outlining the virus's lifecycle from cell entry to genome replication and virion assembly. The chapter further addresses the disease burden of rotavirus, highlighting its significant impact on global health, particularly among young children and in regions with limited healthcare access. Current therapeutic approaches in managing infections are reviewed. Diagnostic methods are also covered, involving both traditional and advanced detection techniques, which are crucial for accurate detection of rotavirus in clinical and environmental samples. Additionally, the chapter identifies significant gaps in rotavirus research, such as the need for improved vaccines and detection strategies. It concludes by outlining the objectives and significance, emphasizing how this research aims to address these gaps and advance the fields of rotavirus therapeutics and diagnostics. Overall, Chapter 1 sets a foundation for understanding the complexities of rotavirus and the necessity for continued research and innovation.

Chapter 2: Reverse Vaccinology-Based Multi-Epitope Vaccine Design Against Indian Group A Rotavirus Targeting VP7, VP4, and VP6 Proteins

Chapter 2 delves into the use of immunoinformatics to develop a peptide-based vaccine against RVA. The chapter outlines the methodology for selecting potential peptide epitopes from RVA to enhance vaccine efficacy. It begins by detailing the process of identifying conserved epitopes from the RVA capsid proteins VP7, VP4, and VP6. The selection criteria include antigenicity, non-allergenicity, non-toxicity, and stability. These epitopes are linked using suitable linkers and combined with an N-terminal beta-

defensin adjuvant to form a multiepitope vaccine. This approach employs computational tools to predict and optimize epitope candidates, aiming to improve vaccine development and immunization strategies.

Chapter 3: Integrated Analysis of RVA VP6: *E. coli* Expression Optimization, Characterization of Inclusion Bodies, and In Silico Prediction of Aggregation-Prone Regions

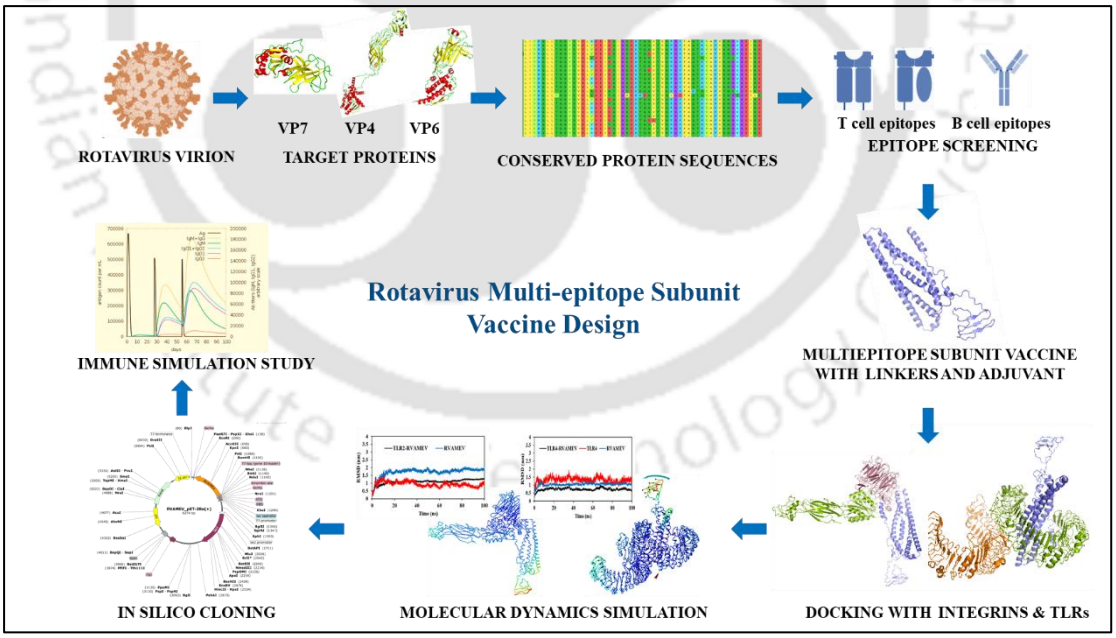
Chapter 3 investigates the expression and characterization of RVA VP6 protein in *E. coli*, addressing the challenges of inclusion body formation and identifying aggregation-prone regions (APRs). The chapter begins with the optimization of VP6 expression in *E. coli* BL21(DE3) cells, and characterization of VP6 inclusion bodies using transmission electron microscopy and employ Thioflavin T assay to study their aggregation behaviour at different induction temperatures. Computational tools used to predict aggregation-prone regions, include both sequence-based tools and structure-based tools. This integrated approach, combining both experimental and computational methods, provides an understanding of VP6 aggregation.

Chapter 4: Development of a Molecular Beacon-Based Genosensor for Detection of RVA

Chapter 4 focuses on developing a molecular beacon-based genosensor for RVA detection, aiming to provide a portable, specific, and cost-effective diagnostic tool. The chapter details the design of a molecular beacon incorporating fluorescein and dabcyl labels, targeting a highly conserved sequence from the human RVA VP6 gene. It describes the characterization of the beacon's secondary and tertiary structures, ensuring the correct stem-loop configuration and absence of internal secondary structures. The chapter also covers the optimization of assay conditions, such as excitation wavelength and incubation time, and evaluates specificity through interference studies with various mismatch and scrambled targets.

Chapter 2

REVERSE VACCINOLOGY-BASED MULTI-EPITOPE VACCINE DESIGN AGAINST INDIAN GROUP A ROTAVIRUS TARGETING VP7, VP4, AND VP6 PROTEINS



2.1 Overview

Over the years, multiple live-attenuated rotavirus vaccines have been developed and approved for inclusion in immunization programs worldwide. Rotashield, the inaugural rotavirus vaccine sanctioned in the United States in 1998, faced association with intussusception, leading to its prompt withdrawal in 1999 (Centre for Disease Control and Prevention (CDC), 1999). Currently, WHO-prequalified rotavirus vaccines encompass Rotarix (GlaxoSmithKline Biologicals; prequalified in 2009), RotaTeq (Merck & Co., Inc.; prequalified in 2008), Rotavac (Bharat Biotech, Hyderabad, India; prequalified in 2018), and Rotasiil (Serum Institute of India Pvt. Ltd., Pune, India; prequalified in 2018) (Burke et al., 2019). Presently, one or more WHO recommended rotavirus vaccines are involved in immunization programs, initiated in 123 countries internationally (https://view-hub.org/vaccine/rota). While the introduction of various licensed rotavirus vaccines has demonstrated a significant decrease in hospitalizations due to rotavirus-induced gastroenteritis in several regions (Aliabadi et al., 2019), the efficacy of these live attenuated vaccines remains inadequate in low-income nations. For example, in the majority of low- to middle-income countries with high rates of rotavirus gastroenteritis, the reported average vaccine effectiveness for Rotarix and RotaTeq is 57% and 45%, respectively. In contrast, in high-income nations experiencing low incidences of rotavirus-induced gastroenteritis, the effectiveness increases to 84% for Rotarix and 90% for RotaTeq (Jonesteller et al., 2017). In India, the Rotavac vaccine exhibited a clinical trial efficacy rate of 54%, whereas RotaSiil demonstrated an efficacy of 67% (Roberto et al., 2021)(Carvalho & Gill, 2019; Kulkarni et al., 2017; O'Ryan, 2017). Moreover, there is a chance of live attenuated vaccines reverting to a more virulent strain,

The genetic and antigenic diversity of rotaviruses, stemming from their segmented genome and interspecies transmission, plays a pivotal role in the variability of vaccine efficacy. This diversity facilitates genetic re-assortment during co-infection, leading to numerous genotype combinations and the emergence of novel genotypes (Martella et

al., 2010). Rotavirus genotype distribution varies significantly across different regions worldwide, resulting in distinct circulating strains that impact vaccine performance. Consequently, currently available live-attenuated rotavirus vaccines often demonstrate suboptimal effectiveness in various geographical areas (Iturriza-Gómara et al., 2011). The variability in vaccine efficacy due to rotavirus genetic diversity, the risk of intussusception, the potential for live-attenuated vaccines to revert to more virulent strains, and contraindications for immunocompromised individuals highlight the need for research and development of innovative rotavirus vaccines.

Subunit vaccines exhibit a propensity for minimal or negligible adverse effects, heightened immunogenicity when coupled with adjuvants, and a favorable safety profile for immunocompromised individuals. Comprising antigenic fragments, such as purified proteins, recombinant proteins, polysaccharides, or peptides, subunit vaccines elicit targeted immune responses. Notably, their production processes are straightforward, rendering them relatively safer than their whole virus counterparts (Vartak & Sucheck, 2016). Addressing the challenge posed by extensive rotavirus diversity necessitates the development of a vaccine targeting highly conserved sequences for effective strategies. Leveraging the existing genomic and proteomic data on rotavirus, a multiepitope-based subunit vaccine can be designed, incorporating conserved protein sequences from region-specific prevalent rotavirus strains, providing a strategic approach to enhance vaccine effectiveness. This strategy involves the application of reverse vaccinology and immunoinformatics methodologies. In recent years, the reverse vaccinology approach has facilitated the delineation of antigenic epitopes of rotavirus, aiding in the design and development of a rotavirus subunit vaccine (Kuri & Goswami, 2023). Nevertheless, to date, no vaccine has been developed utilizing the in-silico approach with a specific focus on targeting RVA strains in Indian subcontinent. This study represents the pioneering effort in designing a multiepitope vaccine for RVA specifically tailored to the Indian subcontinent, utilizing Indian RVA strains. The investigation undertakes the task of designing a multiepitope vaccine for

RVA, utilizing conserved RVA capsid proteins, specifically VP7, VP4, and VP6, derived from strains prevalent in the Indian subcontinent. Rotavirus VP7 and VP4 proteins stand out as pivotal factors in orchestrating protective immunity, making them compelling targets for vaccine design (Hoshino & Kapikian, 2000). Additionally, VP6, being highly abundant and the most conserved capsid protein, constitutes the middle layer of RV particles and has demonstrated both antigenicity and immunogenicity (Afchangi et al., 2019). Employing an in-silico approach, we identified T-cell (helper T lymphocytes, HTL and cytotoxic T lymphocytes, CTL) epitopes and linear B cell epitopes within conserved Indian RVA capsid protein sequences. These epitopes were then subjected to epitope assembly to formulate a vaccine construct, which underwent comprehensive assessment utilizing physicochemical parameters, immunogenicity evaluation, and three-dimensional structures. Additionally, in-silico immune simulation studies were conducted to ascertain the immune response profile of the designed vaccine. However, further validation through in vitro and in vivo examination is imperative to corroborate the computational findings.

2.2 Methodology Section

2.2.1 Sequence retrieval and phylogenetic analysis

The protein sequences of VP7, VP4, and VP6 derived from circulating RVA strains within the geographical region of India were procured from the Virus Variation Resource hosted by the National Center for Biotechnology Information (NCBI) (Hatcher et al., 2017). The search set was defined specifying “species” as “Rotavirus A”, “Host” as “Human”, “Region/Country” as “India” and “segment” as VP7, VP4 or VP6, one at a time. A total of non-redundant 1083 sequences of VP7, 923 sequences of VP4 and 472 sequences of VP6 were shortlisted. Additional screening of sequences was conducted, focusing on the inclusion of exclusively full-length sequences originating from fecal samples, which served as representations of clinical isolates. Further refinement was performed to include only full-length sequences isolated from fecal

samples, representing clinical isolates. Subsequently, carefully curated database was established, comprising 128 VP7 sequences, 116 VP4 sequences, and 127 VP6 sequences from clinically isolated human group A rotaviruses originating in the Indian subcontinent. The sequences of VP7, VP4, and VP6 were obtained and compiled into individual files for multiple sequence alignment using MEGA 7.0.18 software (Kumar et al., 2016). ClustalW algorithm was used for multiple sequence alignment. Subsequently, a phylogenetic tree was generated by implementing the neighbour-joining method, with a bootstrap value set to 1000 to assess the robustness of the tree topology. Further, corresponding protein sequences of all the viral capsid proteins were also subjected to ClustalW alignment in Mega 7.0.18 software, and their respective protein sequence conservancy were visualized using BioEdit v. 7.0.9 (Hall, 1999) and Weblogo3 (http://weblogo.threeplusone.com/) (Crooks et al., 2004). Representative sequences from distinct clusters within the phylogenetic tree were subsequently chosen for additional downstream analysis.

2.2.2 Prediction of Antigenicity

The representative protein sequences of VP7, VP4, and VP6 chosen from phylogenetic analyses were subjected to antigenicity prediction using the VaxiJen v2.0 online server (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) (Doytchinova & Flower, 2007). The prediction was carried out with a set threshold of 0.4. This computational tool was employed to evaluate the antigenicity levels of each individual protein sequence. Subsequently, sequences demonstrating the highest antigenicity were selected for further investigation in the context of epitope prediction.

2.2.3 Linear B-Cell Epitope Prediction

Antigenic/immunogenic epitopes refer to specific segments of antigens that are recognized by antibodies produced by immune B-cells. Bcepred, a web-based bioinformatics tool, has been tailored to delineate linear B-cell epitopes within protein sequences. This computational platform affords users the opportunity to forecast B-cell

epitopes by analyzing diverse physico-chemical attributes, including hydrophilicity, flexibility/mobility, accessibility, polarity, exposed surface, and turns, either in isolation or as integrated parameters. Our analysis reveals that the server achieves an impressive predictive accuracy of 58.7% when leveraging a composite of flexibility, hydrophilicity, polarity, and surface properties, employing a threshold of 2.38. This level of predictive precision surpasses that of other several linear B-cell epitope prediction servers, underscoring its scientific significance in the field of epitope prediction (http://crdd.osdd.net/raghava/bcepred) (Saha & Raghava, 2004). In our investigation, we harnessed the predictive capabilities of the Bcepred server for the prospective identification of linear B-cell epitopes using combination of all the four parameters.

2.2.4 T-cell Epitope Prediction

In the context of vaccine epitope identification, the mapping of T-cell epitopes assumes paramount importance. Within the spectrum of T-cell epitopes, CTL and HTL epitopes exhibit a pivotal role in triggering immune responses. Consequently, we embarked upon the task of identifying T-cell epitopes characterized by robust binding affinities for major histocompatibility complexes (MHC): MHCI (bind to CTL epitopes) and MHCII (bind to HTL epitopes). Employing the NetCTL 1.2 webserver (https://services.healthtech.dtu.dk/services/NetCTL-1.2/), we executed predictions concerning CTL epitopes that manifest binding affinities with MHC1. The NetCTL 1.2 server is employed for the anticipation of cytotoxic T lymphocyte (CTL) epitopes within protein sequences. This methodology estimates peptide binding to MHC class I molecules, the proteasomal C-terminal cleavage process, and the assessment of Transporter associated with antigen processing (TAP) transport efficiency, and is based on HLA class-I super types, i.e. A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58, B62. This server facilitates the precise projection of CTL epitopes, constrained to a defined set of 12 MHC class I super types. The prediction of MHC class I binding and

proteasomal cleavage is conducted through the implementation of artificial neural networks, while the evaluation of TAP transport efficiency is achieved using a weight matrix analysis (Larsen et al., 2007). The default set values of 0.15, 0.05 and 0.75 were used for weights on C terminal cleavage, TAP transport efficiency and epitope identification were set for epitope identification.

The antigenic capsid protein sequences were submitted to IEDB T-cell epitope MHCII tool (http://tools.iedb.org/mhcii/) for HTL epitope prediction. This computational tool employs a range of methods that integrates NN-align, SMM-align, and Combinatorial library techniques, to forecast MHC Class II epitopes. The predefined threshold for epitope selection was set at 4 with peptide length 15 and the full HLA human reference set used for the prediction of peptide binding. Within the framework of IEDB's methodology, a reference set of alleles was employed as a reference for peptide binding predictions. The selection of peptides with the lowest consensus scores was based on the principle endorsed by the Immune Epitope Database and Analysis Resource (IEDB) server, indicating that a lower percentile rank correlates with higher binding affinity (Bui et al., 2005; Nielsen et al., 2007; Sturniolo et al., 1999).

2.2.5 In silico Characterization of Epitopes

The main criteria for epitopes to be considered as viable vaccine candidates include their capacity to induce an immune response by being sufficiently antigenic, as well as being non-allergenic, non-toxic, and non-homologous to human proteins. The antigenicity of the epitopes ultimately selected was determined using the VaxiJen tool (http://www.ddgpharmfac.net/vaxijen/VaxiJen/VaxiJen.html) (Doytchinova & Flower, 2007), consensus allergenicity prediction was determined by AllerTop v. 2.0 (https://www.ddg-pharmfac.net/AllerTOP/) (Dimitrov et al., 2014) and AllergenFP v.1.0 server (http://ddg-pharmfac.net/Allergen FP) (Dimitrov et al., 2014), the toxicity through Toxinpred (http://crdd.osdd.net/raghava/toxinpred/) (Gupta et al., 2013). Furthermore, physicochemical characterization of the individual epitopes was

performed using the ProtParam tool (https://web.expasy.org/protparam/) (Wilkins et al., 1999), facilitating the computation of crucial molecular parameters such as molecular weight, theoretical isoelectric point (p*I*), aliphatic index, and the grand average of hydropathicity (GRAVY). The conservancy analysis of the selected epitopes was performed using the epitope conservancy analysis tool of IEDB server (https://www.iedb.org/conservancy/) (Vita et al., 2019). This tool quantifies the extent to which an epitope is conserved among a collection of protein sequences, at a designated identity threshold. In this instance, the linear sequence conservancy of the VP7 epitopes was assessed at a 60% sequence identity threshold level. The epitopes meeting all specified criteria were subsequently employed in the formulation of the vaccine construct.

2.2.6 Construction of Multiepitope-Based Vaccine: Vaccine Assembly

For the successful construction of the vaccine, we systematically integrated the predicted CTL, HTL, overlapping CTL and HTL epitopes and linear B-cell epitopes. The N-terminal beta-defensin was utilized as an adjuvant, and the EAAAK linker served to conjugate the adjuvant with HTL/CTL epitopes. The interconnection of HTL/CTL epitopes was established using the AAY and the KK linker. The GGGGS linker was employed to connect the linear B-cell epitopes.

2.2.7 In silico characterization of the Rotavirus A Multi-Epitope Vaccine (RVAMEV)

The assessment of vaccine antigenicity was conducted employing the VaxiJen v2.0 online server (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.htm) with a set prediction threshold of 0.5. The determination of consensus allergenicity predictions involved utilization of the AllerTop v. .0 server (https://www.ddg-pharmfac.net/AllerTOP/) (Dimitrov et al., 2014) and the AllergenFP v.1.0 server (http://ddg-pharmfac.net/AllergenFP) (Dimitrov et al., 2014). Furthermore, toxicity predictions were executed through the ToxIBTL server (http://server.wei-group.net/ToxIBTL) (Wei et al., 2022). The evaluation of diverse physicochemical

properties of the vaccines was systematically performed using the ProtParam online server (https://web.expasy.org/protparam/) (Wilkins et al., 1999). Homology evaluation between RVAMEV and the human proteome was systematically conducted employing the NCBI BLAST (Basic Local Alignment Search Tool), specifically utilizing blast-p (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) (McGinnis & Madden, 2004). The principal objective of this analysis was to rigorously examine the designed vaccine's homologous relationships and correlation with the human proteome. The vaccine sequence, formatted in FASTA, functioned as the primary query for the BLAST exploration. Our inquiry was systematically directed towards the non-redundant protein database, and to ensure the utmost specificity, we judiciously tailored the search parameters, exclusively delving into protein sequences from *Homo sapiens* (taxid 9606), *Lactobacillus johnsonii* (taxid 33959), *Lactobacillus rhamnosus* (taxid 47715), and the *Lactobacillus casei* group (taxid 655183). This rigorous selection process was pivotal in confining our investigation to sequences originating from organisms of prime relevance, facilitating a meticulous evaluation of potential cross-reactivity and other pertinent facets germane to vaccine development.

2.2.8 Secondary and Tertiary Structure Prediction of the Vaccine Construct

The secondary structure of the designed vaccine construct was generated using the online tool PRISPRED 4.0 (http://bioinf.cs.ucl.ac.uk/psipred/) and RaptorX Property servers (https://predictprotein.org). PRISPRED utilizes a combination of machine learning and neural network techniques to analyze amino acid sequences, offering accurate predictions for alpha-helices, beta-strands, and coils in protein structures (McGuffin et al., 2000). The tertiary or 3D structures of the vaccine was generated using I-TASSER (Roy et al., 2010). I-TASSER predicts protein structures and annotates functions by submitting the primary amino acid sequence. It uses template-based modelling, refines models through molecular dynamics simulations, and incorporates

experimental data. Known for accuracy, especially for targets without structural homologs, it provides concise models with functional annotations.

2.2.9 3D Structure Refinement and Validation

The predicted 3D structure of the constructed vaccine was subjected to structural refinement by galaxy refine webserver (http://galaxy.seoklab.org/refine) (Heo et al., 2013). The platform functions as an expedient, user-friendly, and effective tool for refining protein structures, intricately designed to optimize and enhance three-dimensional protein models. It meticulously refines structures, with a focus on local regions, resulting in models of heightened accuracy and reliability. The refined models underwent thorough evaluation by diverse validation tools to discern the optimal model and assess its overall quality. The tools used for structural validation include PROCHECK (https://saves.mbi.ucla.edu/), ERRAT server (https://saves.mbi.ucla.edu/) (Colovos & Yeates, 1993) and PROSA web tool (https://prosa.services.came.sbg.ac.at/prosa.php) (Wiederstein & Sippl, 2007). ProCheck systematically evaluates stereochemical quality by assessing parameters such as bond lengths, bond angles, and dihedral angles. Significantly, it generates the Ramachandran plot, offering valuable insights into the conformational space of amino acid residues and assisting in identifying structural irregularities for subsequent refinement. In addition, the ERRAT server (https://saves.mbi.ucla.edu/) (Colovos & Yeates, 1993) operates as a verification algorithm for protein structure designed to assess the advancements of crystallographic model building and refining. Maintained by the National Health Institute at the University of California, USA, this server plays a crucial role in ensuring structural integrity. The determination of the Z-score was accomplished through the PROSA web tool from the Centre of Applied Molecular Engineering, Division of Bioinformatics, University of Salzburg, Salzburg, Austria. This tool quantifies the deviation of the total energy of the structure regarding an energy pattern derived from random conformations found in native proteins (Wiederstein & Sippl, 2007).

2.2.10 Molecular Docking Analyses of Multiepitope-Based Vaccine with Integrins and Toll-Like Receptors (TLRs)

Rotavirus entry into the host cell entails a complex series of events, including the proteolytic cleavage of the VP4 spike protein into VP5 and VP8 fragments, with interaction involving integrins (αvβ3) (Guerrero et al., 2000; Zárate et al., 2004). To explore the interactions between the designed RVAMEV and integrins, molecular docking was performed using the vaccine structure with integrin receptors, integrin αVβ3 (PDB ID: 4O02) and integrin αIIbβ3 (PDB ID: 2vdp). TLRs are key mediators in gut inflammatory pathways, linking adaptive and innate immunity by orchestrating responses to various pathogen-derived ligands. TLR2 and TLR4, members of the TLR family, are involved in recognizing viral structural and non-structural proteins, triggering the production of inflammatory cytokines (Zhou et al., 2021). Thus, the RVAMEV was docked with TLR2 (PDB ID: 5D31) and TLR 4 (PDB ID: 4G8A). Docking was conducted using ClusPro v.2.0 docking program (www.cluspro.bu.edu) with default settings (Comeau et al., 2004). The server categorizes the clusters of docked complexes according to their centre and lowest energy scores. The PRODIGY tool of the HADDOCK (https://haddock.science.uu.nl/) web server were used to determine the binding affinity (ΔG in kcal/mol) of the docked complexes. The server ranks the clusters of docked complexes based on their centre and lowest energy scores. The docked structures were visualized by PyMol tool (Schr€odinger, 2015). In addition, PDBsum (Laskowski et al., 2018) was used to map the interacting residues between the docked chains.

2.2.11 Molecular Dynamics Simulation (MDS) Studies

The stability of vaccine-receptor (RVAMEV- Integrin αVβ3, RVAMEV- Integrin αIIbβ3, RVAMEV-TLR2 and RVAMEV-TLR4) complexes was investigated utilizing the iMODS server (https://imods.iqf.csic.es/) (Lopéz-Blanco et al., 2011)(López-Blanco et al., 2014). Employing normal mode analysis (NMA), the server calculated protein internal coordinates, crucial for evaluating stability. The best docked complexes

of the RVAMEV with integrins and TLRs, one at a time, served as the input for this analysis. The results were illustrated through deformability plots, B-factor values, eigenvalues, covariance matrices, and elastic network models. In addition to NMA, MDS was conducted using GROMACS 2019 (Abraham et al., 2015) for TLRs and RVAMEV complexes. Docked TLR and RVAMEV complexes were placed in a cubic box and solvated with TIP3P water molecules. The Amber ff99SBILDN force field was used for protein parameters. System neutrality was achieved with K^+ ions, and an ionic strength of 0.15 M was maintained using the Joung-Cheatham ion model (Joung & Cheatham, 2008). Initial energy minimization included 50,000 steps of the steepest descent method with 1000 kJ/mol nm² restraints on protein heavy atoms, followed by unrestrained minimization. Equilibration was done in phases, and production simulations ran for 100 ns using NVT and NPT ensembles. Temperature was set at 300 K with velocity rescaling (0.1 ps coupling time), and pressure was maintained at 1 atm using a Parrinello-Rahman barostat (2 ps coupling time). The leapfrog algorithm integrated equations of motion with a 2.0 fs time step. Electrostatic interactions were calculated using Particle Mesh Ewald (PME) with 1.0 nm cut-offs for Coulomb and van der Waals interactions. Periodic boundary conditions were applied, and hydrogen bond lengths were constrained with the LINCS algorithm (Hess et al., 1997). Analysis was performed with GROMACS tools, and graphs were generated using Excel 2016.

2.2.12 Codon Adaptation and In Silico Cloning of Vaccine

A tactic for enhancing the expression of the recombinant RVAMEV protein was employed through the implementation of codon optimization. This optimization proves essential given the degeneracy inherent in the genetic code, which allows for multiple codons to represent most amino acids. The reverse transcription of the vaccine protein sequence was performed using EMBOSS Backtranseq (www.ebi.ac.uk/) to obtain the DNA coding sequence, while codon adaptation for the *E. coli* K-12 codon system, was performed using the Java Codon Adaptation Index (JCAT) (http://www.jcat.de/) to ascertain the Codon Adaptation Index (CAI) value and GC content. The finalized

sequence of the vaccine construct, following codon optimization, was seamlessly integrated into an expression vector using the Snapgene tool (https://www.snapgene.com/free-trial/). The selected plasmid vector, pET-28a (+), was specifically chosen for this purpose, being tailored for the expression of N-terminally 6 x His-tagged proteins. This N-terminal tagging strategy is strategically advantageous, particularly in terms of purification and protein recovery, aligning with the established benefits associated with N-terminal tags in these processes.

2.2.13 Immune Simulation and Prediction of the Immune Profile of the Vaccine Construct

The computational assessment of the in silico immune response profile for the RVAMEV vaccine was conducted through the utilization of the C-IMMSIM web server (Rapin et al., 2010). In the anticipation of the immunological response, the C-IMMSIM web server employs position-specific scoring matrices derived from a machine learning methodology. The recommended temporal interval between the primary and secondary vaccine administrations is subject to variation, encompassing durations of 8 weeks, 3 months, or 6 months (Rapin et al., 2010). In this investigation, the forecasted immune response profile involves the administration of three vaccine doses, each spaced at 4-week intervals. The simulation parameters for the vaccine comprise 300 steps (where one step equates to 8 hours of real-life time), and the specified intervals for the three vaccine doses are set at 1 day, 4 weeks, and 8 weeks, while maintaining constancy in the remaining parameters.

2.3 Results and Discussion

2.3.1 Sequence Retrieval and Evolutionary Relationship Analysis

The Human RVA capsid protein sequences available at GenBank, NCBI were filtered out based on full length sequence and isolation source as feces/stool to represent clinical cases. A total of 128 VP7, 116 VP4 and 127 VP6 full length gene sequences were

downloaded from the source (**Table 2.1**). Details on accession number, sequence length, genotype, host, country, isolation source and collection date are recorded in supplementary **Table A1**, **Table A2** and **Table A3** of **Appendix 1.**

The phylogenetic trees of VP7, VP4 and VP6 are represented in **Figure 2.1**. Phylogenetic analysis divided Indian RVA VP7 sequences clustered into six genotypes, G1, G2, G3, G10, G9 and G12 (**Figure 2.1.A**). These sequences were isolated from the year 2006 to 2017. Majority of these sequences belonged to the genotype G1 (53.90%), followed by G9 (18.75%), G2 (14.06%), G12 (10.15%), G3 (2.34%) and G10 (0.78%). Indian Rotavirus VP4 sequences clustered into four genotypes, P4, P6, P8, P11 (**Figure 2.1.B**).

Table 2.1. Sequence filtration criteria for studying molecular epidemiology of RVA in the Indian subcontinent.

Criteria (Nucleotide sequence)	**Total no. of sequence hits for VP7**	**Total no. of sequence hits for VP4**	**Total no. of sequence hits for VP6**
Sequence type: Full and partial length **Isolation source**: any	1083	923	472
Sequence type: Full length sequences **Isolation source**: any	399	161	218
Sequence type: Full length **Isolation source**: feces	48	14	9
Sequence type: Full length **Isolation source**: stool	80	102	118
Sequence type: Full length **Isolation source**: feces and stool	128	116	127

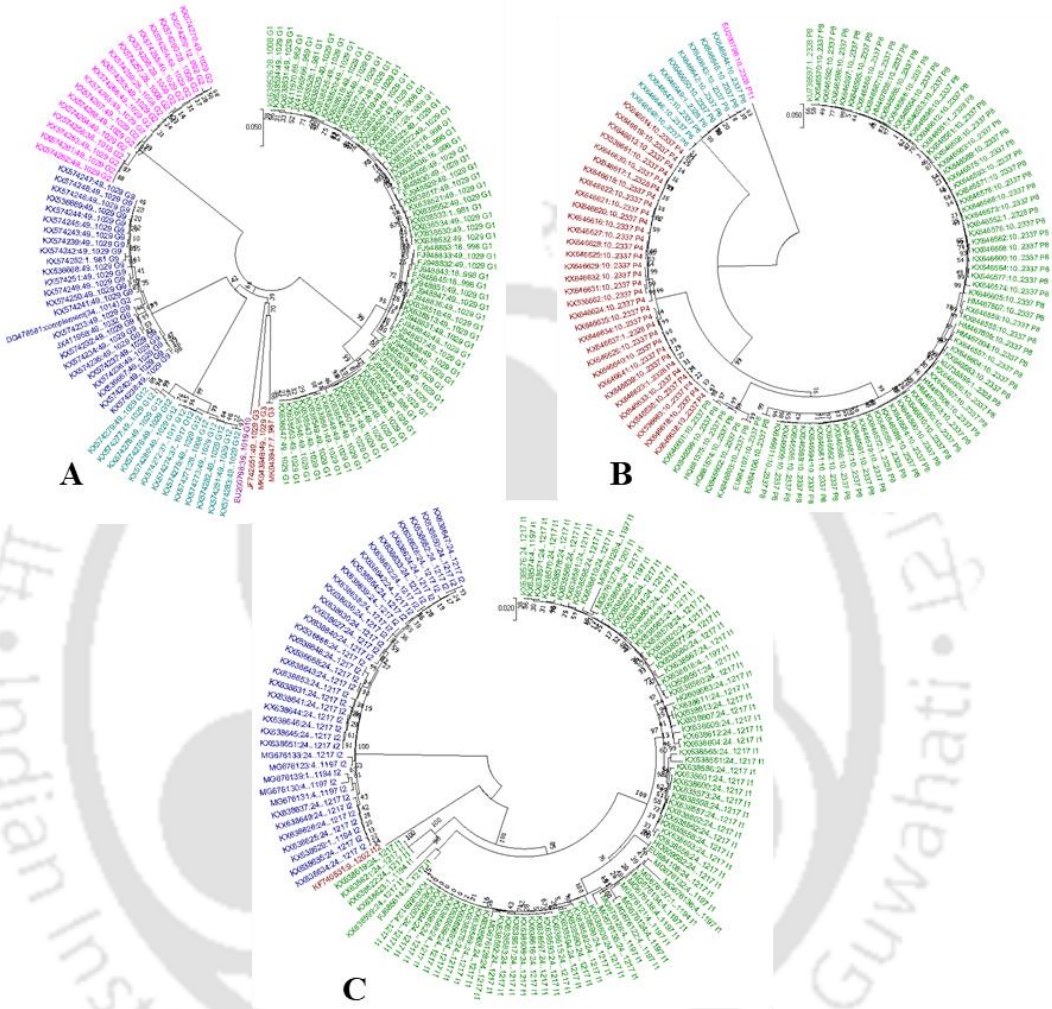


Figure 2.1. Phylogenetic tree of A. VP7 (green: G1, brown: G3, purple: G10, cyan: G12, indigo: G9, pink: G2), B. VP4 (green: P8, brown: P4, cyan: P6, pink: P11) and C. VP6 (green: I1, brown: I12, indigo: I2). The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length equal to 1.48711758 for VP7, 1.21897196 for VP4 and, 0.85416131 for VP6 are shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches.

These sequences were isolated from the year 2006 to 2013. Majority of these sequences belonged to the genotype P8 (63.79%) followed by P4 (27.58%), P6 (7.75%) and P11 (0.86%). This analysis indicates that G1P[8] strain of human RVA is the most prevalent strain in India, in accordance with the global scenario. Phylogenetic analysis of Indian RVA VP6 full length gene sequences exhibited three genotypes namely genotype I1 (69.29%), I2 (29.92%) and I12 (0.78%) (**Figure 2.1.C**). They represented a collection of VP6 sequences from 1992 to 2016. Protein sequence alignment and conservancy of selected sets of VP7, VP4 and VP6 sequences have been elucidated in **Figure A1**, **Figure A2** and **Figure A3** of **Appendix 1**, respectively. Alignment of protein sequences shows a high level of conservation amongst themselves, each for protein sequence sets of VP7, VP4 and VP6 with slight sequence variability in certain positions. Sequence conservancy served as one of the criteria for screening and selection of epitopes.

Table 2.2. VaxiJen scores of representative sequences of VP4, VP7 and VP6 of RVA.

Protein	Group and Genotype	Nucleotide accession number	Protein accession number	Vaxijen Score (threshold value 0.4)
VP4	RVA P8	KX646603.1	ASU62303.1	0.4849
	RVA P4	KX646621.1	ASU62321.1	0.4770
	RVA P6	KX646648.1	ASU62348.1	0.4579
	RVA P11	EU200796.1	ABY64684.1	0.4509
VP7	RVA G1	KX638530.1	AST25359.1	0.5220
	RVA G9	KX574246.1	ASR91770.1	0.4813
	RVA G2	KX574258.1	ASR91782.1	0.5102
	RVA G3	MK043947.1	QAR17781.1	0.4171
	RVA G10	EU200798.1	ABY64686.1	0.4864
VP6	RVA I1	MG676136.1	AYM47564	0.5091
	RVA I2	KX638625.1	AST25454.1	0.5028
	RVA I12	KF740531.1	AIC82074.1	0.4866

2.3.2 Prediction of Antigenicity

Representative conserved RVA VP7, VP4 and VP6 protein sequences from each cluster were submitted to Vaxijen 2.0 server keeping the threshold score as 0.4 for the prediction of antigenicity. Higher the Vaxijen score, higher is the probable antigenicity. The Vaxijen scores for representative RVA VP4, VP7, and VP6 are presented in **Table 2.2**. The sequences with the highest Vaxijen antigenic scores were chosen for further analysis.

2.3.3 Linear B-Cell Epitope Prediction

The Bcepred prediction method relies on a dataset of 1029 experimentally validated continuous B-cell epitopes and an equal number of non-epitopes, encompassing proteins from viruses, bacteria, protozoa, and fungi. As a result, the Bcepred server emerges as a dependable approach for forecasting linear B-cell epitopes, demonstrating an accuracy of 58.70%, higher compared to its counterparts for linear B-cell epitope prediction. Amino acid properties, specifically hydrophilicity, flexibility, polarity, and exposed surface form the basis of such Bcepred predictions (Saha and Raghava, 2004). The selected representative protein sequences of VP4, VP7 and VP6 with highest Vaxijen scores were submitted to Bcepred server and a total of 118 epitopes for VP4 (7 to 30-mer length), 39 epitopes for VP7 (7 to 23-mer length) and 59 epitopes for VP6 (7 to 29-mer length) linear B-cell epitopes were predicted using this approach.

2.3.4 T-cell Epitope Prediction

MHCI (CTL) and MHCII (HTL) epitopes were predicted employing NetCTL 1.2 web server (https://services.healthtech.dtu.dk/services/NetCTL-1.2/) (Larsen et al., 2007) and IEDB T-cell epitope MHCII tool (http://tools.iedb.org/mhcii/) (Vita et al., 2019) for HTL epitope prediction webserver using the default settings, respectively. The NetCTL 1.2 web server integrates forecasts for the three main phases of the MHCI antigenic presentation pathway, which include proteasomal processing, transporter-associated antigen processing (TAP), and MHCI binding. The efficiency of TAP transport is established using a weight matrix, while the prediction of MHC-peptide

binding and proteasomal C-terminal cleavage is performed using an artificial neural network. The representative sequences of all the three capsid proteins were submitted to the IEDB epitope prediction server for T-cell MHCII binding predictions choosing a reference set of alleles. According to IEDB server, a lower percentile rank for an output peptide indicates higher affinity with MHCII. A multitude of CTL and HTL peptide epitopes were predicted for all the three capsid proteins. Thus, an array of epitopes with high NetCTL score and low percentile rank were chosen for further analysis.

2.3.5 In Silico Characterization of Epitopes

The fundamental criteria for epitopes to be deemed suitable as a prospective vaccine construct involve their capacity to elicit a sufficiently robust antigenic response while maintaining attributes of non-allergenicity, non-toxicity with respect to the host. The selected B-cell, CTL and HTL epitopes were subjected to investigations concerning antigenicity with VaxiJen (http://www.ddgpharmfac.net/vaxijen/VaxiJen/VaxiJen.html) (Doytchinova & Flower, 2007). Allergenicity prediction was carried out using AllerTop v.2.0 (https://www.ddg-pharmfac.net/AllerTOP/) (Dimitrov et al., 2014) and AllergenFP v.1.0 server (http://ddg-pharmfac.net/Allergen FP) (Dimitrov et al., 2014) and the epitopes predicted as non-allergen by both of the serves were screened for further evaluation. The peptide toxicity prediction was carried out by Toxinpred (http://crdd.osdd.net/raghava/toxinpred/) (Wei et al., 2022). Epitopes with high Vaxijen score, non-allergenic and non-toxic in nature were shortlisted and subjected to individual physicochemical characterization using the ProtParam tool (https://web.expasy.org/protparam/) (Wilkins et al., 1999). The attributes such as molecular weight, theoretical isoelectric point (p*I*), aliphatic index, and the grand average of hydropathicity (GRAVY) were recorded. Additionally, the conservancy analysis of the selected epitopes was performed using the epitope conservancy analysis tool of IEDB server (https://www.iedb.org/conservancy/) (Vita et al., 2019). Subsequently, a comparative analysis of CTL and HTL epitopes was conducted to

detect any overlapping, identical, or non-overlapping sequences that satisfied all criteria. Ultimately, a total of 20 epitopes (B-cell, CTL and HTL epitopes) meeting criteria for high antigenicity, non-allergenicity, non-toxicity, stability, and high conservancy were identified. Sequences meeting these criteria were further subjected to in silico epitope assembly for vaccine design. The final selected epitopes used for in silico vaccine construction are detailed in **Table 2.3.**

Table 2.3. Selected epitopes and their attributes for multiepitope vaccine design.

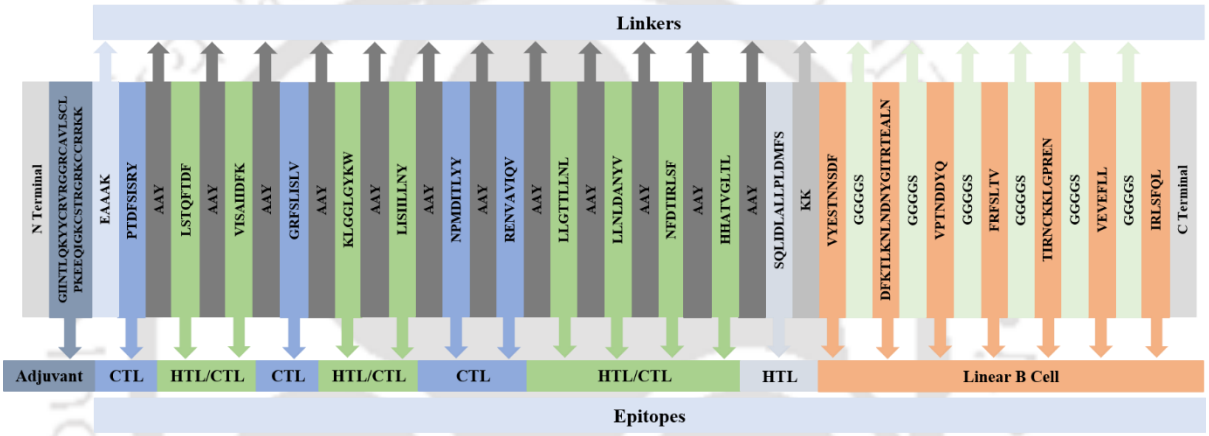
RVA Source Protein	Nature of Epitopes	Epitopes	Amino Acid Position in source proteins	Length	Mol.Wt (Da)	Isoelectric point	Instability Index (<40 is stable)	Stability	Aliphatic index	GRAVY Score
VP4 P8	Overlapping HTL and CTL epitopes	LSTQFTDF	412-419	8	958.04	3.8	-19.53	stable	48.75	0.025
		VISAIIDFK	719-727	9	1005.22	5.81	-16.19	stable	173.33	1.567
		KLGGLGYKW	284-292	9	1021.23	9.7	3.73	stable	86.67	-0.4
	Non-overlapping CTL epitopes	PTDFSISRY	335-343	9	1085.18	6.26	13.53	stable	43.33	-0.656
		GRFSLISLV	467-475	9	991.2	9.75	8.89	stable	162.22	1.4
	Non-overlapping HTL epitope	SQLIDLALLPLDMFS	516-530	15	1676	3.56	35.01	stable	162.67	1.087
VP7 G1	Overlapping HTL and CTL epitope	LISIILLNY	13-21	9	1061.33	5.52	30.29	stable	260	2.144
	Non-overlapping CTL epitope	NPMDITLYY	137-145	9	1129.29	3.8	11.02	stable	86.67	-0.189
		RENVAVIQV	226-234	9	1027.19	6	0.51	stable	151.11	0.433
VP6	Overlapping HTL and CTL epitopes	LLGTTLLNL	65-73	9	957.18	5.52	-0.54	stable	216.67	1.522
		LLNLDANYV	70-78	9	1034.18	3.8	8.89	stable	173.33	0.622
		NFDTIRLSF	284-292	9	1112.25	5.84	-10.8	stable	86.67	0.1
		HHATVGLTL	316-324	9	948.09	6.92	-0.54	stable	130	0.6
VP4 P8	B-cell linear	VYESTNNSDF	92-101	10	1175.17	3.67	-9.39	stable	29	-1.06
		DFKTLKNLNDNYGITRTEALN	725-745	21	2440.7	6.12	-4.18	stable	79.05	-0.957
		VPTNDDYQ	475-482	8	950.96	3.56	14.04	stable	36.25	-1.675
		FRFSLTV	425-431	7	869.03	9.75	8.57	stable	97.14	1.086
VP7 G1	B-cell linear	TIRNCKKLGPREN	245-257	13	1528.79	10.05	4.18	stable	60	-1.477
VP6	B-cell linear	VEVEFLL	259-265	7	848.01	3.79	8.57	stable	194.29	1.686
		IRLSFQL	288-294	7	876.07	9.75	8.57	stable	167.14	0.871

2.3.6 Construction of Multiepitope-Based Vaccine: Epitope Assembly

Vaccine design and construction was accomplished by systematically integrating a total of 20 predicted overlapping and non-overlapping CTL and HTL epitopes, and linear B-cell epitopes in different orders. In the present study, Beta defensin (GIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK) has been used as an adjuvant. It has been demonstrated that defensin adjuvant-containing vaccines activate the basic innate antiviral immune response and mediate other immunomodulatory effects against a variety of viruses, both *in vivo* and *in vitro* (Kim et al., 2018)(Park et al., 2018). EAAAK linker was used to join the N-terminal β-defensin sequence with CTL epitope. EAAAK (Glu-Ala-Ala-Ala-Lys) is a peptide linker that forms a rigid α-helix and has a closed packed backbone due to intramolecular hydrogen bonding. When contrasted to flexible linkers, rigid linkers have a number of advantages. EAAAK linkers provide effective functional domain separation by maintaining a constant distance between the epitopes with little interference, preserving each one's unique functional characteristics (Chen et al., 2013). This aids in a bifunctional fusion protein's efficient domain separation (Arai et al., 2001). The AAY linker was employed to link CTL and overlapping HTL/CTL epitopes. In mammalian cells, the AAY (Ala-Ala-Tyr) linker functions as a cleavage site for proteasomes. Consequently, epitopes linked with the AAY linker undergo effective disjunction within cells, leading to a reduction in junctional immunogenicity. The multi-epitope vaccine's immunogenicity is further enhanced by the AAY linker (Yang et al., 2015). The KK (Lys-Lys) linker was utilized to connect the HTL epitope with B-cell epitopes. This Lysine-based linker acts as the substrate for Cathepsin B, a lysosomal protease involved in the processing of antigenic peptides for their presentation on the cell surface via MHCII restricted pathways. Additionally, it functions in mitigating junctional immunogenicity by precluding the production of antibodies against the peptide sequence that individual epitopes may induce when linked linearly (Yano et al., 2005). Additionally, KK linkers boost immunogenicity (Li et al., 2016). The connection of linear B-cell epitopes was achieved using the GGGGS (Gly-Gly-Gly-Gly-Ser) linker.

Comprising small or polar amino acids such as Gly and Ser, the GGGGS linker offers notable flexibility and solubility. This property enhances stability and folding, making it an excellent option for fusion protein domains that necessitate precise movements or interactions (Chen et al., 2013). Multiple such arrangements of epitopes were obtained and subjected to in silico characterization to choose the best multiepitope vaccine sequence. The RVAMEV architecture is delineated in **Figure 2.2**.

Figure 2.2. Graphical map of the designed RVA multi-epitope vaccine construct. The



vaccine construct includes (left to right) an adjuvant, CTL, HTL, or overlapping HTL/CTL, and linear B cell epitopes. The adjuvant and the first CTL epitope were linked by the EAAAK linker, CTL and overlapping HTL/CTL epitopes were added together by AAY linkers, HTL epitopes were linked by KK linkers, and linear B-cell epitopes were linked by GGGGS linkers.

2.3.7 In silico Characterization of the RVAMEV

The multiepitope assembled subunit vaccine was characterized for overall antigenicity with Vaxijen 2.0 server (Doytchinova & Flower, 2007) with a threshold of 0.5 and was found to be antigenic with a score of 0.7698. Non-allergenicity of the construct was confirmed using AllerTOP (Dimitrov et al., 2014) and AllergenFP (Dimitrov et al., 2014) webservers. Toxicity of the entire vaccine sequence was predicted using ToxIBTL server and was found to be non-toxic. Further, physicochemical

characterization of the vaccine protein sequence was carried out using ProtParam tool (Wilkins et al., 1999). It contains 313 amino acids with a molecular weight of 33707.5 Da, indicating good antigenic nature. The isoelectric point (p*I*) of RVAMEV was 9.16. RVAMEV was categorized as stable as the instability index was 20.44. The aliphatic index was 92.36 showing the proportional volume of the aliphatic side chains. The protein sequence exhibited a GRAVY value of 0.048, indicating the hydrophobic nature of the vaccine. The half-life of the protein was calculated as greater than 20 hours for yeast, 30 hours for mammalian-reticulocytes, and over 10 hours for *E. coli*. The vaccine protein sequence was further subjected to homology assessment against the human genome and selected gut microbes to rule out any potential cross reactivity using Blastp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) (McGinnis & Madden, 2004) to eliminate the possibility of cross reactivity. Protein sequences from the following organisms were included for BlastP: *Homo sapiens* (taxid 9606), *Lactobacillus johnsonii* (taxid 33959), *Lactobacillus rhamnosus* (taxid 47715), and the *Lactobacillus casei* group (taxid 655183). The results revealed no noteworthy similarities, indicating a lack of substantial homology or potential cross-reactivity between the vaccine protein and the proteome of aforementioned biological entities, except for beta-defensin component of the sequence. This is because beta defensin has been purposefully added in the vaccine sequence as it serves as an adjuvant, augmenting immunogenicity to enhance overall efficacy (**Appendix A, Figure A4**.).

2.3.8 Secondary and Tertiary Structure Prediction of the Vaccine Construct

The PsiPred server (McGuffin et al., 2000) was employed to predict the secondary structure of the multiepitope vaccine, as illustrated in **Figure 2.3**. Solvent accessibility, a fundamental characteristic influencing structural stability, was evaluated using the ProteinSol webserver (http://protein-sol.manchester.ac.uk.) (Hebditch et al., 2017). The solubility of the vaccine was predicted, yielding a score of 0.434, indicative of a value proximal to the average solubility score of 0.5 (**Figure 2.3**). To further elucidate the

molecular architecture, the tertiary structure of the vaccine was constructed utilizing I-TASSER (Roy et al., 2010) and subsequently visualized using the PyMol tool (Schr€odinger, 2015). The three-dimensional representation of the vaccine structure is presented in **Figure 2.4.A**.

2.3.9 3D Structure Refinement and Validation

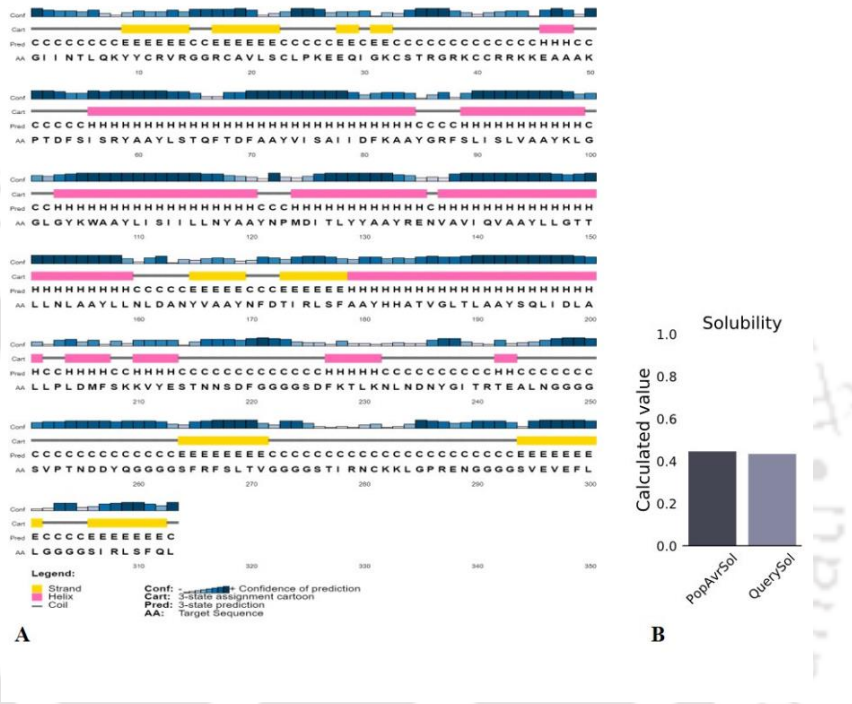


Figure 2.3. RVAMEV. A. secondary structure prediction using Psipred. B. Protein solubility prediction using Protein-sol.

The predicted 3D structure of the constructed vaccine was subjected to structural refinement by GalaxyRefine webserver (http://galaxy.seoklab.org/refine) (Heo et al., 2013). All the generated models were evaluated by diverse three-dimensional structure validation tools to screen the optimal model and assess its overall quality. The tools used for structural validation include Procheck (https://saves.mbi.ucla.edu/), ERRAT server (https://saves.mbi.ucla.edu/) (Colovos & Yeates, 1993) and PROSA web tool (https://prosa.services.came.sbg.ac.at/prosa.php) (Wiederstein & Sippl, 2007). The Ramachandran plot generated by PROCHECK exhibited 92.2% of residues in most

favoured regions, 5.2% residues in additionally allowed regions, 0.4% residues in generously allowed regions and 2.2% residues in disallowed regions. A score of above 90% for most favoured regions in Ramachandran plot indicated the protein structure to be stable and valid. ERRAT overall quality factor was determined to be 93.574.

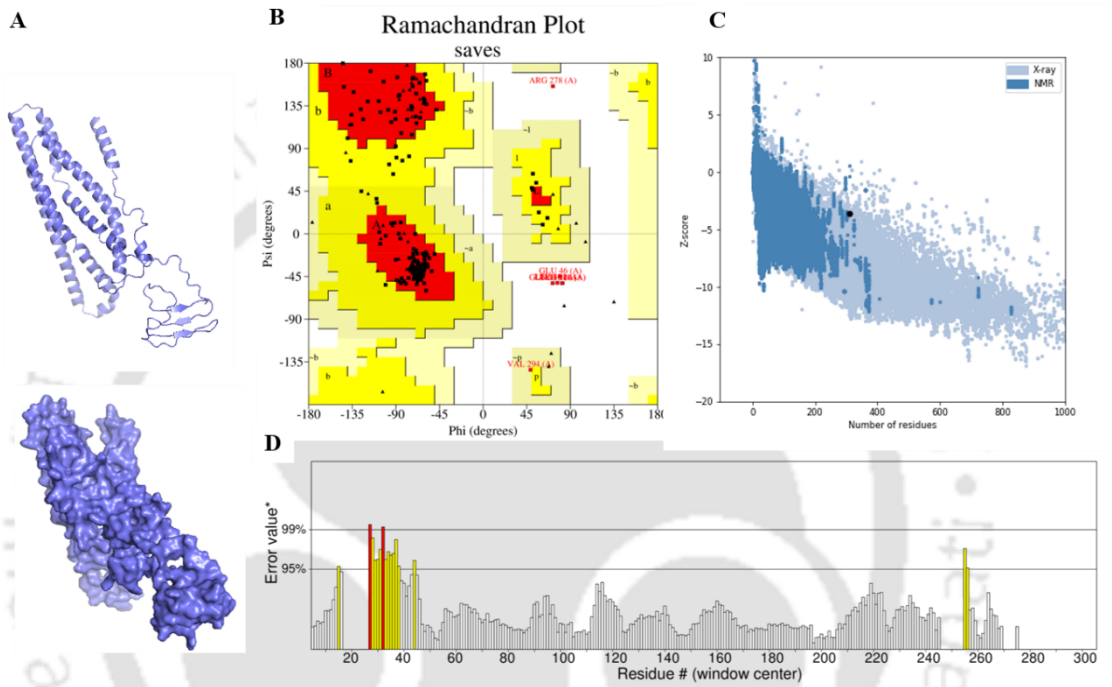


Figure 2.4. Tertiary structure prediction of RVAMEV by I-TASSER (shown in slate blue colour) and 3D structure validation using SAVES server. B. Ramachandran plot: Residues in most favoured regions 92.2%. C. ProSA Z score: -3.56. D. ERRAT2 overall quality factor: 93.574.

ERRAT serves as an "overall quality factor" for non-bonded atomic interactions, with higher scores signifying superior quality. ProSA generated a Z-score of 3.56 with the vaccine structure falling in the range of experimentally determined protein structures. The energy discrepancy between a protein's native fold and the average energy of misfolded structures in a collection is measured by the Z-score, represented in terms of

the standard deviation of this distribution. The validated vaccine three-dimensional structure was subjected to docking studies with integrin receptors and TLRs. **Figure 2.4** presents a summary of vaccine structure validation. Further, in order to ensure the surface accessibility of epitopes, localization or mapping of individual epitopes in the three-dimensional multiepitope vaccine structure was carried out using PyMol tool (Schr€odinger, 2015) and is represented in the **Figure 2.5.** All the selected 20 epitopes for vaccine assembly exhibited surface accessibility in the RVAMEV three-dimensional structure.

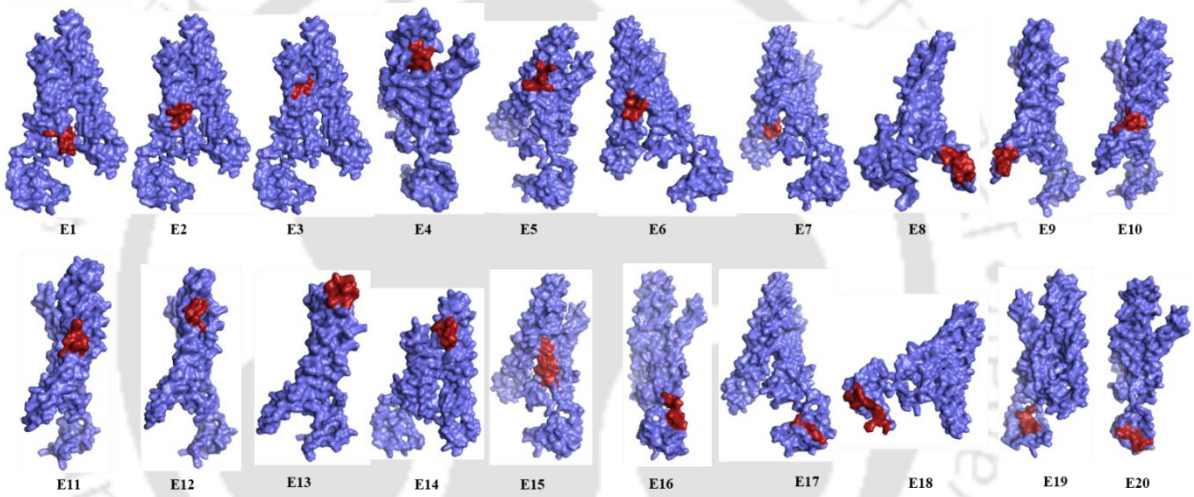


Figure 2.5. Mapping of epitopes (shown in red firebrick colour) in the designed RVAMEV structure (shown in slate blue colour). E1: PTDFSISRY, E2: LSTQFTDF, E3: VISAIIDFK, E4: GRFSLISLV, E5: KLGGLGYKW, E6: LISIILLNY, E7: NPMDITLYY, E8: RENVAVIQV, E9: LLGTTLLNL, E10: LLNLDANYV, E11: NFDTIRLSF, E12: HHATVGLTL, E13: SQLIDLALLPLDMFS, E14: VYESTNNSDF, E15: DFKTLKNLNDNYGITRTEALN, E16: VPTNDDYQ, E17: FRFSLTV, E18: TIRNCKKLGPREN, E19: VEVEFLL, E20: IRLSFQL.

2.3.10 Molecular Docking Analyses of Multiepitope-Based Vaccine with Integrins and Toll-like Receptors (TLRs)

Rotavirus entry into a host cell involves multiple events, including the proteolytic cleavage of the VP4 spike protein into VP5 and VP8 fragments, with interaction involving integrins (αvβ3) (Guerrero et al., 2000)(Zárate et al., 2004). To explore the interactions between the designed RVAMEV and integrins, molecular docking was performed using the vaccine structure with integrin receptors, Integrin αIIbβ3 (PDB ID: 2vdp) and Integrin αVβ3 (PDB ID: 4O02). αVβ3 along with αIIbβ3 (Hynes, 1987) (Varner & Cheresh, 1996) (Byzova et al., 1998), constitute the two members of the β3 subfamily of integrins. They share the same β3 subunit and their α subunits demonstrate a 36% amino acid sequence identity. Integrin αvβ3 is prominently expressed on activated endothelial cells and new-born vessels (Liu et al., 2008), while integrin αIIbβ3 is expressed on the surface of platelets, megakaryocytes, basophils, mast cells, and a subset of tumor cells (Bledzka et al., 2012). Of particular interest, Integrin β3 is also found in the gastrointestinal tract. Studies indicate its involvement in diverse gut processes, such as cell senescence, inflammation, and the regulation of intestinal stem cells (Tun et al., 2023) (Won et al., 2022). TLRs (Toll-like receptors) play a crucial role in adaptive immunity, serving as the initial defence against infections. When they interact with various microbes and viruses, TLRs identify the pathogen-associated molecular patterns (PAMPs) on these microorganisms. This recognition initiates the activation of innate immunity and coordinates the orchestration of the humoral immune response (Pasare & Medzhitov, 2005). TLR2 and TLR4 are present in a variety of immune cells, such as neutrophils, monocytes, macrophages, and dendritic cells (O'Mahony et al., 2008). TLR2/4 on dendritic cells activate helper (CD4+) T-cells by recognizing ligands, inducing chemokine receptor expression, and promoting dendritic cell migration to lymph nodes. This process provides signals for helper T-cell activation through antigen presentation and enhancement of costimulatory molecules. TLR signaling also triggers cytokine production, further driving helper T-cell activation and

differentiation (Duan et al., 2022). The primary location for TLR2 and TLR4 expression within the intestinal mucosa is found in epithelial cells (Frolova et al., 2008). The docking analysis of RVAMEV with integrin αVβ3 demonstrates the formation of 1 salt bridge, 10 hydrogen bonds, and 88 non-bonded contacts with the αV chain, along with 5 salt bridges, 12 hydrogen bonds, and 156 non-bonded contacts with the β3 chain. The predicted binding affinity (ΔG) for these interactions is estimated to be -18.6 kcal/mol. In the case of integrin αIIbβ3, the docking study reveals 2 salt bridges, 6 hydrogen bonds, and 118 non-bonded contacts with the αIIb chain, along with 2 salt bridges, 6 hydrogen bonds, and 60 non-bonded contacts with the β3 chain. The predicted binding affinity (ΔG) for these interactions is estimated to be -13.4 kcal/mol (**Figure 2.6.**).

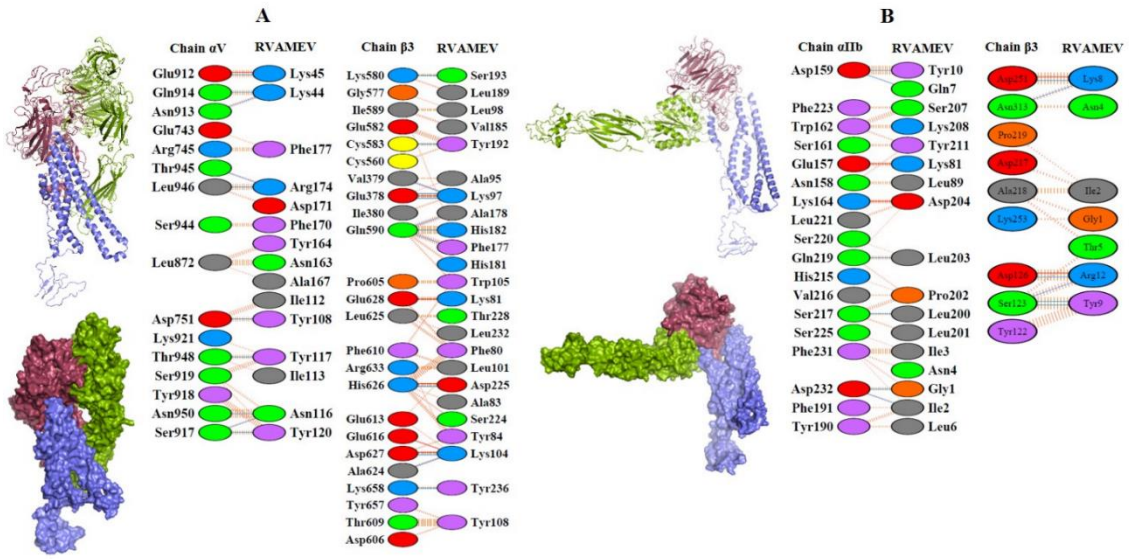


Figure 2.6. Docking of RVAMEV (shown in slate blue color) with A. Integrin αVβ3 (PDB ID: 4O02) (Chain αV shown in splitpea green color and Chain β3 shown in raspberry red color) and B. Integrin αIIbβ3 (PDB ID: 2vdp) (Chain αIIb shown in splitpea green color and Chain β3 shown in raspberry red color).

To assess the interaction of RVAMEV with TLRs, the structurally validated vaccine three-dimensional structure was used for docking with toll-like receptors, TLR 2 (PDB

ID: 5D31) and TLR 4 (PDB ID: 4G8A), known to bind to viral proteins and initiate immune signalling (Lester & Li, 2014). RVAMEV exhibited interaction with both TLR 2 and TLR4. Formation of 1 salt bridge, 5 hydrogen bonds and 139 non-bonded contacts with binding affinity (ΔG) of -199 kcal/mol has been predicted between TLR 2 as a receptor and RVAMEV as a ligand. Docking of TLR 4 chains as a receptor and RVAMEV as a ligand revealed formation of 2 hydrogen bonds and 18 non-bonded contacts with TLR 4 chain A, while 2 salt bridges, 11 hydrogen bonds, and 191 non-bonded contacts with TLR 4 chain B, exhibiting a binding affinity (ΔG) of -19.5 kcal/mol (**Figure 2.7**).

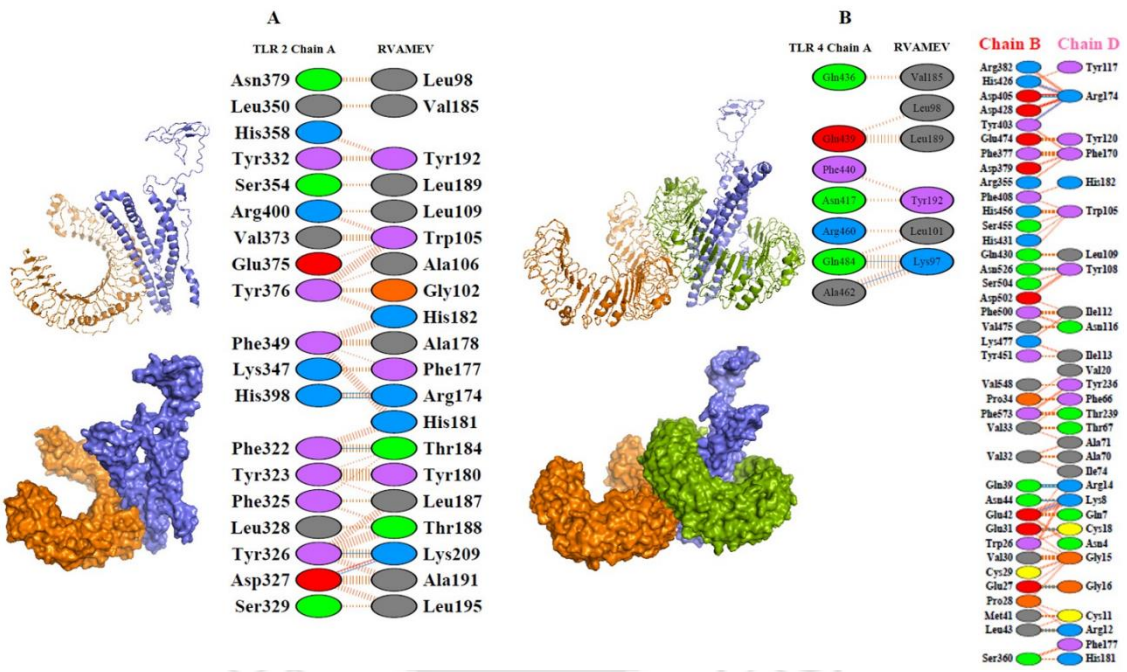


Figure 2.7. Docking of RVAMEV (shown in slate blue color) with A. TLR2 (PDB ID: 5D31) (shown in orange color) and B. TLR4 (PDB ID: 4G8A) (Chain A shown in orange color and Chain B shown in splitpea green color).

2.3.11 Molecular Dynamics Simulation

To evaluate the stability and dynamics of docked vaccine-receptor complexes, molecular dynamics simulations were conducted using the iMODS web server. **Figures 2.8.A, 2.9.A, 2.10.A, and 2.11.A** depict the mobility of the RVAMEV-Integrin αVβ3, RVAMEV- Integrin αIIbβ3, RVAMEV-TLR2 and RVAMEV-TLR4 complexes, respectively, with arrows indicating the direction of movement for each protein residue. Deformability, evidenced by hinges in the main chain resulting from individual residue distortions, was minimal in the RVAMEV- Integrin αVβ3/Integrin αIIbβ3/TLR2/TLR4 complexes, as shown in **Figures 2.8.B, 2.9.B, 2.10.B, and 2.11.B**. Experimental B-factors, derived from PDB field and NMA data, elucidated the relationship between mobility of the docked composite NMA and the PDB score, as depicted in **Figures 2.8.C, 2.9.C, 2.10.C, and 2.11.C**. The eigenvalues of the RVAMEV- Integrin αVβ3, RVAMEV- Integrin αIIbβ3, RVAMEV-TLR2 and RVAMEV-TLR4 complexes were determined to be $3.680451e^{-06}$ (**Figure 2.8.D**), $2.554042e^{-06}$ (**Figure 2.9.D**), $2.684141e^{-06}$ (**Figure 2.10.D**), and $3.912712e^{-06}$ (**Figure 2.11.D**), respectively, confirmed the stability of the complexes. Variance plots associated with each normal mode illustrated individual and cumulative variance, depicted in **Figures 2.8.E, 2.9.E, 2.10.E, and 2.11.E.** Covariance analysis assessed residue coupling in vaccine-receptor complexes, with covariance plots characterizing correlated (red), non-correlated (white), or anti-correlated (blue) atom motions in dynamic regions of complex molecules, as shown in **Figures 2.8.F, 2.9.F, 2.10.F, and 2.11.F**. Additionally, the elastic network model of vaccine-receptor complexes evaluated stiffness, with darker grey indicating stiffer regions and lighter dots signifying flexible regions, displayed in **Figures 2.8.G, 2.9.G, 2.10.G, and 2.11.F**. In summary, these results indicate stable binding interactions characterized by a compact conformation and minimal fluctuations in the vaccine-receptor complexes.

Further, the stability of the docked TLRs and the RVAMEV vaccine construct was assessed using 100 ns MD simulations with GROMACS software. **Figure 2.12A** presents the root mean square deviation (RMSD) plot for TLR2 and the RVAMEV vaccine. The average RMSD values observed for TLR2, RVAMEV, and the docked TLR2-RVAMEV complex were 0.94 nm, 1.86 nm, and 1.15 nm, respectively. These results indicate that TLR2 maintained stability throughout the simulations, whereas the RVAMEV construct displayed flexibility due to the presence of epitopes and linker regions forming loops. The RMSD of RVAMEV stabilized after 80 ns, while the TLR2-RVAMEV complex achieved stability at 20 ns, suggesting that the interaction with TLR2 enhances the stability of the vaccine structure (**Figure 2.12.A**). **Figure 2.12.D** shows the RMSD plot for TLR4 and the RVAMEV vaccine. The average RMSD values for TLR4, RVAMEV, and the docked TLR4-RVAMEV complex were 1.33 nm, 1.04 nm, and 0.74 nm, respectively. These values indicate that TLR4 remained stable throughout the simulations, with the RVAMEV construct again showing flexibility due to the epitopes and linkers. The RMSD of RVAMEV stabilized after 20 ns, while the TLR4-RVAMEV complex reached stability after 80 ns, indicating that interaction with TLR4 further stabilizes the vaccine structure (**Figure 2.12.D**).

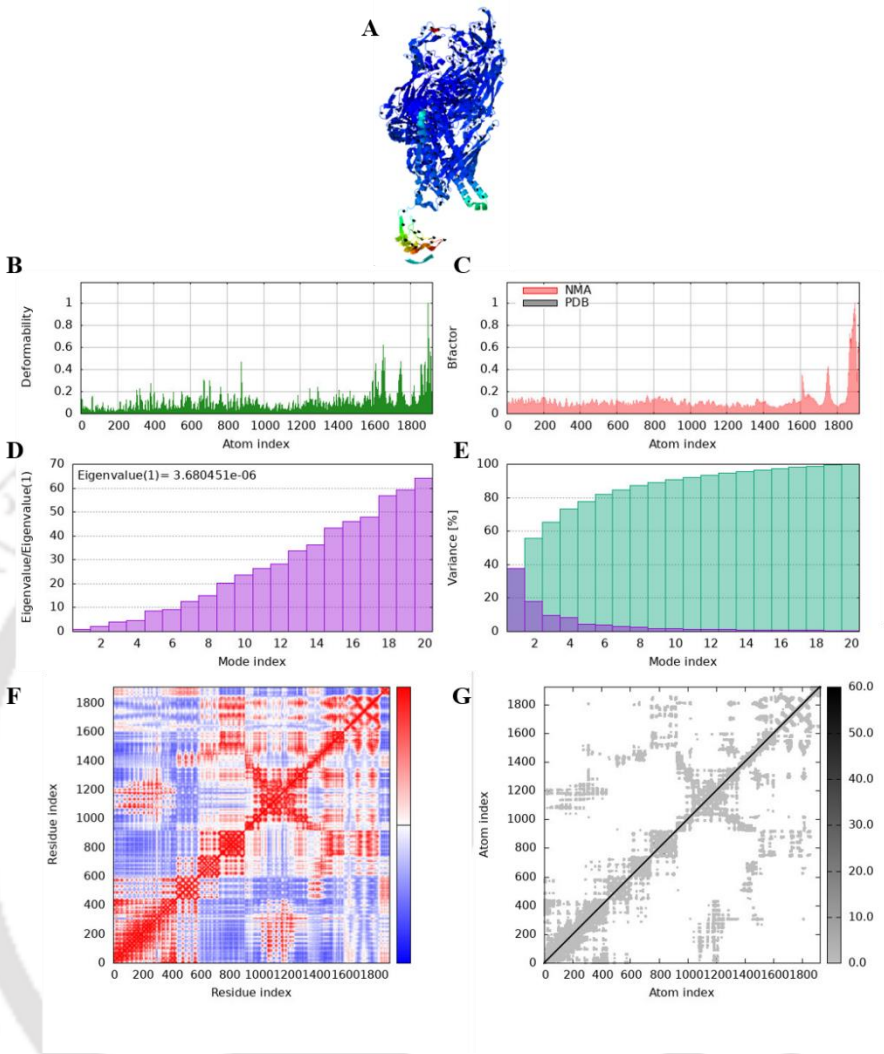


Figure 2.8. Molecular dynamics simulation of RVAMEV-Integrin αVβ3 complex. (A). Mobility of complex, (B). Deformability plot, (C). B-factor plot, (D). Eigen values plot, (E). Structural variance plot where individual (purple) and cumulative (green) variances are shown as coloured bars, (F). Covariance matrix shows correlated (red), non-correlated (white), and anti-correlated (blue) motions of paired residues, (G). Elastic network model where darker grey indicates stiffer regions.

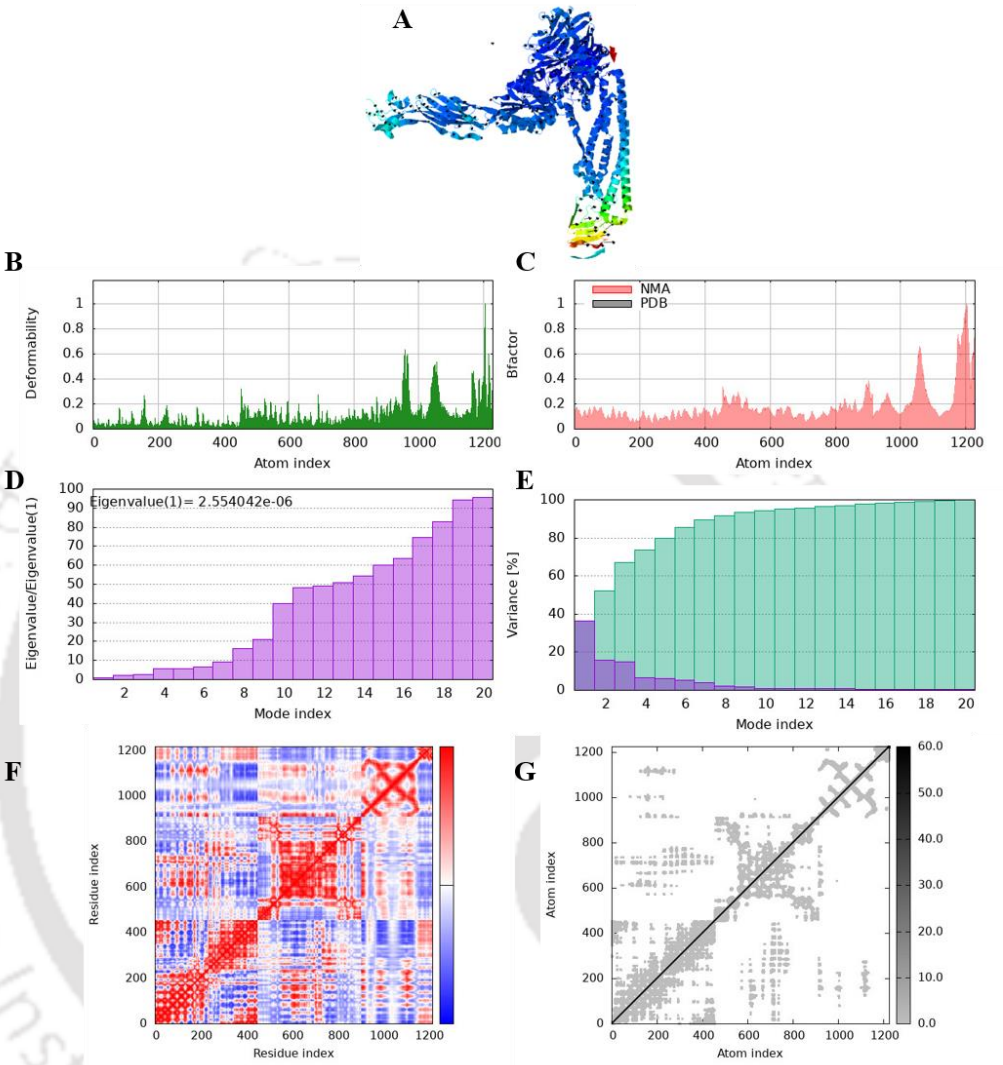


Figure 2.9. Molecular dynamics simulation of RVAMEV-Integrin αIIbβ3 complex. (A). Mobility of complex, (B). Deformability plot, (C). B-factor plot, (D). Eigen values plot, (E). Structural variance plot where individual (purple) and cumulative (green) variances are shown as coloured bars, (F). Covariance matrix shows correlated (red), non-correlated (white), and anti-correlated (blue) motions of paired residues, (G). Elastic network model where darker grey indicates stiffer regions.

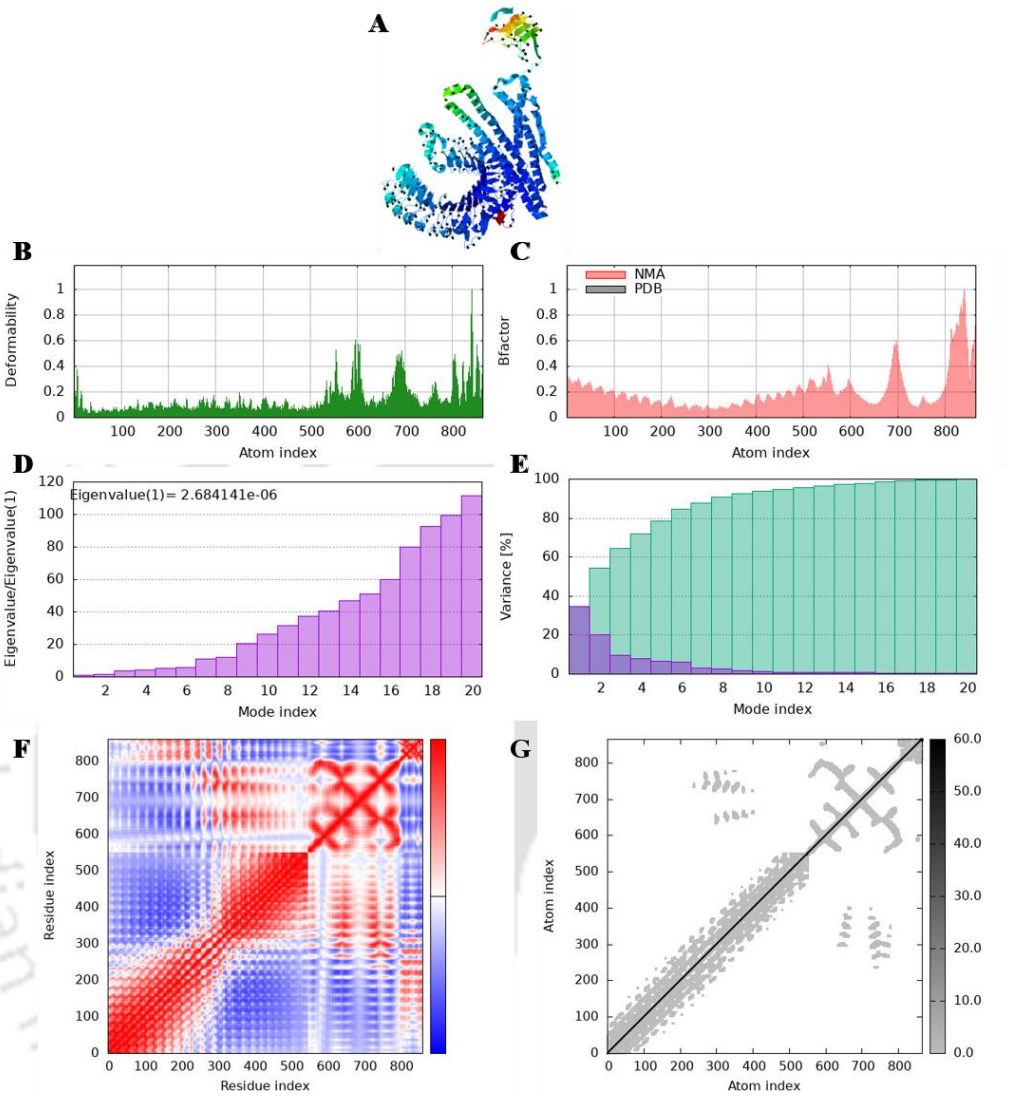


Figure 2.10. Molecular dynamics simulation of RVAMEV-TLR2 complex. (A). Mobility of complex, (B). Deformability plot, (C). B-factor plot, (D). Eigen values plot, (E). Structural variance plot where individual (purple) and cumulative (green) variances are shown as coloured bars, (F). Covariance matrix shows correlated (red), non-correlated (white), and anti-correlated (blue) motions of paired residues, (G). Elastic network model where darker grey indicates stiffer regions.

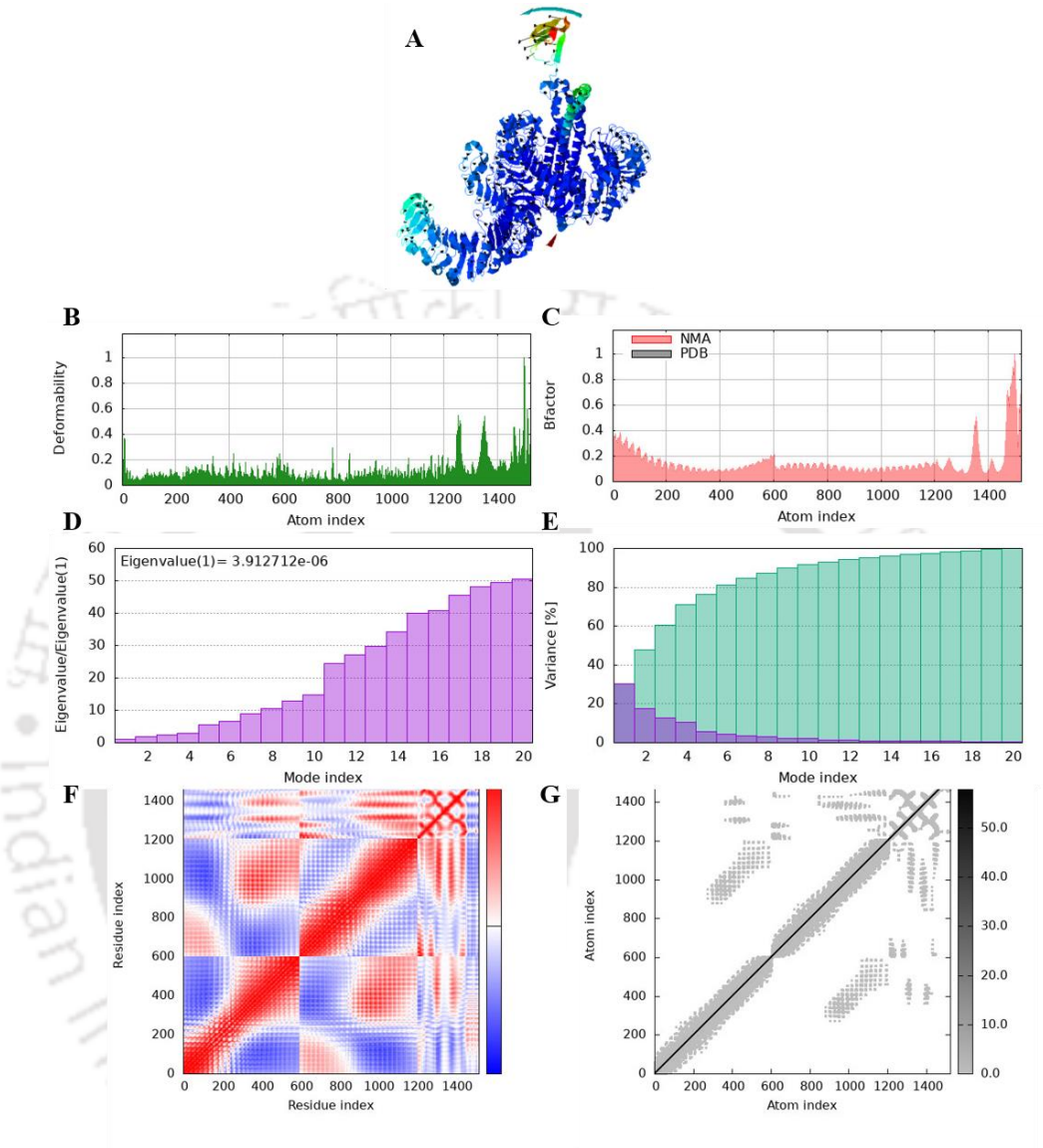


Figure 2.11. Molecular dynamics simulation of RVAMEV-TLR4 complex. (A). Mobility of complex, (B). Deformability plot, (C). B-factor plot, (D). Eigen values plot, (E). Structural variance plot where individual (purple) and cumulative (green) variances are shown as coloured bars, (F). Covariance matrix shows correlated (red), non-correlated (white), and anti-correlated (blue) motions of paired residues, (G). Elastic network model where darker grey indicates stiffer regions.

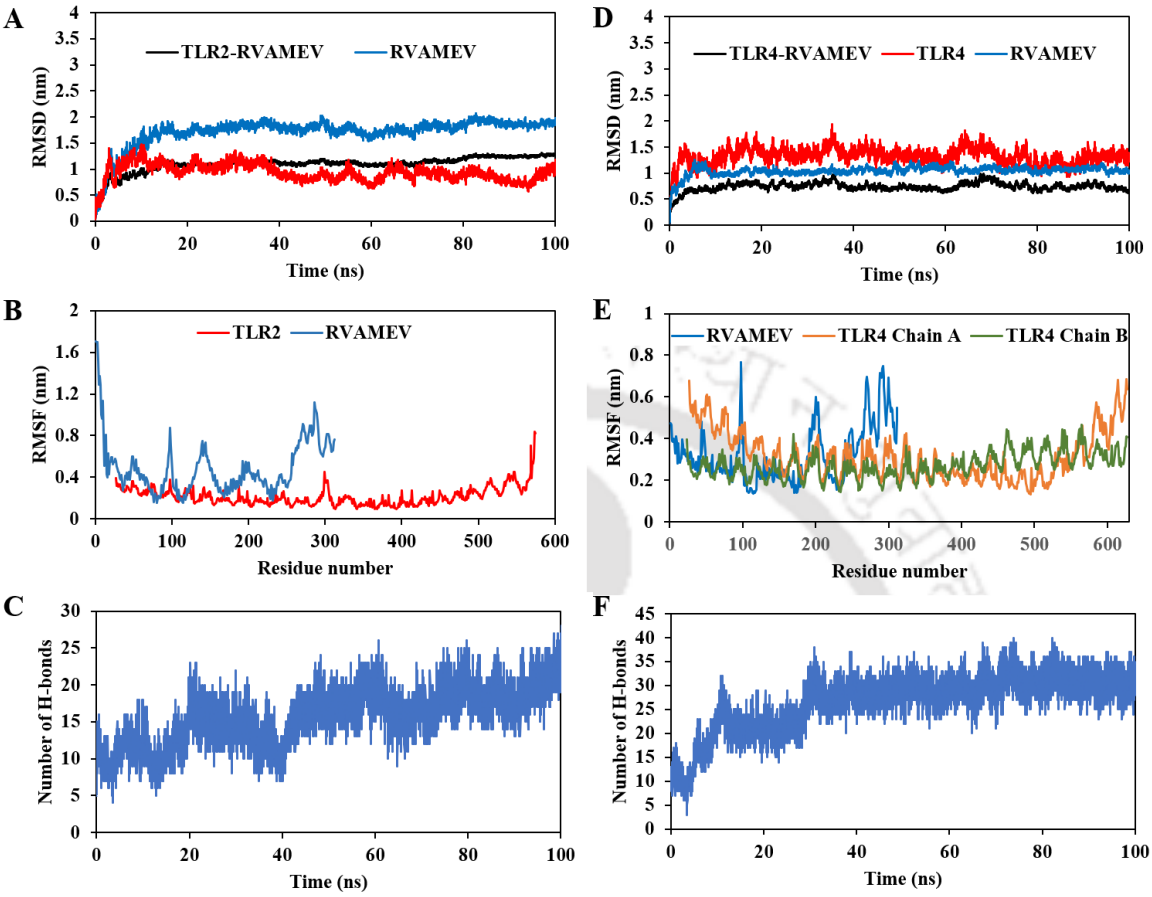


Figure 2.12. Molecular dynamics simulation of TLRs and RVAMEV for 100 ns. (A) Time evolution of the backbone RMSD of TLR2, RVAMEV and TLR2-RVAMEV complex. (B) Backbone RMSF plots of TLR2 and RVAMEV. (C) Time evolution of the number of hydrogen bonds between TLR2 and RVAMEV. (D) Time evolution of the backbone RMSD of TLR4, RVAMEV and TLR4-RVAMEV complex. (E) Backbone RMSF plots of TLR4 and RVAMEV. (F) Time evolution of the number of hydrogen bonds between TLR4 and RVAMEV.

The root mean square fluctuation (RMSF) parameter was analyzed separately for TLR2 and the vaccine constructs to identify regions with high flexibility. **Figure 2.12B** illustrates that the RMSF values for Cα atoms of TLR2 remained stable across the overall structure, except for higher fluctuations in residues 280-310 (associated with

RVAMEV). The average RMSF values were 0.21 nm for TLR2 and 0.48 nm for RVAMEV, underscoring the rigidity of TLR2 and the overall flexibility of the RVAMEV construct (**Figure 2.12.B**). Similarly, the RMSF parameter for TLR4 and the vaccine construct is shown in **Figure 2.12E**. The RMSF values for Cα atoms of TLR4 remained stable, with higher fluctuations observed in residues 250-310 (associated with RVAMEV). The average RMSF values for TLR4 chain A, chain B, and RVAMEV were 0.31 nm, 0.27 nm, and 0.32 nm, respectively, highlighting the rigidity of TLR4 and the flexible nature of the RVAMEV construct (**Figure 2.12.E**).

The hydrogen bond analysis showed that the number of hydrogen bonds between TLR2 and RVAMEV varied from 4 to 27, with an average of 19 hydrogen bonds in the last 20 ns (**Figure 2.12C**). For the TLR4-RVAMEV complex, the number of hydrogen bonds ranged from 3 to 40, with an average of 30 hydrogen bonds in the last 70 ns (**Figure 2.12.F**). The RVAMEV vaccine construct demonstrates both flexibility and stability, which are essential for effective immune recognition and response. The enhanced stability of the RVAMEV construct upon interaction with TLR2 and TLR4 supports its potential efficacy as a vaccine candidate. This comprehensive analysis underscores the promising structural and dynamic properties of the RVAMEV vaccine, paving the way for further experimental validation and development.

2.3.12 Codon Adaptation and in Silico Cloning of Vaccine

In silico cloning of the designed vaccine construct into a suitable vector was conducted to assess the possible and efficient expression of the multi-epitope vaccine protein in a bacterial system such as *E. coli* K-12. Obtaining a codon-optimized version of the reverse-translated cDNA of the vaccine was essential for this purpose. The JCat tool was utilized, resulting in the generation of a 939 bp long cDNA sequence specific to the *E. coli* K-12 strain (Grote et al., 2005). To achieve effective protein expression in *E. coli*, it is essential for the codon adaptation index (CAI) value to be > 0.8, and the GC content must fall within the range of 30–70% (Kirienko et al., 2004). Our construct exhibits a CAI of 0.96 and a GC content of 51.54%, indicating proficient protein

expression in the *E. coli* K-12 bacterial cell system. Subsequently, a recombinant plasmid was created by integrating the codon-adapted RVAMEV cDNA sequences into the pET-28a(+) plasmid (Sathiamoorthy & Shin, 2012) using SnapGene tool version 7.1.0. The codon sequence of the final vaccine, RVAMEV, is highlighted in orange, representing the 939 bp gene sequence generated by the JCat server. This codon sequence is inserted between the XhoI (158) and BamHI (1103) restriction sites of a pET28a+ vector, resulting in a clone with a total length of 6274 bp (**Figure 2.13**).

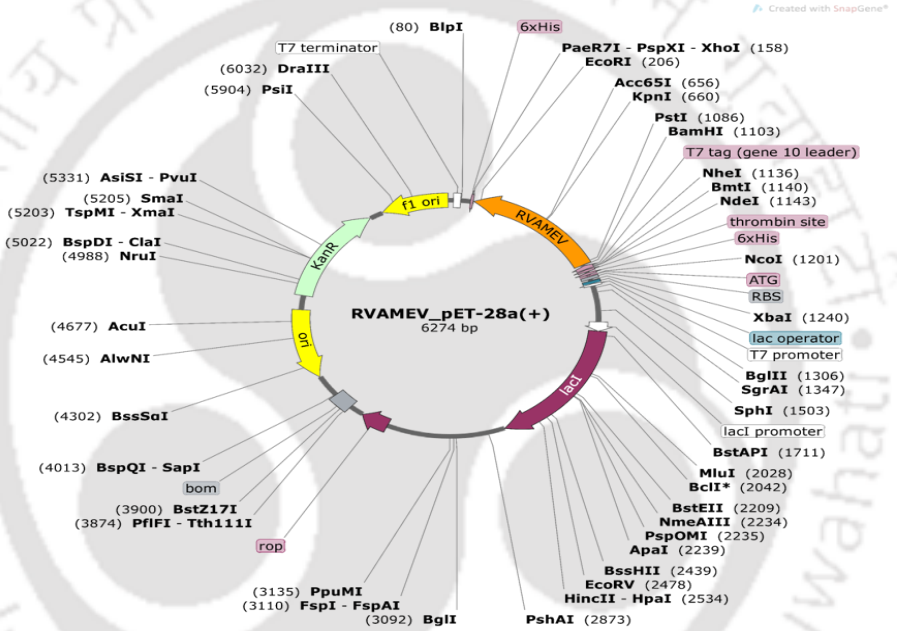


Figure 2.13. In silico cloning of RVAMEV (shown in orange) into pET28a(+) vector.

2.3.13 Immune Simulation and Prediction of the Immune Profile of the Designed Vaccine

In silico immune simulation of the RAMEV was carried out using CIMMSIM server. According to predictions, the antibody titre, encompassing IgM+IgG, IgM, IgG1+G2, and IgG1, shows a noticeable increase during the second and third doses, contrasting with the first dose where no significant increase is observed (**Figure 2.14.A**). Subsequent to each injection, there is a discernible rise in the population of B-memory cells. Additionally, each dose results in an augmentation of B isotypes IgM and subclass

IgG1 (**Figure 2.14.B**). A marginal increase in B isotype IgG2 is also noted in the third injection dose, which aligns with the notion that RVAMEV, as a soluble epitope construct, stimulates these subclass antibodies crucial for effector action. Therefore, it is anticipated that the RVAMEV vaccine may activate B-cells and cause a humoral immune response.

With each administration, RVAMEV elevates the overall population of memory helper T-cells (Th cells) and active Th cells. **Figure 2.14.C** illustrates that while Th memory cells experience a substantial increase during the first and second doses, their count remains relatively stable during the subsequent second and third doses. The initial two injection doses exhibit an increase in the populations of both resting and active Th cells; however, during the third dose, the population of active Th cells increases while the population of resting Th cells essentially remains unchanged (**Figure 2.14.D**). These findings unequivocally demonstrate that the vaccine has the potential to trigger an adaptive immune response. With each injection dosage, RVAMEV boosts the overall population of memory helper T-cells (Th cells) as well as active Th cells. **Figure 2.14.C** illustrates that while Th memory cells exhibit a notable increase during the first and second doses, their count stays relatively unchanged during the second and third doses. The first two injection doses show an increase in the populations of both resting and active Th; however, during the third dose, the population of active Th increases while the population of resting Th essentially stays the same (**Figure 2.14.D**). These findings unambiguously show that the vaccine may trigger the adaptive immune response. Moreover, as shown in **Figure 2.14.E**, the RVAMEV candidate increases the synthesis of TGF-b, Interferon-g, Interferon-b, Interleukin-12, and Interleukin-10.

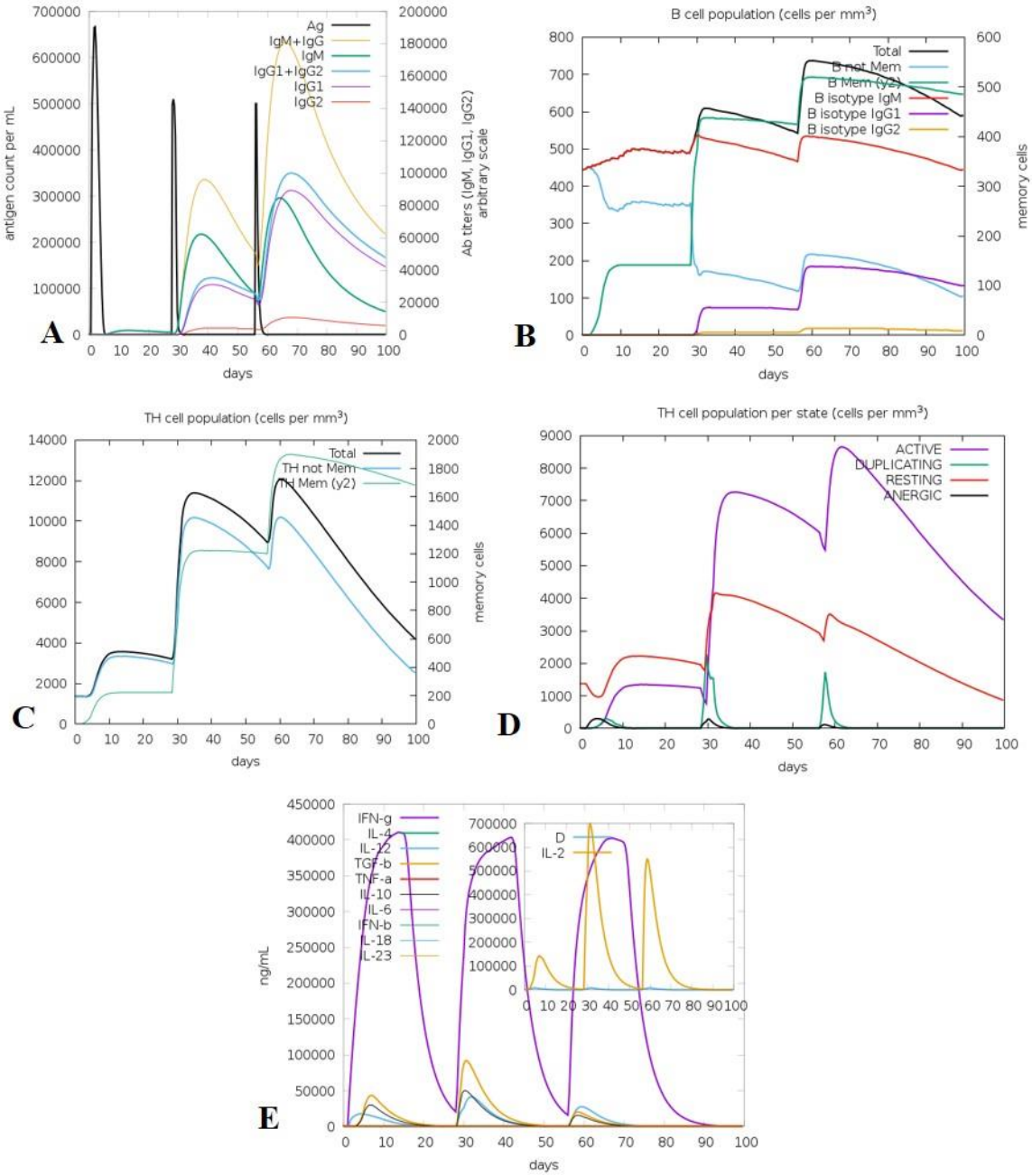


Figure 2.14. Representation of the immune simulation profile for the RVAMEV candidate vaccine, illustrating (A) the count of antigens and antibody titres along with Ig subclass distribution; (B) population dynamics of B-cells; (C) T helper (T_h) cell

activity; (D) distribution of T_h cell populations across various states; and (E) levels of cytokines and interleukins.

After the first two doses, the levels of these cytokines and interleukins rise initially, but after the third dose, their concentration falls compared to the first two shots. In addition, there is a noticeable rise in interleukin-2 levels in the first two doses; however, in the third dose, interleukin-2 levels remain relatively unchanged (**Figure 2.14.E**). Using RVAMEV vaccination, a predominant Th1 immune response profile can be anticipated based on the in silico results. Based on overall immune simulation studies, the proposed RVAMEV candidate vaccine has the potential to induce a robust humoral and cellular immune response.

2.4 Discussion

Reportedly, the efficacy of Rotarix and RotaTeq vaccines against heterotypic RV genotypes is particularly limited, especially in low-income countries including India. Computational immunology stands at the forefront of vaccine design, particularly through reverse vaccinology, amidst the backdrop of evolving viruses and advancements in computational technologies. While a few studies have been conducted to create a subunit vaccine against RVA, region-specific subunit vaccine development has been limited. The substantial burden of rotavirus disease, characterized by the diverse and fluctuating distribution of RV strains across various years and regions, coupled with the suboptimal efficacy of currently licensed vaccines, underscores the necessity for the creation of a vaccine that is both efficacious and safe. This envisioned vaccine should offer region-specific protection and heterotypic immunity against all circulating Rotavirus strains. Empirical evidence derived from investigations into natural infection and the outcomes of monovalent Rotarix vaccination studies substantiates the concept that a human strain can instigate a broad cross-reactive immune response to diverse human genotypes, thereby engendering heterotypic

protection (Velázquez et al., 1996; Bishop et al., 1983; Steele et al., 2012; Jiang et al., 2013).

G1P[8] has been identified as the predominant genotype among Indian rotavirus strains based on phylogenetic analysis in the current study, while a recent epidemiological data from five states of India reports G3P[8] as an emerging strain in the post-vaccination era, alongside continued circulation of G1P[8] (Varghese et al., 2021). While comprehensive epidemiological studies covering all twenty-eight states and eight Union territories of India are currently unavailable, it is essential to evaluate the available data from these five states. Therefore, to assess the relevance of selecting the VP7 G1 genotype in this context, sequence similarity between representative VP7 G1 (AST25359.1) and VP7 G3 (QAR17781.1) sequences was analysed using NCBI protein BLAST, revealing 100% query coverage and 81.90% sequence identity. Notably, two out of three finalized VP7 G1 epitopes (LISIILLNY, NPMDITLYY, TIRNCKKLGPREN) were found to be conserved in VP7 G3, with a single amino acid substitution (Ile17Val) observed in one epitope (LISIILLNY to LISVILLNY). This high similarity suggests potential cross-reactivity, where immunity induced against LISIILLNY may confer partial protection against LISVILLNY. Experimental validation is required to confirm this cross-immunity prospect. Nonetheless, these findings highlight the potential of selected VP7 epitopes to elicit immune responses against both G1 and emerging G3 strains.

The current multivalent vaccine was primarily formulated targeting Indian RVA VP7, VP4 and VP6 capsid proteins. Rotavirus VP7 and VP4 proteins are known to elicit an immune response by inducing production of neutralizing antibodies in the host (Hoshino & Kapikian, 2000). Additionally, the most conserved capsid protein VP6 (Afchangi et al., 2019) has also been subjected to epitope screening. In earlier studies, Devi et al. (Devi et al., 2020) examined all structural and non-structural proteins of group A rotavirus, Dutta et al. (Dutta et al., 2020) focused on VP7 and VP8 proteins, and Jafarpour et al. (Jafarpour et al., 2015) restricted their investigation to VP4 and VP6

proteins, Usman et al. (Usman et al., 2023) used VP4 and VP7, while Sharma et al. (Sharma et al., 2023) focused on VP4 and VP7 specifically from neonatal strains. These studies either used reference sequence of rotavirus genes or global sequences for their study. There has been only one case of region-specific multivalent vaccine development study by Shuvo et al. (Shuvo et al., 2018) where a vaccine targeting the VP4 and VP7 proteins of group A rotavirus prevalent in Bangladesh was developed. Our main objective was to design RVA multi-epitope vaccine with reference to the Indian subcontinent that may also have the potential for heterotypic protection. Phylogenetic study of Indian RVA strains revealed G1 and P[8] to be the most predominant strains circulating in the Indian subcontinent. The most conserved sequences of VP7 (G1), VP4 (P[8]) and VP6 were used for screening of HTL, CTL and linear B-cell epitopes. Multitude of HTL, CTL and linear B-cell epitopes were screened and shortlisted on the acceptable parameters of high conservancy, high antigenic scores, stability, non-allergenic and non-toxic to the host. A total of 20 different epitopes. Beta defensin adjuvant and appropriate linkers (EAAAK, AAY, KK, GGGGS) were used to assemble the epitopes into a complete subunit vaccine. Linkers primarily serve to preserve junctional immunogenicity between epitopes. Cleavable linkers facilitate the presentation of epitopes upon cleavage, while flexible linkers permit mobility. RVAMEV demonstrated an instability index of 20.44, below the threshold value of 40, indicating stability in vitro, along with an aliphatic index of 92.36. These physicochemical attributes underscore the potential of our multivalent construct as a promising vaccine candidate. Additionally, the entire vaccine sequence was predicted to be antigenic with an antigenic score of 0.7698, and was non-allergenic and non-toxic to the human host. Three-dimensional structure of the vaccine was determined using I-TASSER, the structure was refined using GalaxyRefine. the model had 92.2% of residues in most favoured regions in Ramachandran plot and ERRAT score of 93.574 suggesting a stable three-dimensional structure.

The refined structure was docked with integrins and TLRs. Rotavirus entry involves the proteolytic cleavage of the VP4 spike protein into VP5 and VP8 fragments, interacting with integrins (αvβ3) (Guerrero et al., 2000; Zárate et al., 2004). Molecular docking was performed using the RVAMEV vaccine structure with integrin receptors, specifically Integrin αIIbβ3 (PDB ID: 2vdp) and Integrin αVβ3 (PDB ID: 4O02). The predicted binding affinity (ΔG) for interactions involving salt bridges, hydrogen bonds and non-bonded contacts is estimated to be -13.4 kcal/mol. TLRs are crucial contributors to adaptive immunity, constituting the primary defence against infections. Through interactions with a diverse array of microbes and viruses, TLRs identify the pathogen-associated molecular patterns (PAMPs) on these microorganisms. The recognition of PAMPs serves as a stimulus for activating innate immunity and coordinating the orchestration of the humoral immune response. TLR2 and TLR4, members of the TLR family, are known for recognition of viral proteins, triggering the production of inflammatory cytokines (Zhou et al., 2021). RVMEV exhibited strong interactions with both TLR2 and TLR4 with binding affinity of -199 kcal/mol and -19.5 kcal/mol, respectively. MD simulation was employed to appraise the stability of the vaccine protein complex. Utilizing the iMODS server, MD was conducted to scrutinize the essential dynamics and stability of the vaccine, focusing on protein normal modes. The analysis revealed no notable distortions in the structure of the vaccine-receptor complex, suggesting a reduced likelihood of deformability and indicating appropriate stiffness in motion. Further, MD simulation of TLRs and RVAMEV were performed using GROMACS and the analysis revealed stable interaction between the TLRs and RVAMEV.

The data from **Figure 2.14.** demonstrates that the RVAMEV vaccine induces a robust immune response. Antibody levels (IgM+IgG) notably increase during second and third doses, while B memory cells rise after each injection. The vaccine stimulates B-cell isotypes crucial for effector function, suggesting activation of humoral immunity. Additionally, RVAMEV boosts memory helper T-cells and cytokine synthesis,

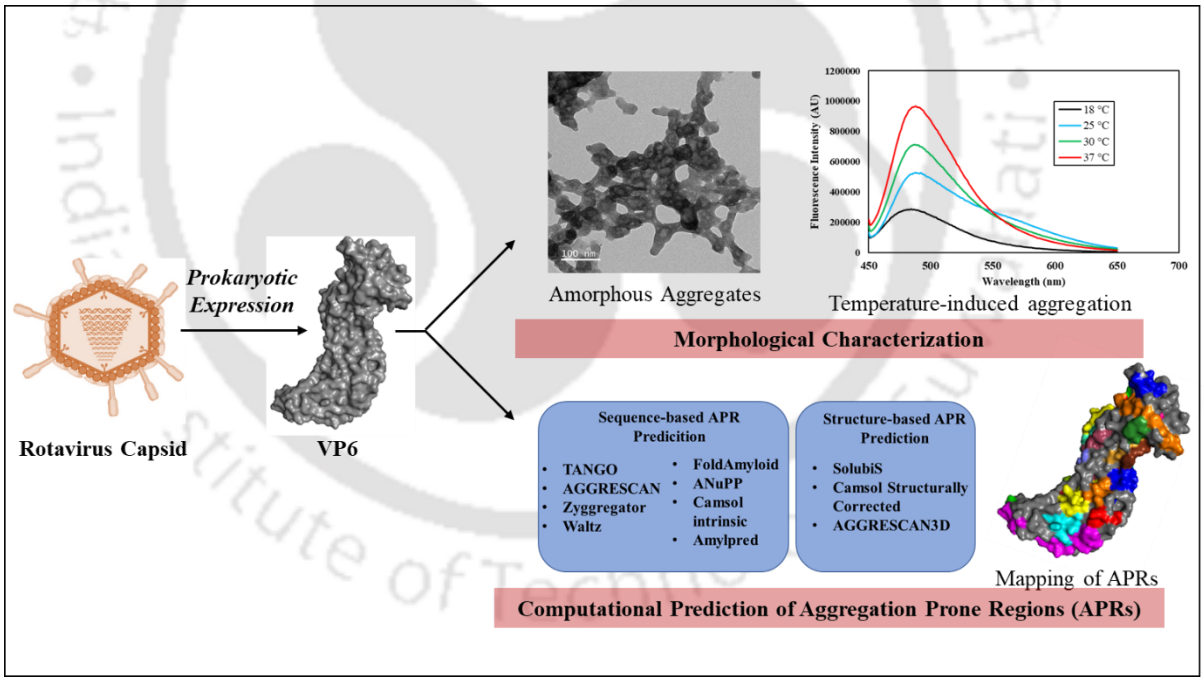
indicating an adaptive immune response. Overall, these findings highlight the vaccine's potential to elicit both humoral and adaptive immunity effectively.

2.5 Conclusion

The study employs reverse vaccinology as an innovative strategy for developing a multi-epitope subunit vaccine targeting Indian RVA, addressing region-specific rotavirus diversity. The investigation identified T-cell (HTL and CTL) and linear B-cell epitopes from conserved capsid proteins (VP7, VP4, and VP6), which were screened, connected with appropriate linkers and an adjuvant to boost immunogenicity. RVAMEV successfully underwent rigorous physicochemical characterization, structural validation, exhibited interactions with integrins and TLRs, and was cloned in silico into a pET28a+ vector for efficient bacterial expression. In silico immune simulation studies suggested its potential to stimulate both humoral and cell-mediated immunity. Computational analysis affirmed the safety, non-allergenicity, and non-toxicity of the vaccine construct, contrasting with live attenuated vaccines associated with known side effects. Results indicated favorable safety, non-allergenicity, and non-toxicity profiles, and the efficacy of the multivalent vaccine in activating cell-mediated as well as humoral immune systems for the RVAMEV. The proposed subunit vaccine candidate appears theoretically overcomes limitations of live-attenuated vaccines, prompting additional *in vitro* and *in vivo* testing to validate computational findings. This study marks the first attempt to develop a region-specific vaccine for the Indian subcontinent.

Chapter 3

Integrated Analysis of Rotavirus A VP6: *E. coli* Expression Optimization, Characterization of Inclusion Bodies, and In Silico Prediction of Aggregation-Prone Regions



3.1 Overview

VP6 is a structural protein, present in the middle of the triple layered capsids of the virion. Notably, Rotavirus structural proteins govern host specificity, cellular entry, enzymatic machinery for viral transcript synthesis, and harbor immunological epitopes. Situated in the middle layer, 260 VP6 proteins engage in extensive interactions with both the outer layer (VP4 and VP7) and the inner layer (VP2) (Jayaram et al., 2004; Crawford et al., 2001; Settembre et al., 2010; Charpilienne et al., 2002). These interactions are essential for the assembly of the viral capsid and transcriptase activity. The integrity of these layers and the orchestrated interplay between their constituent proteins are fundamental for the overall rotavirus structure.

The VP6 protein, which constitutes 51% of the total rotavirus virion mass (Estes, 2007) demonstrates a hierarchical arrangement when expressed *in vitro* in appropriate expression systems. The native VP6 protein exhibits self-assembly with structural polymorphism and forms trimers, hexagonal structures and further assembles into diverse architectures such as nanotubes, nanospheres as VLPs, or nanosheets (Estes, 2007; Estes et al., 1987; Bugli et al., 2014; Choi et al., 1999; Kapikian & Hoshino, 2001). Such polymorphic VP6 self-assembly is influenced by factors such as pH and ionic strength of divalent cations (Ca^{2+} and Mg^{2+}) (Lepault et al., 2001; Ready & Sabara, 1987).

Recombinant expression of VP6 has been pivotal in advancing the understanding of rotavirus pathogenesis, vaccine development, and diagnostic applications. Various expression systems, including prokaryotic models and several eukaryotic platforms, have been employed for VP6 production. Despite the diversity of these systems, prokaryotic expression remains the most accessible and cost-effective method for laboratory-scale and industrial purposes. However, VP6 expression in prokaryotic hosts often results in the formation of inclusion bodies, posing a significant challenge to its research and application. This underscores the need for an in-depth structural understanding of VP6 aggregation propensities, which may contribute to inclusion body formation. Characterizing VP6's aggregation propensity can facilitate understanding of its behavior during prokaryotic expression and open avenues for protein engineering of

soluble variants, thereby advancing research on rotavirus VP6 in pathology, therapy, and diagnostics.

The establishment of protein's native conformation mostly involves intramolecular non-covalent interactions. However, the high translation rate in cells continuously generates a large number of unfolded polypeptides, exposing hydrophobic side chains and hydrogen bond donors/acceptors to the surrounding solvent. This enables both intramolecular and intermolecular interactions of protein segments, leading to the formation of either native proteins or aggregated proteins, respectively (Radford, 2008). Consequently, protein aggregation, forming inclusion bodies, dominates in many cases of heterologous protein expression in prokaryotes, such as *E. coli* (Jahn & Radford, 2005). Because protein aggregation is associated with structural compromise of the otherwise native structure, the inclusion bodies are significantly rendered non-functional. Recovery of functional target protein necessitates unfolding and refolding procedures, which can be cumbersome and inefficient (Ventura & Villaverde, 2006). Protein aggregation is facilitated by specific sequences in the polypeptide chain that initiate the aggregation process, leading to the enrichment of intermolecular beta-sheet structures. While amyloid-like aggregates are characterized by cross-beta sheet structures, protein aggregates can encompass both amorphous and amyloid-like forms. (Willbold et al., 2021; Narayanan et al., 2004). These sequential determinants of protein aggregation are called aggregation-prone regions (APRs) or aggregation hotspots (Wang et al., 2008; Narayanan et al., 2004; Rousseau et al., 2004; Conchillo-solé et al., 2007). The identification of APRs can aid in protein engineering through sequence modification of the target protein, achieved by mutating residues in aggregation hotspots. This approach aims to enhance soluble expression of the target protein by reducing its aggregation propensity and increasing its solubility (Ventura & Villaverde, 2006; Ventura, 2005). Insights into the mechanisms of aggregation, derived from experimental and computational findings, may facilitate the development of reliable algorithms for prediction of APRs.

In this work, computational prediction of APRs in the VP6 protein of RVA has been performed in response to the observation of VP6 inclusion body formation during heterologous recombinant expression in *E. coli*. The determination of the intrinsic aggregation propensity of VP6 can contribute to our understanding of its aggregation

mechanisms and open avenues for future protein engineering efforts to generate more soluble forms during recombinant expression. This, in turn, can support fundamental and practical research concerning the pathology, therapy, or diagnostics of rotavirus VP6. A series of computational APR prediction tools were employed here to analyze both the sequence and structure of VP6. Various algorithms were selected to mitigate biases stemming from training sets, parameterization, and the unique characteristics inherent to each method. Utilizing this approach, diverse sequence-based as well as structure-based computational methods were applied. These computational techniques play a pivotal role in forecasting APRs within proteins, which is crucial for both protein characterization and biotechnological applications. Emphasis was placed on APRs identified by two or more prediction tools, underscoring their potential significance in elucidating protein behavior and devising strategies for protein solubilization and aggregation inhibition.

3.2 Methodology Section

3.2.1 Chemicals and Reagents

The RVA VP6 gene sequence was codon optimized for bacterial expression system, synthesized, then cloned at the Xho I and BamH I restriction sites of the pET-28a(+) plasmid by Genscript®. Restriction enzymes were procured by New England Biolabs and Taq Red PCR mix was obtained from Sigma Aldrich respectively. The gene specific primers were synthesized from Bioserve (India). Nickel- Nitrilotriacetic acid (Ni-NTA) Hi-TRAP column was procured from GE healthcare. The mouse anti-RV inner capsid protein primary antibody and anti-mouse IgG specific-peroxidase secondary antibody was obtained from Novus Biologicals. The 3,3'-Diaminobenzidine was obtained from Amresco. The polyvinylidene fluoride (PVDF) membrane (Hybond P) was procured from Amershan. Thioflavin T was procured from Sigma Aldrich. All other chemicals used in the current work were of analytical grade. All solutions were prepared using nanopure water, MilliQ (18.2 MΩ.cm), (Millipore Co., USA).

3.2.2 Cell Transformation and Optimization of VP6 Expression in *E. coli* BL21(DE3)

Competent *E. coli* BL21(DE3) cells were prepared by $CaCl_2$ treatment. The pET-28a(+)-RVAVP6 recombinant plasmid was transformed into the prepared competent *E. coli* BL21(DE3) cells with a heat-shock treatment at 42 °C for 90s. These transformed cells were grown in Luria Bertani (LB) broth (HiMedia, India) supplemented with kanamycin (Sigma) (50µg/ml) on shaker-incubator at 37°C (180 rpm) for 1 hour followed by plating on Luria-Bertan agar plate supplemented with kanamycin (50µg/ml), isopropyl beta-D-thiogalactoside (IPTG) (0.1 mM) and X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) (40 µg/ml) followed by incubation at 37 °C for about 16 hours and observed for blue-white screening and validated with restriction digestion (NdeI and/or XhoI) and PCR using gene specific primers.

The recombinant *E. coli* BL21(DE3) clones were subjected to optimization of RVAVP6 protein expression. Recombinant *E. coli* expression host, BL21(DE3), was inoculated in 5 mL (1:100) of LB medium supplemented 50µg/ml of kanamycin and cultivated overnight at 37°C in shaker-incubator. This preculture was used to inoculate (1:100) different sets of 10 mL LB medium, which were cultivated at 37 °C until reaching a 600 nm optical density (OD) of 0.6–0.8, followed by induction with IPTG (0.5 mM) and post induction temperature of 37 °C for 12 hours. Cells were collected by centrifugation at 10,000 *g* at 4°C for 10 minutes, the culture medium was discarded and cells resuspended in lysis buffer containing 300 mM NaCl, 50 mM sodium phosphate buffer (SPB), pH 8. The homogenates were sonicated for 10 times for 60 seconds, with 30 seconds' interval between each sonication, and then centrifuged at 12,000 rpm to separate supernatants and pellets. The crude supernatant and pellet fractions (pellets resuspended in resuspension buffer consisting of 300 mM NaCl, 50 mM SPB, pH 8) from bacterial protein extractions were confirmed by Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis (SDS-PAGE) analysis.

To optimize the overexpression of the recombinant protein (VP6), several induction conditions - inducer concentration (0.0 mM to 1M), post-induction temperature (18 °C, 25 °C, 30 °C and 37 °C), post-induction incubation time with and without PMSF treatment (0.2mM, 0.4mM, 0.6mM and 0.8mM) during protein extraction, and urea treatment were tested.

3.2.3 VP6 purification and Western Blot

The over-expressed recombinant VP6 protein localized as inclusion bodies in the pellet fraction or the insoluble fraction of the protein extract as revealed by SDS-PAGE analysis. Therefore, recombinant VP6 protein was purified by Ni-NTA affinity chromatography after solubilizing the inclusion bodies in a urea solubilization buffer (50 mM SPB, pH 8.0, 300 mM NaCl, 8M Urea). Briefly, after equilibrating of the Ni-NTA column was carried out using lysis equilibration buffer (50 mM SPB, pH 8.0, 300 mM NaCl, 8M Urea), followed by passage of an equal volume of lysis equilibration buffer with urea-treated cell lysate agarose column after equilibration. For on-column renaturation of VP6, the column was washed with five bed volumes of wash buffers with step-down urea concentrations (50 mM SPB, pH 8.0, 300 mM NaCl, 8 M to 0 M Urea, 160 mM Imidazole). The recombinant VP6 protein was subsequently eluted with two bed volumes of elution buffer (50 mM SPB, pH 8.0, 300 mM NaCl, 500 mM Imidazole). The eluates collected at each step were analyzed by SDS-PAGE. The elution with purified recombinant VP6 was subjected to western blotting.

Briefly, recombinant VP6 protein was subjected to SDS-PAGE and protein bands were transferred over Polyvinylidene difluoride (PVDF) membrane (MDI, India) using constant voltage of 25 V, 300 mA, for 3 to 4 hours at 4 °C. The unoccupied sites on PVDF membrane were blocked by 3% BSA at room temperature for 3 hours. The protein bands were allowed to bind with mouse anti-RV inner capsid protein monoclonal antibody (NB110-37243) (1:1000 in 3% BSA) at 4 °C overnight, followed by incubation with anti-mouse IgG specific-peroxidase secondary antibody (1:5000 in 3% BSA) after washing with PBST (0.1% Tween-20 in PBS). The protein bands were detected using 3,3-Diaminobenzidine (DAB) (Sigma-Aldrich, USA).

3.2.4 Field Emission Transmission Electron Microscopy (FETEM) of VP6 Inclusion Bodies and Purified VP6

Samples for FETEM analysis constituted the VP6 inclusion body and purified renatured VP6 protein. Each sample was diluted, and then 5 µL was placed on a carbon-coated copper grid, followed by staining with 2% uranyl acetate solution which was dried at room temperature using vacuum pump desiccator for 8 hours. The grids with the sample were examined on JEOL JEM- 2100F at an accelerating voltage of 100 kV.

3.2.5 Thioflavin T (ThT) Binding Assay

VP6 was expressed in BL21(DE3) at four different temperatures separately with optimized inducer concentration for 8 hours. Cells were harvested by centrifugation at 10,000 g for 10 min (4 °C). Cell pellets were resuspended in buffer 1 (50 mM SPB, pH 8.0, 300 mM NaCl) to form suspensions of equal OD 600 nm. Inclusion bodies were purified using several sonication and washing steps. Purified VP6 inclusion bodies purified from equal quantity of cells were then subjected to ThT assay to study amyloid nature of VP6 inclusion bodies expressed at different temperatures. Briefly, inclusion bodies expressed at various temperatures were diluted in 50 mM SPB (pH 8.0) containing 300 mM NaCl to achieve uniform optical density suspensions (measured at 350 nm, OD 350 nm = 0.3). Purified inclusion bodies were incubated with 50 mM ThT for 15 min. Fluorescence spectra were recorded on FluoroMax-4 (HORIBA Scientific). Samples were excited at $\lambda_{440\ nm}$ and emission spectra were recorded in the range of $\lambda_{450–650\ nm}$ with excitation and emission slit widths set to 5 nm. Data were smoothed using 5-point Savitzky-Golay smoothing. ThT without protein served as negative control. Difference spectra (after deducting spectra of negative control) were used for analysis. The final analysis was based on the average of three independent spectra.

3.2.6 Physicochemical Characterization and Structure Prediction of Rotavirus A VP6 Protein

Human RVA VP6 protein sequence (accession number YP_002302229.1) in FASTA format was retrieved from the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/) (Sayers et al., 2022). The physicochemical characterization was performed using ProtParam Expasy tool (https://web.expasy.org/protparam/) (Gasteiger et al., 2005). The Protein-Sol webserver (https://protein-sol.manchester.ac.uk/) was used to predict protein solubility of VP6 in *E. coli* expression system (Hebditch et al., 2017).

Secondary structure was predicted by SOPMA. The three-dimensional structure of the chosen VP6 protein sequence has not been determined experimentally to date. However, crystal structure of VP6 isolated from Bovine rotavirus strain RF is available in RSCB PDB (1QHD) (Berman et al., 2000). This structure exhibits 96.98% amino acid sequence identity, as confirmed by NCBI protein BLAST analysis and has therefore been used as a template for modelling the structure of VP6 under study using

SWISS-MODEL (Waterhouse et al., 2018) (https://swissmodel.expasy.org). The generated three dimensional structure was subjected to structural validation using the tools Procheck (Laskowski et al., 1993), ProSA (Wiederstein & Sippl, 2007), Verify 3D (Eisenberg & Lüthy, 1997) and ERRAT (Colovos & Yeates, 1993) and were visualized using the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC. Further, this structure was compared with the one generated using AlphaFold2 (Jumper et al., 2021) (https://cosmic-cryoem.org/tools/alphafold2/). To compare the quality of the predicted structures, we analyzed their Ramachandran plots using the MolProbity server (http://molprobity.biochem.duke.edu/). Additionally, we performed pairwise structure alignment using the RCSB PDB alignment tool (rcsb.org/alignment) to assess the structural congruence with the reference structure (PDB ID 1QHD).

3.2.7 Sequence-based Detection of APRs in VP6

The amino acid sequence of VP6 was subjected to analysis using a range of sequence-based APR detection tools. These tools employ diverse prediction methodologies, which are described in the subsequent sections.

3.2.7.1. TANGO

TANGO employs a statistical mechanics-based algorithm to predict the nucleating regions for protein aggregation. Additionally, it analyzes the impact of mutations and environmental conditions on the propensity of these regions to undergo aggregation (Buck et al., 2012). TANGO conducts computations to evaluate the occurrence of cross-beta aggregation in peptides and denatured proteins. It considers the relative propensities of different structural states, including beta-turn, alpha-helix, and beta-sheet, which compete with each other (Buck et al., 2012.). To determine the propensity of each conformation, TANGO utilizes a partition function and predicts cross-beta aggregation. This prediction assumes that the core regions involved in the aggregation process are fully buried and fulfil their hydrogen-bond potential. The TANGO aggregation score considers environmental factors such as stability, pH, ionic strength, protein concentration, and the presence of denaturant trifluoroethanol (TFE). The default parameters of pH, temperature, and ionic strength were used for APR prediction of VP6. In terms of data interpretation, any residue exhibiting an aggregation score exceeding 5% over a span of 5-6 residues is considered a potential APR.

3.2.7.2. AGGRESCAN

AGGRESCAN functions as a computational framework aimed at the recognition of aggregation-nucleating segments, commonly referred to as "aggregation hot spots," in polypeptides. This platform relies on a scale of aggregation propensity, which is established based on empirically validated hotspots observed in a diverse assortment of proteins, encompassing both natively unfolded and pathogenic variants such as Ab42, synuclein, prion, and amylin. By harnessing this scale, AGGRESCAN enables the precise identification and characterization of specific regions within protein sequences that exhibit a heightened likelihood for aggregation. For APR utilizing AGGRESCAN, the sequence of VP6 was submitted to the AGGRESCAN database (http://bioinf.uab.es/aggrescan/) (Conchillo-solé et al., 2007). Various parameters were obtained, including the aggregation-propensity value (a^3v) for each amino acid, the average of a^3v within a sliding window (a^4v), a graphical representation of the protein's aggregation profile, the area of the aggregation profile above the hotspot threshold (HST) in a specific hotspot, and a graphical depiction of the peak area. Additionally, putative aggregation hot spots were identified, defined as regions containing five or more residues with an a4v value surpassing the HST. The Hotspot Area (HSA) was calculated as the aggregation profile area divided by the number of residues in the entry sequence. Furthermore, the Normalized Hotspot Area (NHSA) per residue was determined by dividing the HSA by the number of residues in the entry sequence.

3.2.7.3. Zyggregator

Zyggregator is a method based on the protein's sequence that computes various characteristics, including the local stability of the individual units (ln *P*i), the formation likelihood of β-rich oligomers (Z_i^{tox}), and the propensity for forming fibrillar aggregates (Z_i^{agg}). These assessments aid in predicting the inherent amyloid aggregation tendency of the protein. The Z_i^{agg} score of 0 corresponds to an aggregation propensity equivalent to that of a random sequence at position *i*, while a score of 1 indicates a one standard deviation increase in the propensity for aggregation. The presence of a Z_i^{agg} score of 1 or higher indicates the involvement of aggregation-prone residues. For the analysis of VP6 using Zyggregator, the sequence of VP6 was submitted to the Zyggregator database, and the aforementioned parameters were computed. Subsequently, a graphical representation was created based on the Z_i^{agg} scores of the individual residues (Tartaglia & Vendruscolo, 2008).

3.2.7.4. Waltz

WALTZ is a computational algorithm designed to predict amyloids in proteins effectively. It demonstrates the ability to differentiate genuine amyloid structures from disordered amorphous aggregates. It is capable of discerning challenging amyloid aggregates, such as those found in prion diseases, which are typically difficult to predict (Oliveberg, 2010). WALTZ employs a comprehensive three-component scoring function. Firstly, it utilizes a position-specific scoring matrix (PSSM) focused on hexapeptides to identify regions prone to amyloid formation. The PSSM is derived from the AmylHex dataset, enriched with numerous experimentally determined amyloidogenic and non-amyloidogenic sequences. Secondly, WALTZ considers the physicochemical properties of amino acids, including hydrophobicity and β-structure-forming propensity. Lastly, it incorporates a position-specific pseudoenergy matrix, obtained from the Sup35 GNNQQNY peptide's crystal structure, which consist of cross-β spine of amyloid fibrils. This matrix, in combination with FoldX, estimates the relative energies involved in the process (Ahmed & Kajava, 2013)(Meric et al., 2017). The training of the WALTZ algorithm involved an extensive dataset comprising experimentally validated amyloid-forming peptides. Approximately eighty percent of the dataset is validated using techniques such as circular dichroism, electron microscopy, infrared spectroscopy (FTIR), and Thioflavin-T binding assays (Meric et al., 2017). The prediction of amyloid-forming peptides can be conducted using any of the three pre-existing thresholds available in the WALTZ database, namely "Best overall performance," "High specificity," and "High sensitivity." Alternatively, customized threshold may also be set. To analyze VP6 using WALTZ, the amino acid sequence of VP6 in FASTA format was submitted to the online WALTZ web server (http://waltz.switchlab.org/) using “Best overall performance” as a threshold.

3.2.7.5. FoldAmyloid

The FoldAmyloid webserver predicts amyloidogenic regions in protein sequences based on two key characteristics: expected probability of hydrogen bond formation and expected packing density of residues. High probabilities of backbone-backbone hydrogen bond formation and dense packing are indicative of amyloidogenic potential. The FoldAmyloid server is accessible at http://antares.protres.ru/fold-amyloid/ (Garbuzynskiy et al., 2009). The primary protein sequence of VP6 was submitted to FoldAmyloid webserver with an averaging frame of 5 and threshold set to 21.4.

3.2.7.6. ANuPP

ANuPP is an ensemble classifier designed to discern potential APRs within peptides and proteins, employing atomic-level features derived from hexapeptides. This method demonstrated an accuracy rate of 83% in its predictions. The ANuPP server is accessible at https://web.iitm.ac.in/bioinfo2/ANuPP/ (Prabakaran et al., 2021). RVAVP6 primary amino acid sequence was submitted in FASTA format with a default threshold of 0.52.

3.2.7.7. Camsol intrinsic

CamSol comprises two distinct algorithms, one based on sequence analysis and the other on protein structure analysis, to accurately determine protein solubility. The initial algorithm, referred to as CamSolintrinsic, focuses on predicting the inherent solubility of the unfolded state of the protein using the amino acid sequence of a protein. The intrinsic solubility analysis of VP6 protein was conducted by submitting its amino acid sequence to CamSolintrinsic web server (http://www-vendruscolo.ch.cam.ac.uk/camsolmethod.html) (Sormanni et al., 2015).

3.2.7.8. Amylpred

Amylpred employs a consensus approach combining multiple strategies to predict amyloid fibril formation features associated with diseases like Alzheimer's, Parkinson's, type II diabetes, and prion disease. This tool also assists in understanding protein folding/misfolding properties and controlling aggregation/solubility in biopharmaceuticals. Amylpred integrates predictions from five methods: Average Packing Density, Possible Conformational Switches, Amyloidogenic Pattern, TANGO, and Hexapeptide Conformational Energy. Amylpred webserver is accessible at http://aias.biol.uoa.gr/AMYLPRED/. The primary amino acid sequence of the query protein served as the input (Tsolis et al., 2013).

3.2.8 Sequence-based Detection of APRs in VP6

Three tools were utilized for the structure-based prediction of APRs in VP6, as outlined in the following section.

3.2.8.1. SolubiS

SolubiS is a comprehensive method developed to identify APRs characterized by high aggregation propensity and low thermodynamic stability using the protein's three-dimensional structure. This approach integrates the TANGO algorithm for predicting β-aggregation-prone regions and utilizes the FoldX empirical force field to assess

protein stability. SolubiS distinguishes between two types of APRs: structural APRs, which are protected from aggregation-triggering events by folding, and critical APRs, which can induce aggregation without major unfolding (van Durme et al., 2016). The three-dimensional structure of VP6 was submitted to SolubiS web-based tool (http://solubis.switchlab.org/) and stretch-plots representing APRs based on the TANGO summed score were generated.

3.2.8.2. Camsol structurally corrected

CamSol structurally corrected implements structural corrections on three-dimensional structure of protein to enhance intrinsic solubility and compute protein solubility based on amino acid proximity and solvent exposure within the structural context (Sormanni et al., 2015). The structural analysis of the VP6 PDB file served as an input for CamSol. Further details regarding CamSol can be accessed at http://www-vendruscolo.ch.cam.ac.uk/camsolmethod.html.

3.2.8.3. Aggrescan3D (A3D)

AGGRESCAN3D is a method designed to predict APRs within natively folded proteins by utilizing structural analysis in both static and dynamic states. This approach focuses on identifying surface-exposed APRs based on amino acid positioning and structural characteristics. During the static mode calculation of AGGRESCAN3D (A3D), the input protein structure is first minimized using the FoldX force field. Following this, the intrinsic aggregation propensity score of each amino acid is adjusted based on its structural context, resulting in a structurally corrected aggregation score for each residue. The 3-D structure of VP6 in PDB format served as the input for AGGRESCAN3D server (http://biocomp.chem.uw.edu.pl/A3D/), using a radius of 10 Å. All the APRs were subjected to amino acid composition analysis using PEPTIDE 2.0 webserver (https://www.peptide2.com/N_peptide_hydrophobicity_hydrophilicity.php).

3.3 Results

3.3.1 Cell Transformation and Optimization of VP6 Expression in *E. coli* BL21(DE3)

E. coli BL21(DE3) was successfully transformed with pET-28a(+)-RVAVP6 recombinant plasmid confirmed by single/double restriction digestion and PCR. VP6

was effectively overexpressed at various tested Isopropyl β-D-1-thiogalactopyranoside (IPTG) concentrations ranging from 0.0 mM to 1 M, with comparable band intensities. Hence, an IPTG concentration of 0.4 mM was considered for the subsequent VP6 protein expression. However, despite successful overexpression, the VP6 protein consistently appeared in the pellet fraction without evident solubilization (**Figure 3.1.**).

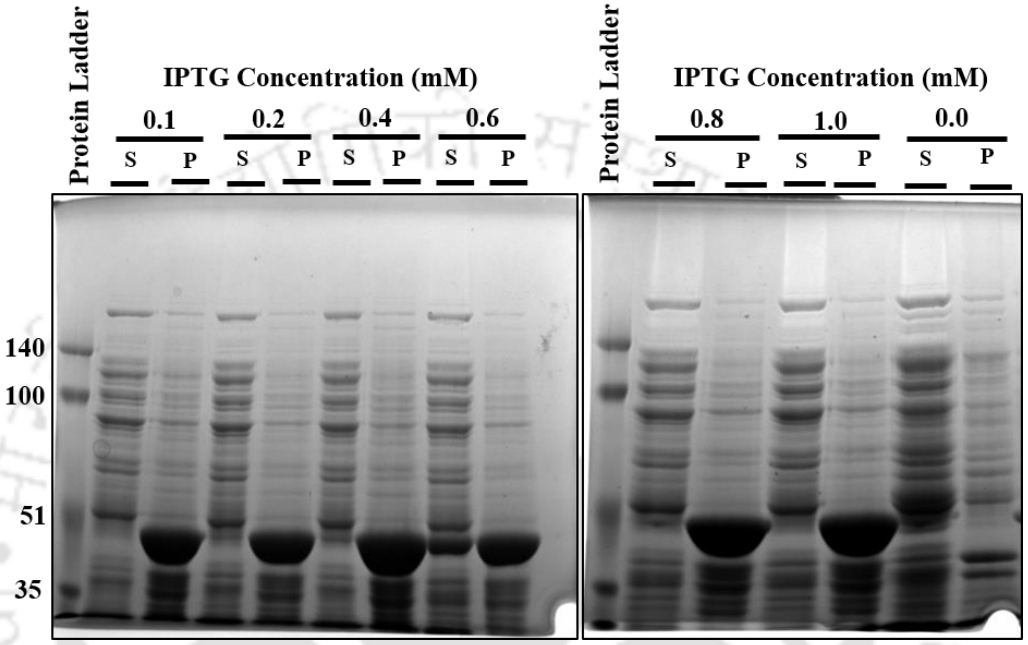


Figure 3.1. SDS-PAGE analysis of the impact of IPTG concentration (0.0 to 1.0 mM) on the expression of RVA VP6 in *E. coli* BL21(DE3). S represents the supernatant fraction, and P represents the pellet fraction.

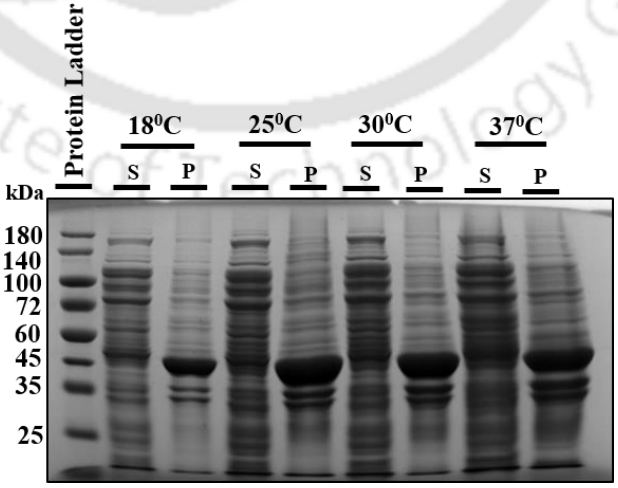


Figure 3.2. SDS-PAGE analysis showing the effect of post-induction temperatures (18°C, 25°C, 30°C, 37°C) on the expression of RVA VP6 in *E. coli* BL21(DE3). S represents the supernatant fraction and P represents the pellet fraction.

To determine the optimum post-induction temperature for overexpression of VP6, the bacterial cultures were subjected to post-induction growth at temperatures of 18 °C, 25 °C, 30 °C and 37 °C. The overexpression was at a higher extent from 25 °C to 37 °C, almost equally. Henceforth, overexpression of VP6 was carried out at 25 °C (**Figure 3.2**).

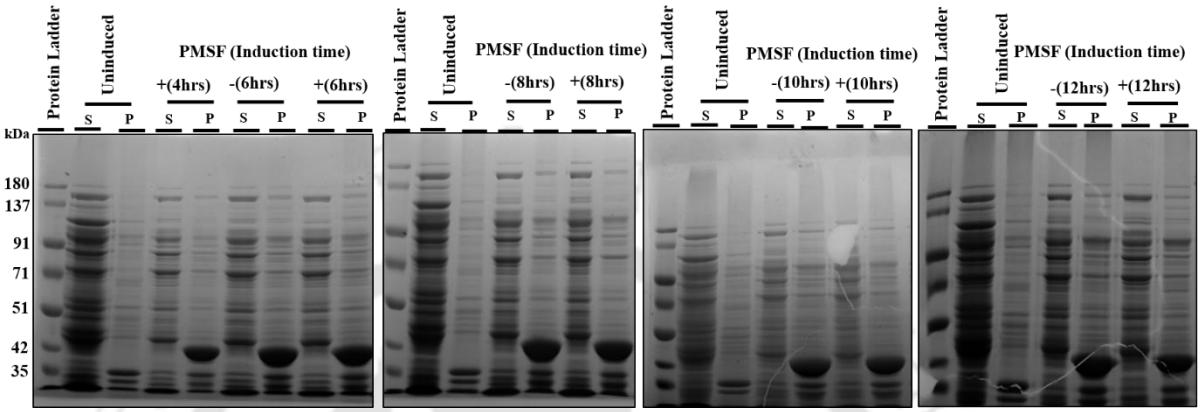


Figure 3.3. SDS-PAGE analysis illustrating the impact of post-induction time (4, 6, 8, 10, and 12 hours) with and without PMSF treatment (0.2 mM) on RVA VP6 expression in *E. coli* BL21(DE3) in supernatant (S) and pellet (P) fractions.

Post-induction incubation time was assessed to determine the optimal time for protein expression and PMSF, a serine protease inhibitor, was used to investigate possible protein degradation of overexpressed VP6 in the soluble fraction during protein extraction, in a time-dependent manner. The effect of post-induction incubation time on solubilizing the protein of interest was tested with and without PMSF treatment (0.2 mM). Protein extraction was carried out at 4 hours, 6 hours, 8 hours, 10 hours, and 12 hours post-induction to assess solubilization efficiency. The overexpression of VP6 protein increased from 4 to 6 hours post-induction, after which the amount of overexpressed protein remained stable until 12 hours of incubation. Interestingly, the presence or absence of PMSF treatment did not affect VP6 overexpression. However, under all conditions tested, the protein consistently appeared in the insoluble fraction (**Figure 3.3.**).

To achieve protein solubilization, urea treatment was tested with concentrations ranging from 1M to 8M in the lysis and resuspension buffers during protein extraction. Analysis of the urea-treated protein samples by SDS-PAGE revealed considerable partial solubilization of VP6 at urea concentrations of 6M to 8M (**Figure 3.4.**).

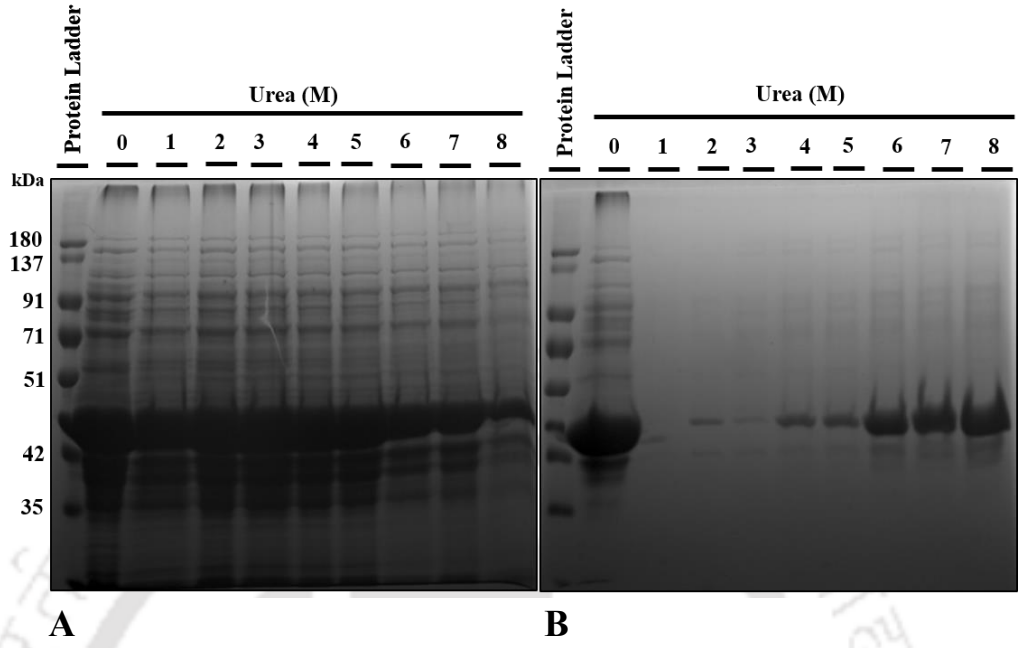


Figure 3.4. SDS-PAGE analysis of the effect of urea treatment on RVA VP6 expression in *E. coli* BL21(DE3), showing A. pellet fraction and B. supernatant fractions from crude cell lysates treated with varying concentrations of urea (0 to 8 M).

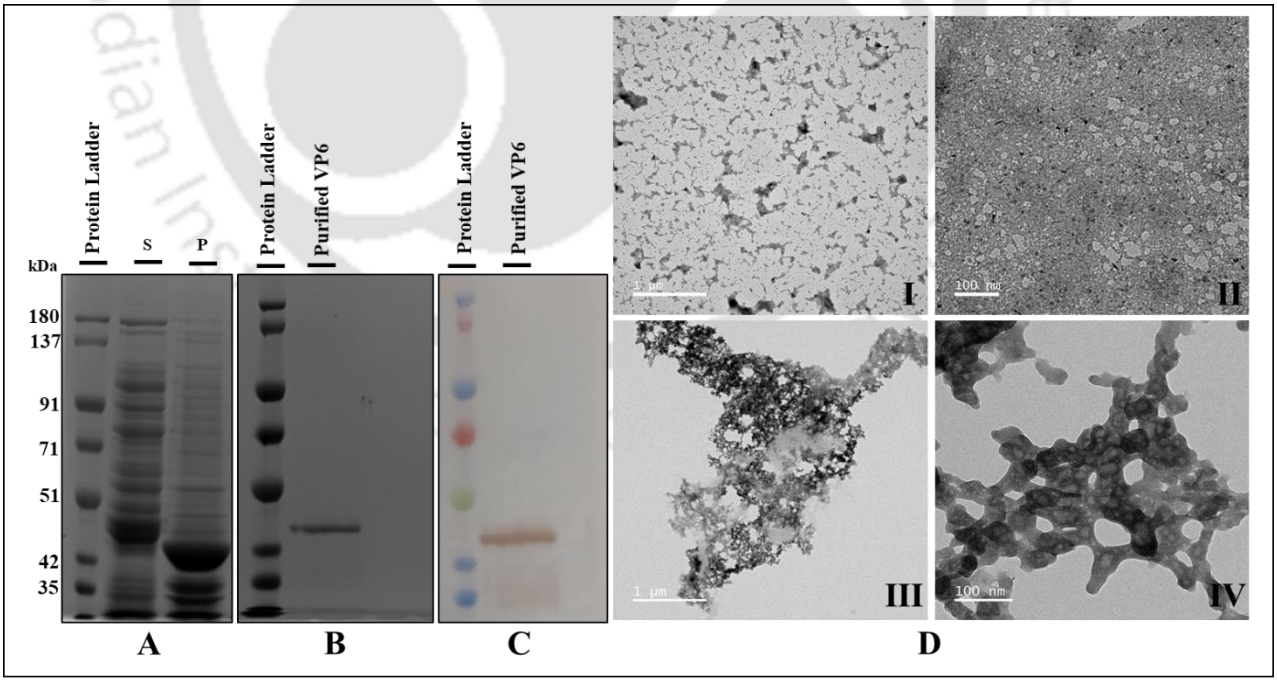


Figure 3.5. Characterization of recombinant RVA VP6 protein. A. SDS-PAGE analysis of crude cell lysates from induced E. coli BL21(DE3)/pET-28a(+)-RVAVP6. S represents the supernatant/soluble fraction while P represents pellet/insoluble fraction. B. SDS-PAGE and corresponding C. western blot analysis of purified RVA VP6 with

VP6 specific monoclonal antibody. D. FETEM images depict VP6 inclusion bodies at magnifications of 1 µm (I) and 100 nm (II), alongside purified VP6 at magnifications of 1 µm (III) and 100 nm (IV), showing amorphous aggregates.

3.3.2 VP6 Purification and Western Blot

The urea-solubilized, overexpressed VP6 was subjected to Ni-NTA column chromatography and coupled with on-column renaturation. The purity of the protein sample eluate was analysed by SDS-PAGE in comparison to the crude lysate (**Figure 3.5.A and B**). The eluted sample containing purified VP6 was subjected to western blot, and a dark brown coloured band was observed at ~45 kDa, corresponding to the molecular weight of VP6 (**Figure 3.5.C**).

3.3.3 Field Emission Transmission Electron Microscopy (FETEM) of VP6 Inclusion Bodies and Purified VP6

FETEM analysis was conducted to assess the structural characteristics of purified VP6 samples (**Figure 3.5.D**). Intriguingly, the FETEM images revealed the presence of amorphous aggregates within the purified VP6 samples from prokaryotic expression system. These aggregates exhibited a lack of defined structure or crystallinity, appearing as irregularly shaped accumulations dispersed throughout the sample. The observation of amorphous aggregates suggests a propensity for VP6 to undergo non-specific interactions and self-association, which may have implications for its stability, functionality, and downstream applications. Such aggregation behaviour suggests potential challenges in the downstream processing and formulation of VP6-based products, necessitating further investigation into the factors influencing VP6 aggregation and strategies for mitigating this phenomenon.

3.3.4 ThT Binding Assay

The examination of VP6 inclusion bodies through ThT assay, at an excitation wavelength of 440 nm, unveiled a peak emission at approximately 487 nm, indicative of amyloidogenic aggregates. Intriguingly, fluorescence intensity analysis revealed a temperature-dependent modulation of amyloid character in VP6 inclusion body formation (**Figure 3.6.A and B**). The plot depicts a linear trend in ThT fluorescence intensity in VP6 inclusion bodies with increasing incubation temperature, suggesting a systematic relationship between incubation temperature and amyloid characteristic of

VP6 inclusion bodies. This finding underscores the thermally sensitive nature of amyloid formation in VP6 inclusion bodies, aligning with the established phenomenon that higher temperatures enhance amyloid aggregation propensity (Chiti & Dobson, 2006). This understanding can contribute to a more comprehensive characterization of the conformational dynamics and thermal stability of VP6 aggregates following expression at different temperatures.

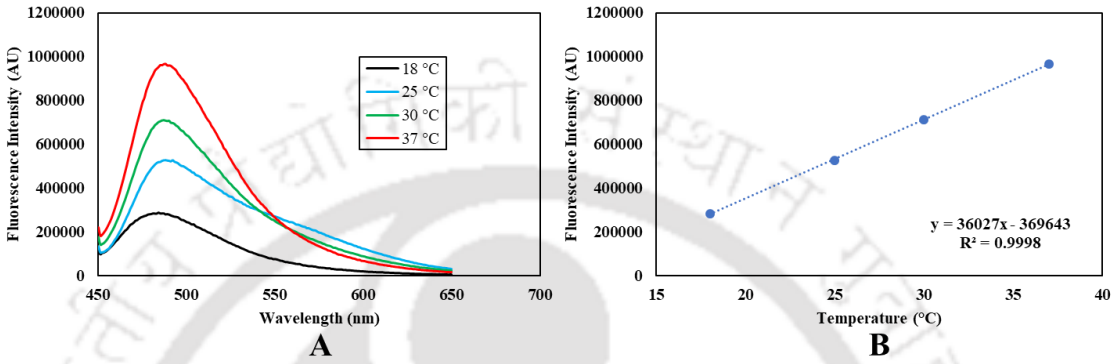


Figure 3.6. Aggregation profiling of VP6 inclusion bodies expressed at different temperatures using ThT binding assay. A. Fluorescence emission spectra of ThT bound to VP6 inclusion bodies (excitation: λ_{440nm}). B. Temperature-dependent fluorescence response of ThT in VP6 inclusion bodies (excitation: 440 nm, emission: 487 nm).

3.3.5 Physicochemical Characterization and Structure Prediction of Rotavirus A VP6 Protein

The retrieved VP6 sequence was submitted in FASTA format to ProtParam for physicochemical characterization. The VP6 protein is 397 amino acids in length with molecular weight of 44.873 kDa. and p*I* of 5.81. Instability index had been predicted to be 40.27, suggesting that the protein may be unstable. This indicates the protein's susceptibility to structural instability. This metric assesses the likelihood of the protein undergoing denaturation or loss of structure under physiological conditions. A higher instability index implies increased vulnerability to unfolding or aggregation, potentially affecting its functional integrity and experimental outcomes. Therefore, this value suggests that VP6 may exhibit characteristics predisposing it to instability, a factor important for understanding its behavior and applications in biological studies. The aliphatic index of a protein is defined as the relative volume occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine). It may be regarded as a positive factor for the increase of thermostability of globular proteins. The aliphatic index of VP6 were

found to be 88.66. The aliphatic index, indicative of the proportion of aliphatic side chains in the protein sequence, highlights its influence on stability, hydrophobicity, and thermal stability. With an aliphatic index of 88.66, VP6 exhibits a notable presence of hydrophobic residues, which can promote aggregation when exposed on the protein surface. Conversely, buried aliphatic residues contribute to stability but may hinder solubility if essential for interactions with aqueous environments. These characteristics underscore how aliphatic amino acids, crucial for stability and function, significantly impact the solubility profile of VP6. Engineering strategies often focus on modifying surface-exposed aliphatic residues to enhance solubility in aqueous solutions, balancing the need for stability with the challenges posed by hydrophobic interactions and potential aggregation pathways. GRAVY value of VP6 were predicted to be -0.144. The GRAVY number of a protein is a measure of its hydrophobicity or hydrophilicity. A negative GRAVY value indicates that the protein is hydrophilic. The negative GRAVY index of VP6, indicating hydrophilicity, contrasts with the presence of numerous exposed hydrophobic patches identified in our study. This discrepancy is intriguing as it suggests a potential complexity in VP6's structural and functional roles. While hydrophobic patches typically correlate with aggregation propensity, the overall hydrophilic nature indicated by GRAVY may imply a specific mechanism of interaction or folding that counters the aggregative tendency of these patches. Understanding this dichotomy can be crucial for elucidating VP6's behaviour in both native and recombinant contexts, influencing strategies for protein engineering and therapeutic development. Its half-life is estimated to be 30 hours (mammalian reticulocytes, in vitro), >20 hours (yeast, in vivo) and >10 hours (*E. coli*, in vivo).

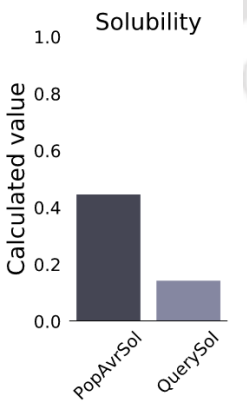


Figure 3.7. RVA VP6 protein solubility prediction using Protein-Sol.

The solubility of VP6 in the *E. coli* expression system was predicted by the Protein-Sol server. The population average for the experimental dataset (PopAvrSol) for *E. coli* expressed proteins is 0.45. Therefore, a scaled solubility value above 0.45 suggests higher predicted solubility than the average soluble *E. coli* protein in the dataset, whereas values below 0.45 indicate lower predicted solubility. Notably, the predicted solubility of VP6 in the *E. coli* expression system, assessed by the Protein-Sol server, was lower compared to other experimentally expressed proteins (**Figure 3.7.**).

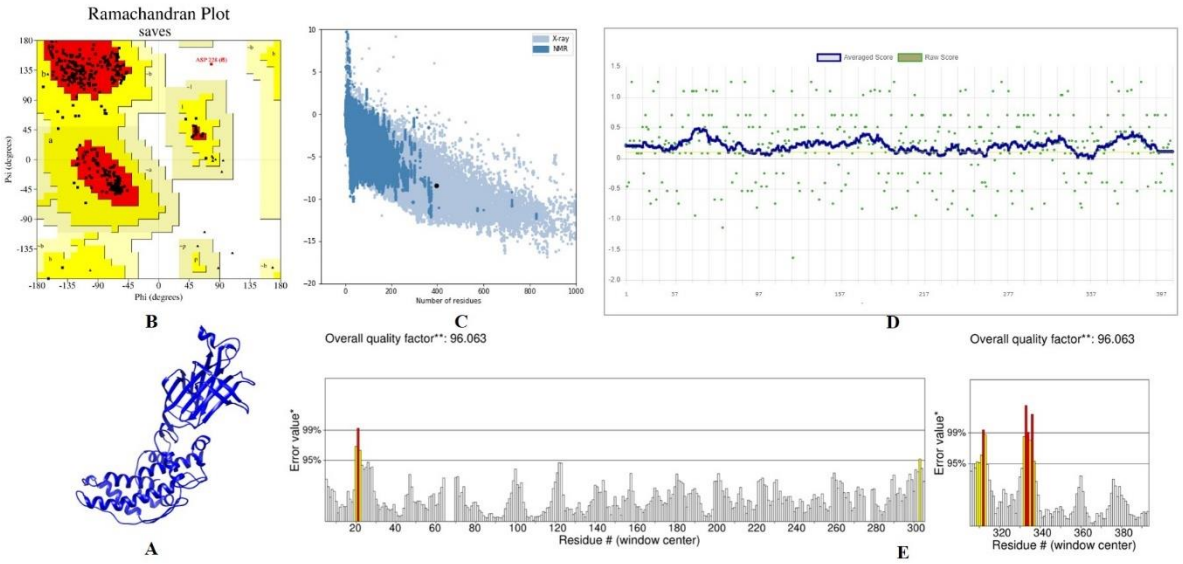


Figure 3.8. Structure Validation of RVA VP6. A. Predicted model of Human RVA VP6. B. Ramachandran Plot. C. ProSA with overall Z-score of -8.42. D. Verify 3D with 87.91% of the residues averaging 3D-1D score >= 0.1. (e) ERRAT score 96.063.

While there is no available complete experimental structural data for RVA VP6 sequence used in the current investigation, experimental structural data for bovine RVA VP6 is available (PDB ID: 1QHD), which exhibit 96.98% amino acid sequence identity to the VP6 sequence used in the current study. The experimentally determined structure of bovine RVA VP6 reveals the following secondary structure composition: alpha helix 27.7%, 3-10 helix 4.5%, pi helix 0.0%, extended strand 31.0%, isolated beta bridge 0.5%, turn 9.3%, bend 7.3%, and coil 19.6%. Given the high sequence identity, it is plausible that the human RVA VP6 sequence exhibits a similar secondary structure composition.

The VP6 three-dimensional structure predicted by SWISS-MODEL underwent structural validation. Procheck analysis showed that 92.4% of residues were in the most favoured regions (A, B, L) according to the Ramachandran Plot. ProSA indicated an

overall Z-score of -8.42, suggesting a structure within an acceptable range of native-like conformations. The protein structure analysis by Verify 3D demonstrated that 87.91% of residues achieved a 3D-1D score >= 0.1, exceeding the 80% threshold. Verify 3D scores assess the compatibility of a protein's 3D structure with its amino acid sequence, indicating higher scores for residues in environments typical of native proteins.

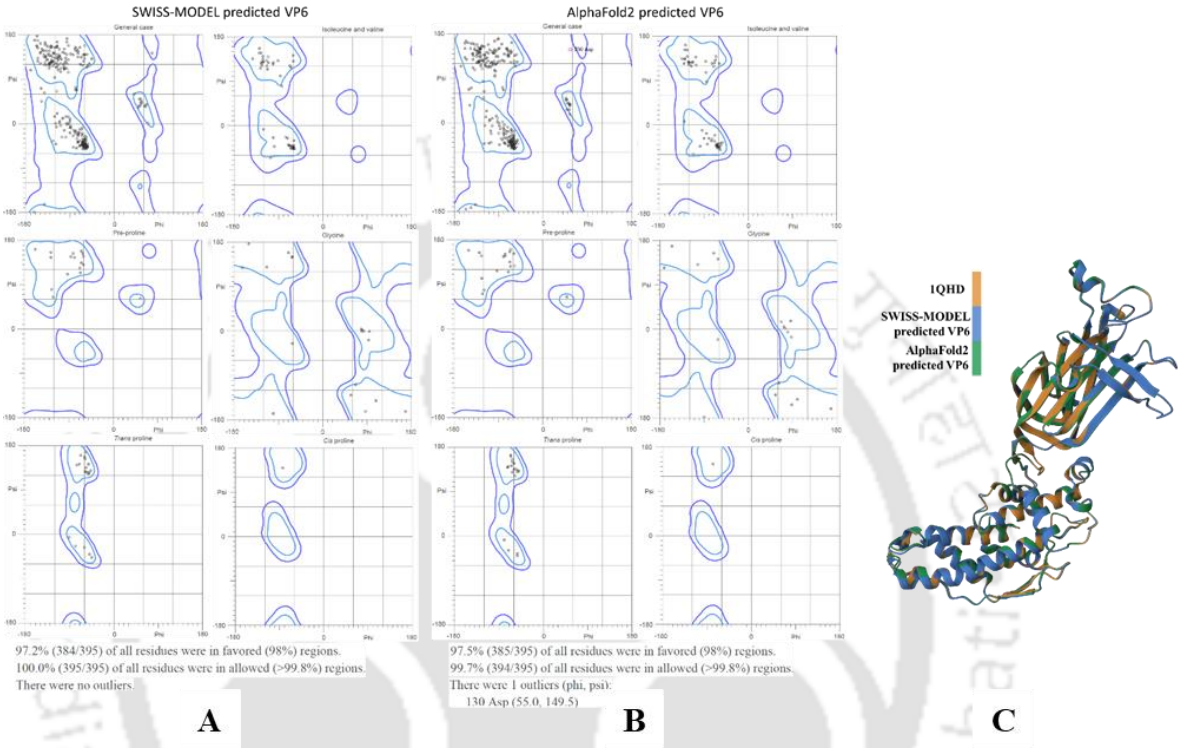


Figure 3.9. Comparison of VP6 predicted structures. Ramachandran plot of A. SWISS-MODEL and B. AlphaFold2 predicted VP6 3D structures obtained by Molprobity webserver. C. Pairwise structure alignment of SWISS-MODEL and Alphafold2 VP6 predicted structures with the reference VP6 structure with PDB ID: 1QHD using RCSB PDB structure alignment.

Additionally, the ERRAT score was 96.063, surpassing the 95% threshold (**Figure 3.8.**). ERRAT evaluates overall structure quality by analyzing non-bonded interactions. The X-axis of the ERRAT graph represents the amino acid positions in the protein sequence. The Y-axis represents the error values, which are indicative of the quality of the non-bonded interactions for each segment of the protein. Lower error values suggest better quality and more reliable interactions. These structural validations collectively support the high quality and reliability of this predicted model of VP6. Further, to compare the quality of the SWISS-MODEL and AlphaFold2 predicted structures, we analysed their Ramachandran plots using the MolProbity server (**Figure 3.9.A and B**).

Additionally, we performed pairwise structure alignment using the RCSB PDB alignment tool to assess the structural congruence with the reference structure (PDB ID: 1QHD) (**Figure 3.9.C**). The SWISS-MODEL predicted VP6 structure exhibited 97.2% (384/395) of all residues in favoured regions, 100.0% (395/395) in allowed regions, and no outliers. The AlphaFold2 predicted VP6 structure showed 97.5% (385/395) of all residues in favoured regions, 99.7% (394/395) in allowed regions, with one outlier (phi, psi): 130 Asp (55.0, 149.5). Pairwise structural alignment of models predicted by SWISS-MODEL and AlphaFold2 with the reference structure (PDB ID 1QHD) showed RMSD (root mean square deviation) values of 0.08 and 0.35, respectively, and a TM-score of 1 for both predicted structures. RMSD measures the alignment of backbone C-alpha atoms between superposed structures in Å, with lower values indicating better alignment. The TM-score (template modeling score) assesses topological similarity between template and model structures, ranging from 0 to 1, where scores above 0.5 typically indicate a similar protein fold. These results indicate that both predicted models are comparable and of reliable quality. The SWISS-MODEL predicted VP6 structure was used for further analysis.

3.3.6 Sequence-based and Structure-Based Detection of APRs in VP6

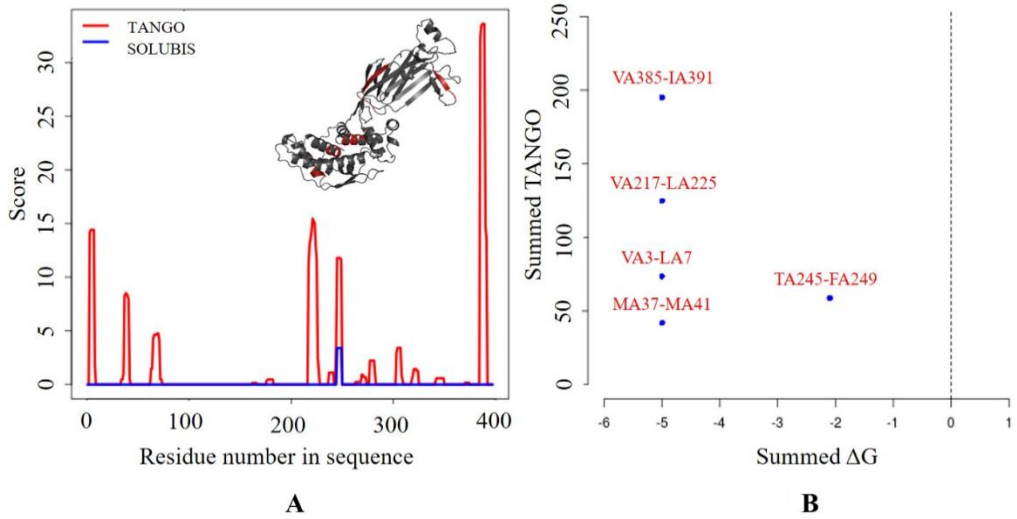


Figure 3.10. Prediction of APRs in VP6. A. Schematic representation of APRs predicted by TANGO in VP6. TANGO aggregation propensity and SolubiS score were plotted against the protein sequence. B. Stretch-plot showing APR aggregation propensity and thermodynamic stability in VP6. Positive summed ΔG indicates surface-exposed APRs, negative SolubiS score denotes structural, buried APRs.

Table 3.1. Summary of predicted APRs with amino acid residue positions and sequences in VP6 protein.

TOOLS APR REGION	APR SEQUENCE	TOOLS APR REGION	APR SEQUENCE	TOOLS APR REGION	APR SEQUENCE
TANGO		FOLDAMYLOID		CAMSOL INTRINSIC	
3 to 7	VLYSL	1 to 6	MDVLYS	1 to 4	MDVL
37 to 41	MIITM	35 to 40	NQMIIT	21	G
217 to 225	VLTTATITL	55 to 65	PIRNWNFNFGL	25	S
245 to 249	TTWFF	67 to 72	GTTLLN	33	Q
385 to 391	VFTVASI	85 to 91	IDYFVDF	36 to 40	QMIIT
AGGRESCAN		93 to 99	DNVCMDE	61 to 64	FNFG
1 to 9	MDVLYSLSK	122 to 129	IKFKRINF	66 to 68	LGT
32 to 39	QQFNQMII	137 to 141	ENWNL	86 to 90	DYFVD
58 to 75	NWNFNFGLLGTTLLNLDA	178 to 183	GTMWLN	94	N
86 to 94	DYFVDFVDN	212 to 219	VPLRRVLT	138	N
119 to 125	LSAIKFK	232 to 236	FSFPR	159 to 164	PYSASF
160 to 164	YSASF	246 to 255	TWFFNPVILR	178 to 182	GTMWL
179 to 186	TMWLNAGS	261 to 265	VEFLL	196	S
190 to 198	VAGFDYSCA	286 to 295	DTIRLSFQLM	198	A
214 to 225	LRRVLTTATITL	305 to 309	AVLFP	211	I
242 to 250	DGATTWFFN	322 to 326	LTLRI	220 to 224	TATIT
257 to 270	NNVEVEFLLNGQII	366 to 370	NWTDL	235	P
272 to 278	TYQARFG	383 to 390	QRVFTVAS	246 to 251	TWFFNP
280 to 287	IVARNFDT	392 to 397	RSMLIK	262	E
289 to 293	RLSFQ	AMYLPRED		266 to 267	NG
303 to 308	AVAVLF	3 to 7	VLYSL	270 to 272	INT
321 to 335	GLTLRIESAVCESVL	35 to 40	NQMIIT	278 to 280	GTI
343 to 348	LANVTS	58 to 65	NWTFDFG.	282	A
350 to 355	RQEYAI	67 to 68	GT	288 to 293	IRLSFQ
385 to 397	VFTVASIRSMLIK	85 to 93	IEYFIDFID	303 to 309	AVAVLFP
ZYGGREGATOR		149 to 153	GFVFH	319 to 323	TVGLT
11 to 15	LKDAR	178 to 185	GTMWLNAG	332 to 333	ES
100 to 101	MV	188 to 189	IQ	344 to 346	ANV
104 to 106	SQR	210 to 214	HIVQL	357 to 358	VG
114 to 118	DSLRK	245 to 253	TTWFFNPII	385 to 389	VFTVA
143 to 147	NRRQR	260 to 266	EVEFLLN	SOLUBIS	
170 to 171	QP	278 to 283	GTIVAR	245-249	TTWFF
376 to 381	PSREDN	290 to 295	LSFQLM	CAMSOL STRUCTURALLY CORRECTED	
18, 228, 230, 299	I, D, E, N	322 to 326	LTLRI	39, 163, 220, 235, 246, 248, 279	I, S, T, P, T, P, F
WALTZ		342 to 348	LLANVTA	AGGRESCAN3D	
1 to 7	MDVLYSL	367 to 373	WTDLITN	23 to 24	LY
35 to 41	NQMIITM	383 to 394	QRVFTVASIRSM	55 to 56	PI
57 to 67	RNWNFNFGLLG	ANUPP		64 to 65	GL
73 to 92	LDANYVETARNTIDYFVDFV	18-43	IVEGTLYSNVSDLIQQFNQMIITMN	68 to 71	TTLL
135 to 141	YIENWNL	84-97	TIEYFIDFIDNVC	121 to 122	GI
207 to 212	QFEHIV	148-167	TGFVFHKPNIFPYSASFTL	159 to 164	PYSASF
245 to 252	TTWFFNPV	173-215	HDNLMGTMWLNAGSEIQVAGFDY SCALNAPANIQQFEHIVQL	181 to 184	WLNA
266 to 289	NGQIINTYQARFGTIVARNFDTIR	244-273	ATTWFFNPIILRPNNVEVEFLLNGQI INT	234 to 239	FPRVIT
304 to 309	VAVLFP	301-339	TPAVNALFPQAQPFQHHATVGLTL RIESAVCESVLADA	252 to 253	VI
370 to 374	LITNY	352-375	EYAIPVGPVFPPGMNWTDLITNYS	39, 117, 211, 226, 246, 248, 250, 278, 281	I, I, I, L, T, Y, N, G, I
				303 to 307	AVAAL
				356 to 358	PVG

TANGO utilizes a statistical algorithm that merges protein sequence information with pertinent physicochemical parameters. This combined methodology plays a pivotal role in discerning APRs within proteins. TANGO analysis identified five segments within VP6 as potential APRs, characterized by aggregation scores exceeding the predefined threshold (5%). These segments, as shown in **Table 3.1** are indicative of their potential

involvement in cross-β aggregation (**Figure 3.10.A**). Notably, the peptide VFTVASI exhibited the highest TANGO β-aggregation score, while the MIITM segment demonstrated the lowest score. To discern between buried and exposed APRs as predicted by TANGO, VP6 was analysed using the SolubiS server. The resulting stretch-plot, generated by SolubiS, depicted the aggregation propensity and local stability of APRs within VP6. Analysis revealed that none of the five APRs identified by TANGO exhibited a positive summed ΔG, indicating the absence of critical, surface-exposed APRs mediating native state aggregation. However, the segment spanning residues 245-249 (TTWFF), characterized by a least negative SolubiS score and therefore its low contribution to thermodynamic stability, was identified as a structural, buried APR implicated in non-native aggregation (**Figure 3.10.B**).

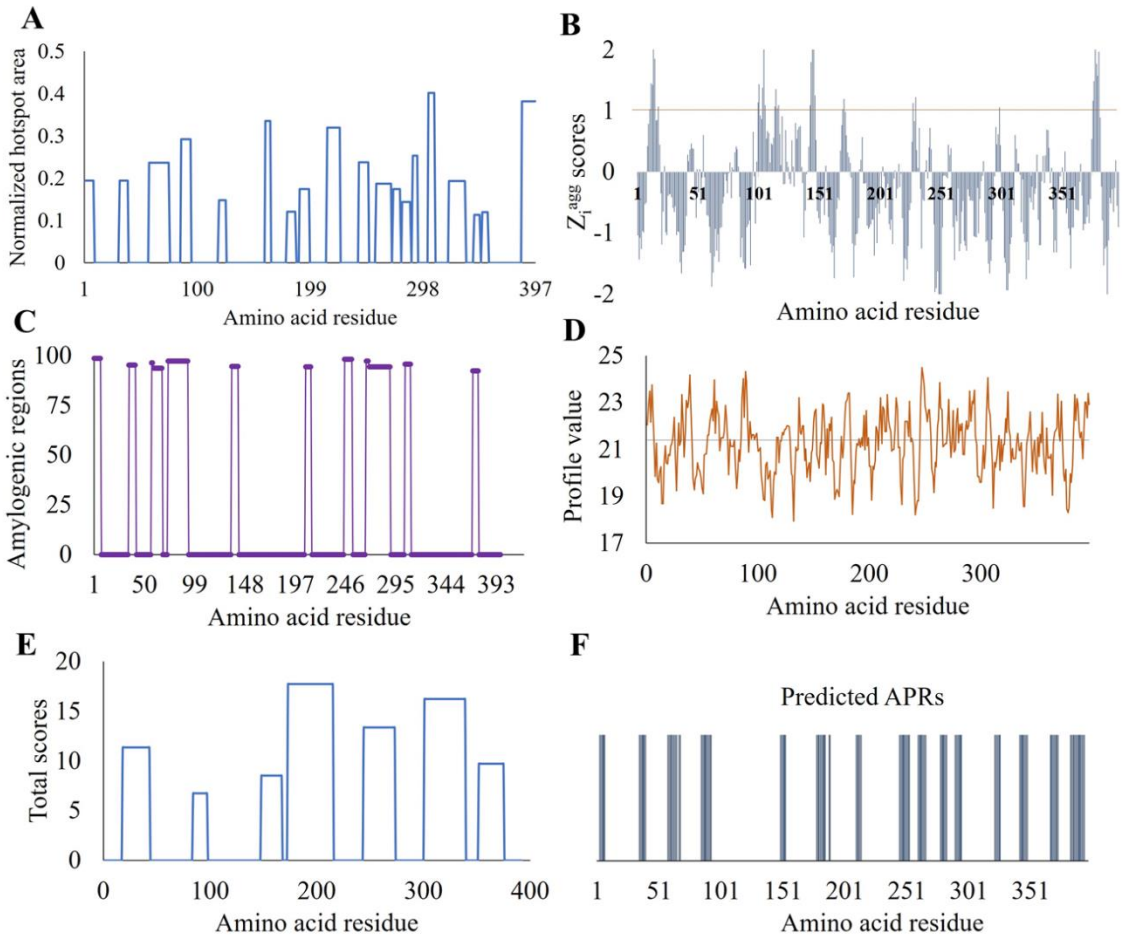


Figure 3.11. Prediction of APRs in VP6. A. AGGRESCAN normalized hot spot area plot. B. Zyggregator scores (Z_i^{agg}). The brown line represents the aggregation tendency threshold. Residues above this threshold are predicted to contribute to fibrillar aggregate formation. C. Amyloidogenic regions identified using WALTZ. D. The amyloidogenic regions identified using FoldAmyloid. E. ANuPP showing predicted total aggregation

scores of identified amyloidogenic regions. F. Amyloidogenic regions identified by AmylPred.

AGGRESCAN predicts APRs by identifying primary sequences of query protein with protein fragments experimentally linked to disease-associated protein aggregation. According to the AGGRESCAN hot spot area data (**Figure 3.11. A**) prepared for the VP6, nineteen aggregation hotspots have been recognized in the VP6 protein (**Table 3.1**).

Zyggregator considers physicochemical conditions like pH, temperature, ionic strength, and trifluoroethanol concentration on aggregation. It predicts local instabilities using the CamP program. Zyggregator also assesses "gatekeeper" residues flanking a sliding window for charged residues that could mitigate aggregation. Zyggregator identified 32 residues (**Table 3.1**) with the Z_i^{agg} score greater than the threshold value (**Figure 3.11.B**).

Analysis using Waltz algorithm revealed ten distinct regions as APRs within the VP6 sequence. These APRs, spanning residues as shown in **Table 3.1**, were identified as potential sites for amyloid formation (**Figure 3.11.C**). Prediction of APRs using FoldAmyloid revealed nineteen aggregation hotspots regions (**Table 1**) (**Figure 3.11.D**). ANuPP, utilizing atomic-level features derived from hexapeptides, predicted 7 APRs as shown in **Table 1**(**Figure 3.11.E**). AmylPred server prediction relies on the concurrence of at least two out of five methods (refer section 3.2.7.8). In VP6, Amylpred identified 17 consensus aggregation hits as shown in **Table 3.1** (**Figure 3.11.F**).

The intrinsic solubility profile of VP6, assessed by CamSol, identified certain residues as poorly soluble (aggregation-prone) (**Table 3.1**) (**Figure 3.12.**). According to the Camsol structurally-corrected solubility profile, the residues I 39, S 163, T 220, P 235, T246, F 248, and T 279 have structurally corrected solubility scores (**Figure 3.12.**). These are the solvent-exposed poorly soluble residues of VP6. These residues represent potential positions for mutations aimed at enhancing solubility. The structurally-corrected solubility scores are derived from the intrinsic solubility scores, considering the proximity of the amino acids in the 3D structure and their solvent exposure. The Aggrescan3D (A3D) profile of VP6 provides a comprehensive assessment of residues implicated in aggregation propensity, utilizing spatial conformation and intermolecular

interaction potentials within the protein structure. Notably, residues identified are shown in **Table 3.1.** Residues 39, 65, 70, 71, 160, 234, 237, 238, 248, 252, 253, and 357 demonstrate significant aggregation propensity as delineated by the A3D analysis (**Figure 3.13**). A compiled information on all the predicted APRs have been provided in **Table 3.1**.

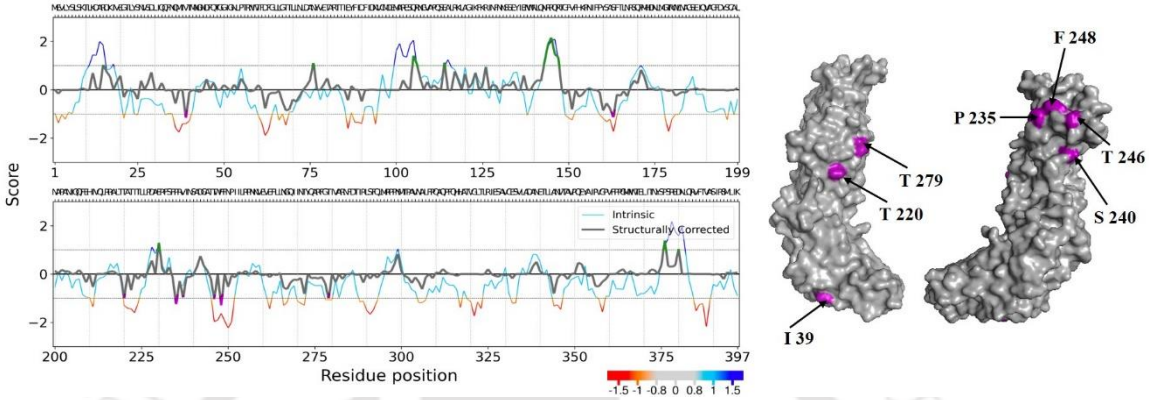


Figure 3.12. Camsol intrinsic and Camsol structurally corrected solubility profiles predicted for VP6. The solvent-exposed poorly soluble amino acids in VP6 predicted by Camsol structurally corrected are shown with the black arrows and marked with the purple colour.

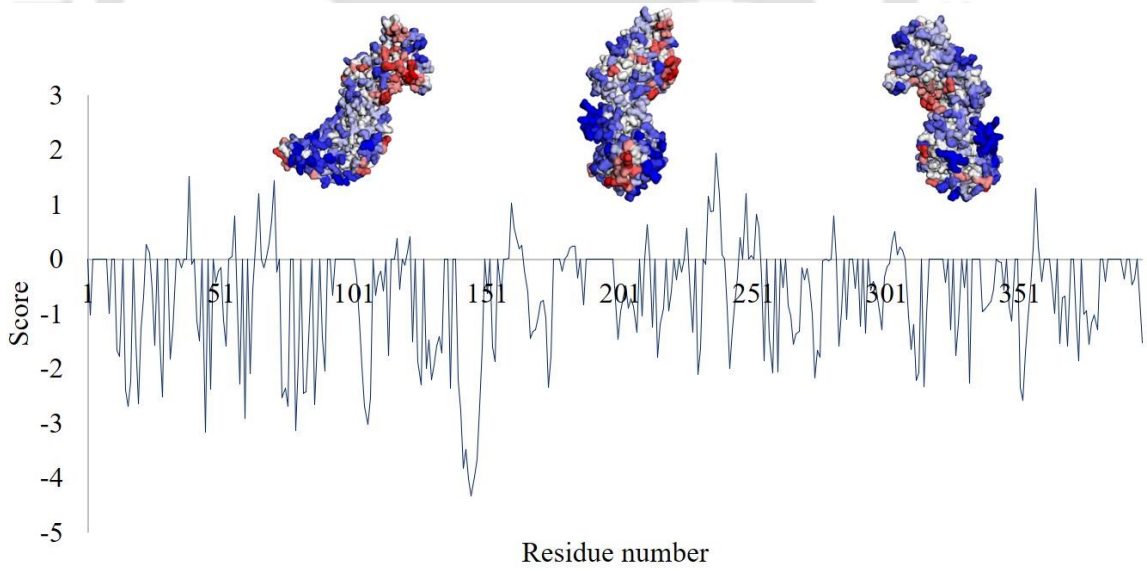


Figure 3.13. A3D structure-based analysis of VP6 aggregation propensity. Positive scores denote aggregation-prone residues, while negative scores indicate solubility-prone residues. The protein surface is color-coded based on the A3D score gradient: red for high aggregation propensity, white for negligible effect, and blue for high solubility regions.

A comprehensive investigation utilizing both sequence-based and structure-based approaches revealed a total of 24 consensus APRs within VP6 (**Figure 3.14.A and B**).

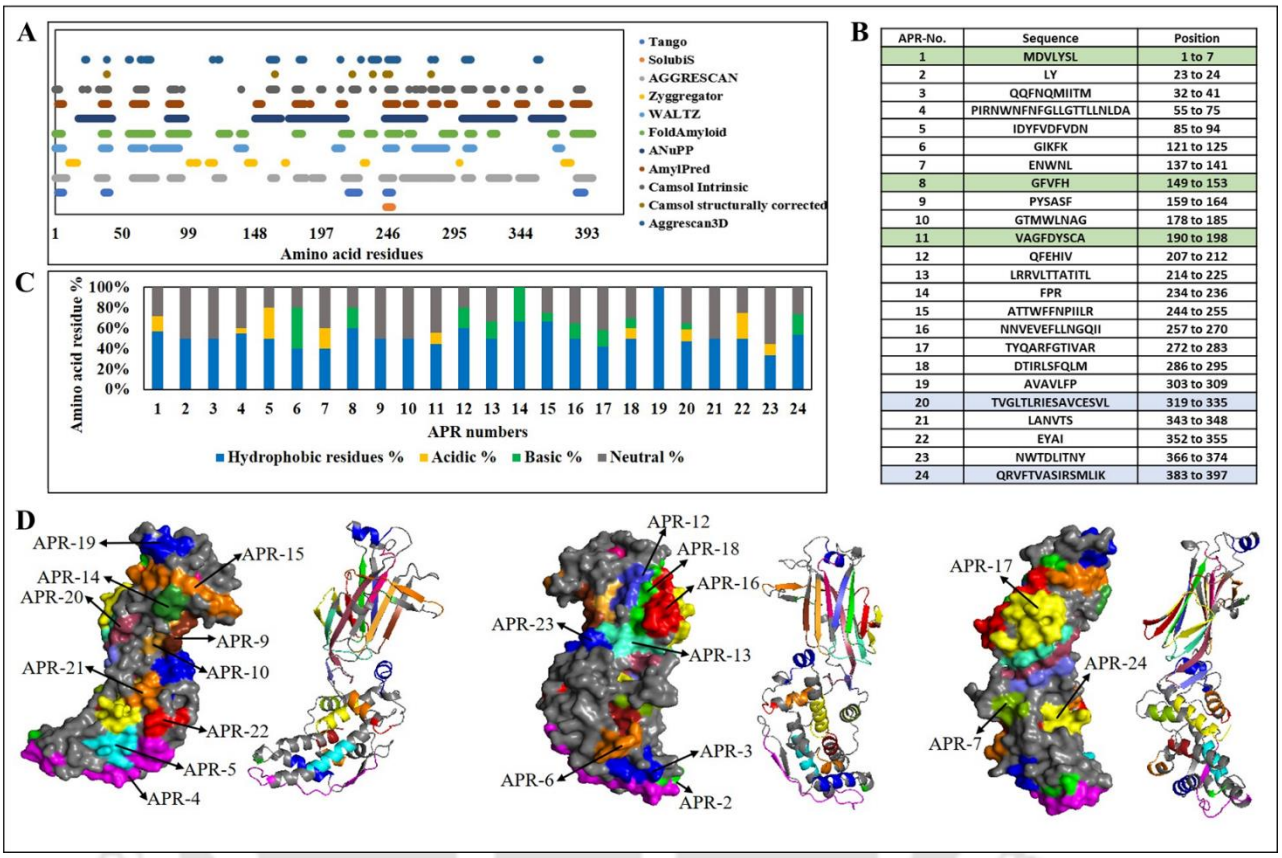


Figure 3.14. A comprehensive summary of all APRs identified by different tools in the RVA VP6 protein. A. Summary of consensus APRs within RVA VP6, as predicted by various sequence-based and structure-based APR predictors. Each predictor's results are indicated by distinct colours, facilitating comparison. B. Table displaying consensus APR sequences predicted by two or more tools and their positions in the RVA VP6 protein. Entries in white indicate totally surface-exposed APRs, those highlighted in green denote totally buried APRs, and light blue entries signify partially surface-exposed APRs. C. Amino Acid Composition of APRs. D. RVA VP6 structure is shown in both space-filling and cartoon models. Consensus APRs are highlighted in different colours, with their assigned APR number corresponding to their APR numbers from the table in B.

These regions, defined as areas consistently predicted by two or more computational tools, offer insights into potential aggregation tendencies of the protein. Six APRs (1,

2, 3, 5, 7, 19, 21) were found to be part of helical structures, while five APRs (4, 6, 22, 23, 24) form a combination of random coil and helix. Additionally, five APRs (8, 9, 10, 12, 18) are located within beta sheet structures, and eight APRs (9, 11, 13, 14, 15, 16, 17, 20) are positioned in regions with a combination of beta sheet and random coil structures in the native VP6 3-D structure. Amino acid composition of APRs exhibited hydrophobic amino acid residues in the range of 33.33 to 100 %, acidic amino acid residues from 0 to 30 %, basic amino acid residues from 0 to 40% and neutral amino acid residues from 0 to 55.56% showing clear predominance of hydrophobic amino acid residues (**Figure 3.14.C**). The predominance of hydrophobic amino acid residues in the APRs suggests that these regions are likely to contribute to protein aggregation and reduced solubility. Hydrophobic interactions are known to drive protein aggregation, leading to the formation of insoluble aggregates such as inclusion bodies. Therefore, the high percentage of hydrophobic residues in the APRs indicates a propensity for reduced solubility of the protein. Notably, a substantial majority of these APRs, comprising 21 out of 24, were mapped to be surface-exposed in VP6 3-D structure (**Figure 3.14.D**).

3.4 Discussion

In this investigation, we evaluated the heterologous prokaryotic expression of codon-optimized RVA VP6 in *E. coli* BL21(DE3) cells. Subsequent protein extraction and solubility analysis revealed an abundant presence of the recombinant protein within the insoluble fraction, contrasting with the preferred localization within the soluble fraction for efficient downstream processing, including protein purification, formulation, characterization, and functional studies. This phenomenon underscores a common challenge associated with high-level heterologous protein expression in bacterial systems, where the propensity for protein inclusion body formation presents a substantial impediment to protein research endeavours (Ventura & Villaverde, 2006). While protein recovery from inclusion bodies is indeed possible, the process entails labour-intensive and time-consuming procedures, resulting in inevitable protein loss at each stage. Additionally, the absence of a universal technique emphasizes the necessity for a trial-and-error approach to identify optimal strategies, thereby highlighting the intricate nature of protein handling in such instances. Therefore, to optimize soluble protein expression, a series of experimental conditions were systematically investigated

for VP6 expression in the prokaryotic host. This included varying the concentration of the inducer (IPTG), post-induction temperature, post-induction incubation time with and without PMSF. These efforts were undertaken with the objective of enhancing protein solubility.

The inducer concentration profoundly influences the soluble protein fraction in *E. coli* expression systems. Higher concentrations typically yield elevated protein expression levels, potentially leading to a greater proportion of protein. Excessively high inducer concentrations can also lead to cellular stress or saturation of protein folding machinery, potentially causing misfolding, aggregation, and a decrease in solubility. Conversely, lower inducer concentrations may result in reduced protein expression but could enhance solubility by allowing for more controlled folding. However, excessively low inducer concentrations may not provide sufficient stimulation for protein expression, leading to inadequate levels of soluble protein. Thus, an optimal inducer concentration is crucial, striking a balance between robust protein expression and efficient folding to maximize soluble protein fraction while minimizing aggregation and cellular stress (Sambrook J, 2001). In the present study, the effect of varying IPTG concentrations on VP6 expression was examined, revealing localization in the insoluble fraction across all tested IPTG concentrations (**Figure 3.1**).

Post-induction temperature in *E. coli* expression systems modulates protein solubility via its effects on protein folding kinetics, chaperone functionality, and potential protein denaturation. Optimal temperatures facilitate correct folding and enhanced solubility, while deviations from this range can induce protein misfolding or aggregation. Generally, protein solubility and activity are reported to be enhanced when expressed at lower temperatures in *E. coli* (Sahdev & Khattar, 2008). While VP6 expression level showed a slight increase with rising temperature (25°C, 30°C, and 37°C) (**Figure 3.2**).

Incubation times of 4, 6, 8, 10, and 12 hours were assessed both with and without PMSF (0.2 mM) to investigate potential degradation during protein extraction of VP6 expressed in the soluble fraction. The results revealed that VP6 was not detectable in the soluble fraction under any of the assessed incubation times, regardless of the presence of PMSF (**Figure 3.3**).

As a final approach, solubilization employing urea was undertaken, demonstrating efficacy solely at higher concentrations ranging from 6 to 8 M for the solubilization of

VP6 (**Figure 3.4**). However, given urea's protein-denaturing properties, a subsequent renaturation step was implemented utilizing on-column renaturation coupled with protein purification via Ni-NTA column chromatography. Subsequent visualization of the renatured protein via FETEM unveiled the formation of amorphous aggregates (**Figure 3.5.**). In contrast, the native VP6 protein demonstrates self-assembly characterized by structural polymorphism, resulting in the formation of trimers, hexagonal structures, and diverse architectures such as nanotubes, nanospheres, or nanosheets (Estes, 2007; Estes et al., 1987; Bugli et al., 2014; Choi et al., 1999; Kapikian & Hoshino, 2001).

Additionally, inclusion bodies generated at different temperatures underwent ThT binding assays. An augmentation in ThT fluorescence emission was observed corresponding to increasing incubation temperature, providing additional confirmation of aggregate formation within the inclusion bodies (**Figure 3.6.**). The analysis of ThT binding to VP6 inclusion bodies at varying incubation temperatures offers insights into the temperature sensitivity of protein aggregation processes associated with VP6 expression. An increase in ThT binding to VP6 inclusion bodies with incubation temperature suggests enhanced formation or stabilization of VP6 aggregates at higher incubation temperatures, indicating a temperature-dependent and potentially accelerated aggregation process. Conversely, a decrease in ThT binding with incubation temperature implies inhibition or destabilization of VP6 aggregates at lower temperatures. This suggests a less favorable or slower aggregation process, increasing the likelihood of proper folding of VP6 into its native structure. This finding is consistent with previous reports that have demonstrated how low expression temperatures promote the production of non-classical inclusion bodies, which exhibit significant biological activity (Ami et al., 2005)(Upadhyay et al., 2012)(Peternel et al., 2008).

The observation of VP6 inclusion body formation during heterologous recombinant expression in *E. coli*, despite extensive optimization of experimental culture conditions, prompted us to investigate the potential root cause of aggregation: the APRs within the VP6 protein. Various computational methods have been developed to identify aggregation-prone regions in amyloidogenic proteins. While certain algorithms like TANGO, AGGRESCAN, AGGRESCAN3D, and CamSol primarily focus on predicting β-aggregation prone regions, others such as Waltz, Zyggregator, and Fold-

amyloid target both β-aggregation and amyloid-fibril prone regions. In this study, we employed multiple tools utilizing diverse algorithms to analyse VP6, striving to alleviate biases due to training datasets, parameterization, and the distinct methodological traits inherent in each database. Finally, we shortlisted APRs that exhibited consensus identification by two or more prediction tools. The prediction tools utilized for this investigation have also been utilized for computational determination of APRs in human keratinocyte growth factor (Dastjerdeh et al., 2022), SARS-CoV-2 proteome (Gour & Yadav, 2021), tumour suppressor protein PTEN (Palumbo et al., 2020), Naked Mole-Rat and Mouse proteome (Besse et al., 2022) in previous reports.

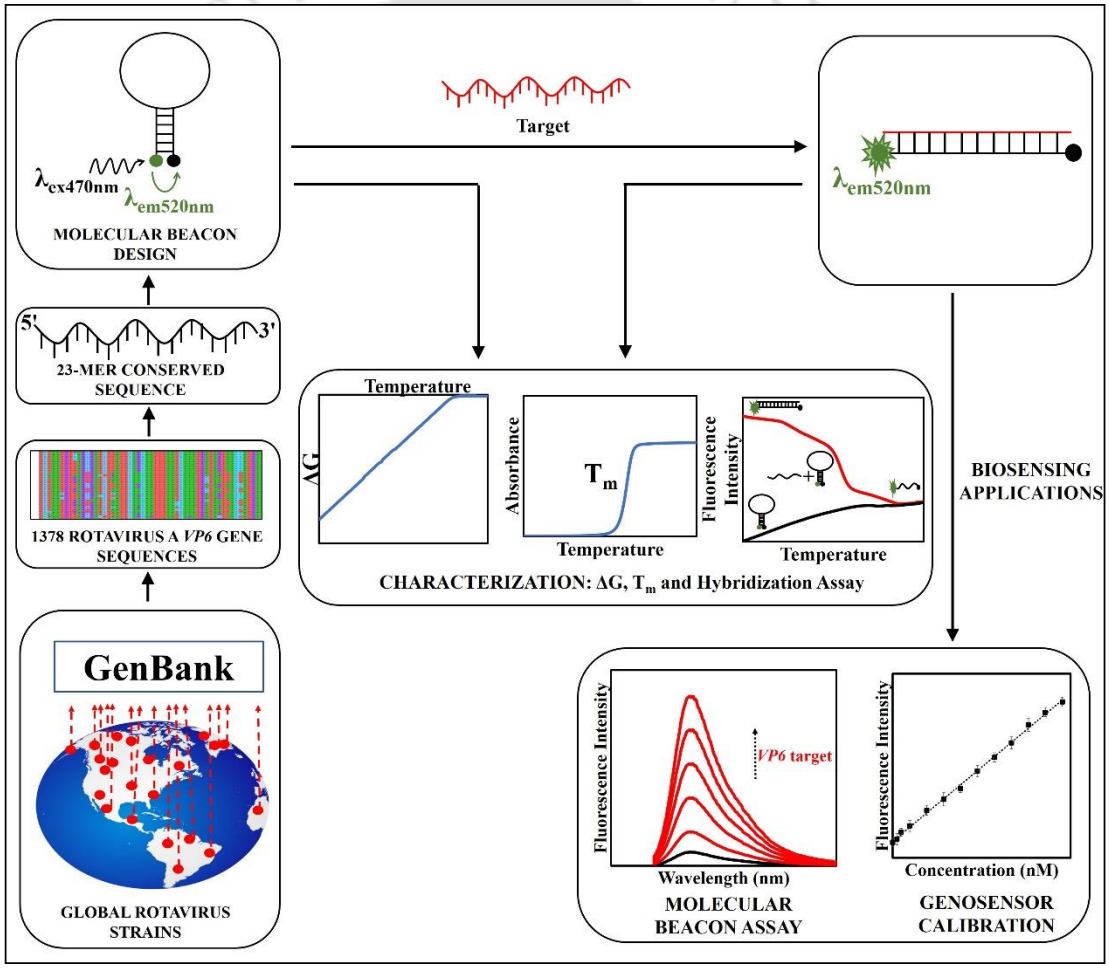
The sequence-based APR prediction tools predicted multiple aggregation hotspot regions in VP6. A total of five, nineteen, ten, nineteen, seventeen, and seven aggregation hotspot sequences were identified by TANGO, AGGRESCAN, Waltz, FoldAmyloid, AmylPred, and ANuPP respectively. 32 residues were identified as aggregation prone residues by Zyggregator, while 30 regions involving residues as well as sequences were identified by Camsol intrinsic. The tools for predicting APRs based on structure restrict the quantity of anticipated APRs depending on the dynamically-exposed hydrophobic or aggregation-prone residues or regions. A majority of the total identified APRs by different tools exhibited overlaps with each other, and the ones in consensus with two or more tools were taken into consideration. In case of overlapping sequences, we considered the residues that were common within the overlapped region. This resulted in the consensus identification of total 24 APRs (**Figure 3.13**). All the consensus epitopes were successfully mapped on the native VP6 three-dimensional structure to assess their surface accessibility. This revealed 21 APRs as solvent-exposed and 3 were found to be buried inside the protein. Furthermore, the consistent predominance of hydrophobic amino acids in all predicted APRs further elucidates their contribution to VP6 aggregation and, consequently, reduced solubility of the protein. Therefore, these 21 surface exposed APRs may promote misfolding or aggregation of VP6 during prokaryotic heterologous expression. Computational identification of these APRs corroborates our experimental observation of VP6 inclusion body or aggregate formation.

3.5 Conclusion

VP6 expression in *E. coli* BL21(DE3) led to its localization in inclusion bodies despite exhaustive optimization of culture conditions. This indicates its inherent tendency for aggregation during prokaryotic expression. Computational investigation revealed 24 APRs, 23 of which are mapped as surface-exposed APRs in native VP6 structure, predominantly enriched in hydrophobic amino acids, corroborating our experimental observations. The identified regions can be considered as the candidate positions to create the VP6 mutant varieties with reduced aggregation propensity through site-directed mutagenesis in prokaryotic expression systems. It is noteworthy that the presence of exposed hydrophobic patches on VP6 is likely crucial for its intermolecular interactions that drive the formation of VP6 trimers and ultimately the assembly of the VP6 capsid structure. These patches may serve as critical interaction sites necessary for maintaining the structural integrity of the viral capsid. Concurrently, our identification of multiple APRs within VP6 provides a basis for potential mutagenesis strategies aimed at enhancing protein solubility, particularly for recombinant expression in *E. coli*. However, the decision to mutate these regions must carefully balance the goal of improving solubility with the imperative to preserve VP6's structural and immunological integrity within the viral capsid. This dual consideration is essential to avoid compromising VP6's functional roles in capsid assembly and its potential immunogenicity. The future scope of this research involves providing a detailed rationale for selecting mutation positions within APRs and experimentally validating these mutations. Specific mutation sites can be determined based on factors such as hydrophobicity, structural motifs, conservation analysis, surface accessibility relevant empirical evidence from prior studies. Subsequently, the study progresses to conducting mutagenesis, expression, and solubility analysis of VP6 mutants through recombinant expression in *E. coli*, aiming to ascertain the potential enhancement of VP6 solubility. Given the pivotal role of VP6 as a prospective rotavirus vaccine candidate owing to its high conservation and abundance in the rotavirus capsid structure, its successful solubilization promises to catalyse further advancements in pathological elucidation, therapeutic interventions, and diagnostic applications. This investigation represents the first comprehensive analysis of APRs within the RVA VP6 capsid protein, marking a notable stride in elucidating its molecular complexities.

Chapter 4

Development of a Molecular Beacon-Based Genosensor for Detection of Human Rotavirus A



4.1 Overview

Rotavirus detection techniques encompass a range of methods including EM (Ong & Chandran, 2005), cell culture (Babaei et al., 2022), immunoassays like ELISA (Peruski & Peruski, 2003) and LAT (Xiang et al., 2020), molecular techniques such as PCR and RT-PCR (Guliy et al., 2019). However, these conventional virus detection methods are limited by the need for specialized, costly instrumentation, the requirement for skilled labour, and their incompatibility with point-of-care settings (Babaei et al., 2022) (Kuri et al., 2020) (Rovida et al., 2013). In this regard biosensors may play crucial role in early diagnosis and management of rotavirus led diseases as these technologies have already emerged as a reliable sensing tool for diverse applications including healthcare sectors (Kuri et al., 2020).

While biosensor technologies have advanced significantly across various fields, their development for RVA detection remains limited (**Table 1.4**, chapter 1), with most relying on antibodies as bioreceptors, which are specific but prone to low stability, cross-reactivity, high costs, and complex production processes. Nucleic acids offer superior stability, specificity through complementary base pairing, cost-effectiveness, and versatility as aptamers, probes, or molecular beacons. However, existing non-antibody-based rotavirus sensors are often inadequately validated and face challenges like complex fabrication, lengthy processing, and high costs. To overcome these limitations, we propose a straightforward and versatile molecular beacon-based detection system for rotavirus.

Molecular beacons are nucleic acid-based probes that utilize a stem-loop structure with a fluorophore and a quencher to detect specific sequences. In the absence of the target sequence, the probe's fluorophore and quencher are in close proximity, resulting in quenching of fluorescence. Upon hybridization with the target sequence, the probe's structure changes, separating the fluorophore from the quencher and allowing fluorescence emission. This principle provides specificity and sensitivity, making molecular beacons a valuable tool for rapid, accurate virus detection with minimal sample preparation. These characteristics enhance the effectiveness of virus detection (Stobiecka & Chałupa, 2015; Bidar et al., 2021).

In the current study, we have made an endeavour to develop a molecular beacon-based fluorometric biosensor for the detection of the human rotavirus *VP6* DNA segment as

the target, chosen for its high conservation across human RVA strains, which ensures enhanced reliability and accuracy in detection. This biosensor employs a fluorescein-dabcyl pair as the fluorophore-quencher system and utilizes FRET to achieve sensitive and specific detection. This system has the potential to integrate with a variety of diagnostic technologies, facilitating the development of highly sensitive and specific optical biosensing platforms.

4.2 Methodology Section

4.2.1 Design and Synthesis of Molecular Beacon and Target Sequences

A total of 1,378 full-length human RVA *VP6* gene sequences were retrieved from the GenBank database of NCBI in FASTA format and aligned using MEGA 7 software. The alignment was analysed to identify conserved regions within the VP6 gene sequence. The complementary molecular beacon probe was designed by first determining its reverse complement using SnapGene software (www.snapgene.com). To construct the molecular beacon stem region, multiple 5-nucleotide stem sequences were randomly generated using Excel's RANDBETWEEN function and concatenated. A suitable stem sequence was then appended to the probe sequence, resulting in the final molecular beacon design.

4.2.2 Structural Profiling of the Molecular Beacon

The melting temperature (T_m) and Gibbs free energy change (ΔG) value of the designed molecular beacon were determined by ApE, A Plasmid Editor (Davis & Jorgensen, 2022) and RNA fold, respectively. Secondary structure predictions for the molecular beacon's loop region, complementary to the target sequence, as well as for the complete molecular beacon incorporating both loop and stem regions, were conducted using the RNAfold webserver (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi)(Gruber et al., 2008). This analysis was performed to identify and assess any potential internal secondary structures within the loop region of molecular beacon as strong secondary structures may interfere with complementary target binding. For tertiary structure prediction, the 3dRNA/DNA web server (http://biophy.hust.edu.cn/new/3dRNA/)(Xiong, et al., 2023) was utilized to obtain a

three-dimensional model of the molecular beacon, providing insights into its spatial configuration and confirming the integrity of the designed structure.

4.2.3 Thermal Stability Assessment of Molecular Beacon-Target Complex

The thermal stability of molecular beacon-target complex was assessed by multiple analyses. Firstly, we used the DINAMelt Quickfold two-state melting hybridization tool to generate a temperature vs. ΔG plot. The molecular beacon and its complementary target sequences were input into the tool. ΔG predictions were conducted one temperature at a time, ranging from 0°C to 100°C in 1°C increments. The resulting ΔG values were plotted against temperature to visualize changes in the stability of the molecular beacon-target complex. Secondly, the thermal stability of the molecular beacon-target complex was predicted using the DINAMelt application, employing the "Hybridization of Two Different Strands of DNA or RNA" function. The analysis was conducted over a temperature range from 0°C to 100°C. Both the molecular beacon and its complementary target were maintained at a concentration of 200 nM, with a sodium ion concentration of 157 mM. Absorbance versus temperature data were generated. Furthermore, the T_m was corroborated by plotting a derivative graph (dA/dT vs. temperature). Thirdly, the T_m of the molecular beacon-target complex was experimentally assessed using a FRET melting assay. This assay was performed over a temperature range from 25°C to 95°C using a Rotor-Gene Q real-time PCR cycler (QIAGEN). Fluorescence intensity versus temperature data were collected to determine the T_m. Additionally, a derivative plot of fluorescence intensity versus temperature (dF/dT vs. temperature) was analysed to accurately ascertain the T_m of the molecular beacon-target complex.

4.2.4 Spectral Analysis of Fluorescein, Dabcyl, and Molecular Beacon-Target Interaction

Absorbance and fluorescence spectra of fluorescein were measured to characterize its excitation and emission properties using Thermo Scientific Varioskan™ LUX Multimode Microplate Reader. Fluorescein, dissolved in 1x phosphate buffered saline (PBS), was analysed to obtain the absorbance spectrum in the range of λ_{370nm} to λ_{700nm}. Fluorescence emission spectra were recorded with an excitation wavelength of λ_{490nm}. For dabcyl, absorbance spectra were measured in 1x PBS buffer over the range of λ_{370nm}

to λ_{560nm}. The spectral overlap between the emission spectrum of fluorescein and the absorbance spectrum of dabcyl was assessed to evaluate the potential efficiency of FRET in the molecular beacon system.

Additionally, fluorescence spectra of the molecular beacon, both in the presence and absence of the synthetic *VP6* target, were recorded to evaluate fluorescence quenching and enhancement. The molecular beacon and synthetic *VP6* target were each prepared at a concentration of 1 µM in 1x PBS. The molecular beacon, the molecular beacon-*VP6* target complex, and the *VP6* target were excited at λ_{490nm}. Fluorescence emission spectra were collected from λ_{480nm} to λ_{650nm} using a fluorescence spectrophotometer (HORIBA scientific Fluorimax-4 instrument). Measurements were performed at room temperature, and fluorescence readings were taken 20 minutes post-mixing to ensure equilibrium conditions.

4.2.5 Optimization of Excitation Wavelength and Incubation Time for Fluorescence Assays with Molecular Beacons

Fluorescence spectra of the molecular beacon with fluorescein-Dabcyl pair were analyzed to determine the optimal excitation wavelengths for target detection. Measurements were conducted using excitation wavelengths ranging from λ_{450nm} to $\lambda_{490\ nm}$, with emission spectra recorded from approximately from λ_{500nm} to λ_{650nm}. Both the molecular beacon and the *VP6* target were prepared at a concentration of 1 µM in 1x PBS. The fluorescence spectra were collected at room temperature following a 20-minute incubation period to ensure stable conditions.

To determine the optimal incubation time, fluorescence measurements were conducted using 200 nM of the molecular beacon with fluorescein-dabcyl and varying concentrations of *VP6* target ranging from 0 nM to 200 nM. The assays were performed at room temperature in 1x PBS. Fluorescence readings were recorded every 10 minutes for a total duration of 70 minutes, with an excitation of λ_{470nm}.

4.2.6 Calibration Curve and Specificity Assessment of Molecular Beacon for VP6 Target

To establish the calibration curve, a 1 µM molecular beacon solution in 1× PBS was prepared. *VP6* target concentrations from 0 nM to 200 nM were mixed with the 200 nM molecular beacon solution, with experiments conducted in triplicate. The mixtures were

incubated at room temperature for 10 minutes. Fluorescence emission spectra were recorded using a fluorescence spectrophotometer with an excitation of λ_{470nm}. Fluorescence intensity values, averaged from triplicate measurements, were plotted against target concentrations to generate the calibration curve. The R-squared (R^2) value was used to assess the curve's fit. The limit of detection (LOD) was calculated using the formula, $LOD=3\sigma_{blank}/s$, where "σ_{blank}" is the standard deviation of blank measurements and "s" is the slope of the calibration curve.

For the interference study, the molecular beacon was prepared at 1 µM in 1× PBS and incubated with synthetic oligonucleotides containing single mismatch, double mismatch, and scrambled sequences, all at the same concentration, with experiments conducted in triplicate. After a 10-minute incubation at room temperature, fluorescence emission spectra were recorded with an excitation of λ_{470nm}. Specificity was assessed by calculating the percent difference in fluorescence intensity between the complementary target and each mismatched or scrambled sequence. The signal-to-noise ratio (SNR) was calculated using the formula: SNR = Δ(target-background fluorescence signal)/√(background signal). The basal fluorescence of molecular beacon without the target was considered as the background signal.

4.3 Results

4.3.1 Design and Synthesis of Molecular Beacon and Target Sequences

A highly conserved sequence of 23 nucleotides, 5'GCTAGGGATAAGATTGTTGAAGG3', was identified and designated as the target from the alignment of full length 1378 human RVA *VP6* gene sequences from NCBI (**Appendix A, Table A.4.**). The reverse complement of this sequence, determined using SnapGene software (www.snapgene.com), was 5'CCTTCAACAATCTTATCCCTAGC3' which served as the probe. A suitable stem sequence was integrated into the probe, producing the final molecular beacon sequence of 33 nucleiotides, with stem and loop regions: 5' [Fluorescein]ATAGTCCTTCAACAATCTTATCCCTAGCACTAT[Dabcyl] 3'. A scrambled control sequence was generated using the OligoRAnd tool (http://mkwak.org/oligorand/index.html), yielding 5'

AGCTCGCGGTCCATACCGCCATC 3'. Single and double mismatch sequences were designed as 5' GCTAGGG**C**TAAGATTGTTGAAGG 3' and 5' GCTAGGGATAA**A**ATTGTTGA**C**GG 3', respectively. The designed molecular beacon, with fluorescein conjugated to the 5' end and dabcyl conjugated to the 3' end, as well as the target, single mismatch, double mismatch and scrambled sequences, were synthesized and obtained from Integrated DNA Technologies (IDT, USA).

4.3.2 Structural Profiling of the Molecular Beacon

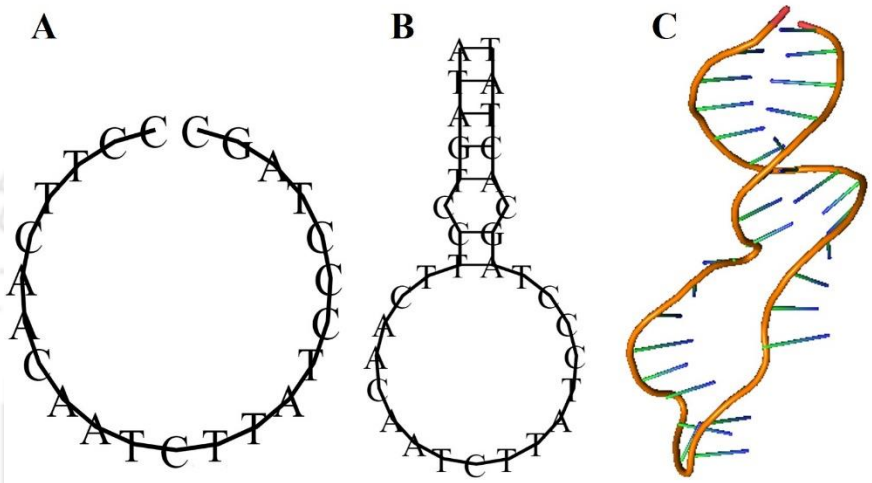


Figure 4.1. Structural characterization of molecular beacon. A. Secondary structure of loop region of molecular beacon (without stem) predicted by RNAfold webserver. B. Secondary structure prediction of molecular beacon (loop and stem) predicted by RNAfold webserver. C. 3D structure prediction of molecular beacon using 3dRNA/DNA webserver, visualized by PyMol.

The T_m and ΔG value of the designed molecular beacon were determined to be 54°C and -0.9 kCal/mol. Secondary structure predictions of the molecular beacon were performed using RNAfold represented in **Figure 4.1**. The probe region, complementary to the target sequence, demonstrated no intramolecular base pairing or internal secondary structure, aligning with its design as a single-stranded segment in the absence of the target. In contrast, the complete molecular beacon, including both the loop and stem regions, exhibited a well-defined hairpin structure, confirming the successful design of a stable stem-loop configuration. This finding supports the intended functionality of the probe within the molecular beacon. Further, analysis of three-dimensional structure of molecular beacon using available tools (Bachu et al., 2023)

and methods can further support molecular beacon design. The tertiary structure predicted by the 3dRNA/DNA web server revealed a stable stem-loop conformation consistent with the design parameters and secondary structure predictions. This structural arrangement suggests that the molecular beacon will maintain a compact and stable conformation under experimental conditions, facilitating effective quenching by holding the fluorescein and dabcyl groups in close proximity. These findings validate the design of the molecular beacon and support its anticipated performance in molecular detection.

4.3.3 Thermal Stability and Melting Profile Study of Molecular Beacon-Target Complexes

The temperature vs. ΔG plot for the hybridization of the molecular beacon-target complex reveals distinct phases of thermal behaviour (**Figure 4.2.A**). In the initial phase (0°C to 60°C), the ΔG values are highly negative, indicating a stable molecular beacon complex. As the temperature increases, ΔG becomes less negative, reflecting a gradual decrease in the stability of the complex. This trend signifies the heating process where the molecular beacon begins to denature and separate from its target. At temperatures above 60°C, ΔG values approach zero, indicating that the complex is predominantly denatured and no longer stable. The curve flattens, showing that further increases in temperature have minimal effect on ΔG as denaturation is nearly complete. Overall, the plot effectively illustrates the denaturation process of the molecular beacon, demonstrating its stability at temperatures up to around 60°C, with the progressive loss of stability with increasing temperature. This analysis is crucial for understanding the thermal stability and behaviour of molecular beacons under varying temperature conditions.

The predicted normalized absorbance measurements over a temperature range of 0°C to 100°C is shown in **Figure 4.2.B**. The normalized absorbance versus temperature graph, known as the melting curve, demonstrated a gradual increase in absorbance with temperature, indicating the transition from a double-stranded to a single-stranded state. The Tm was calculated to be approximately 59°C, determined from the midpoint of this

transition. This smooth and well-defined transition suggests that the molecular beacon-target complex is homogeneous and exhibits stable molecular interactions.

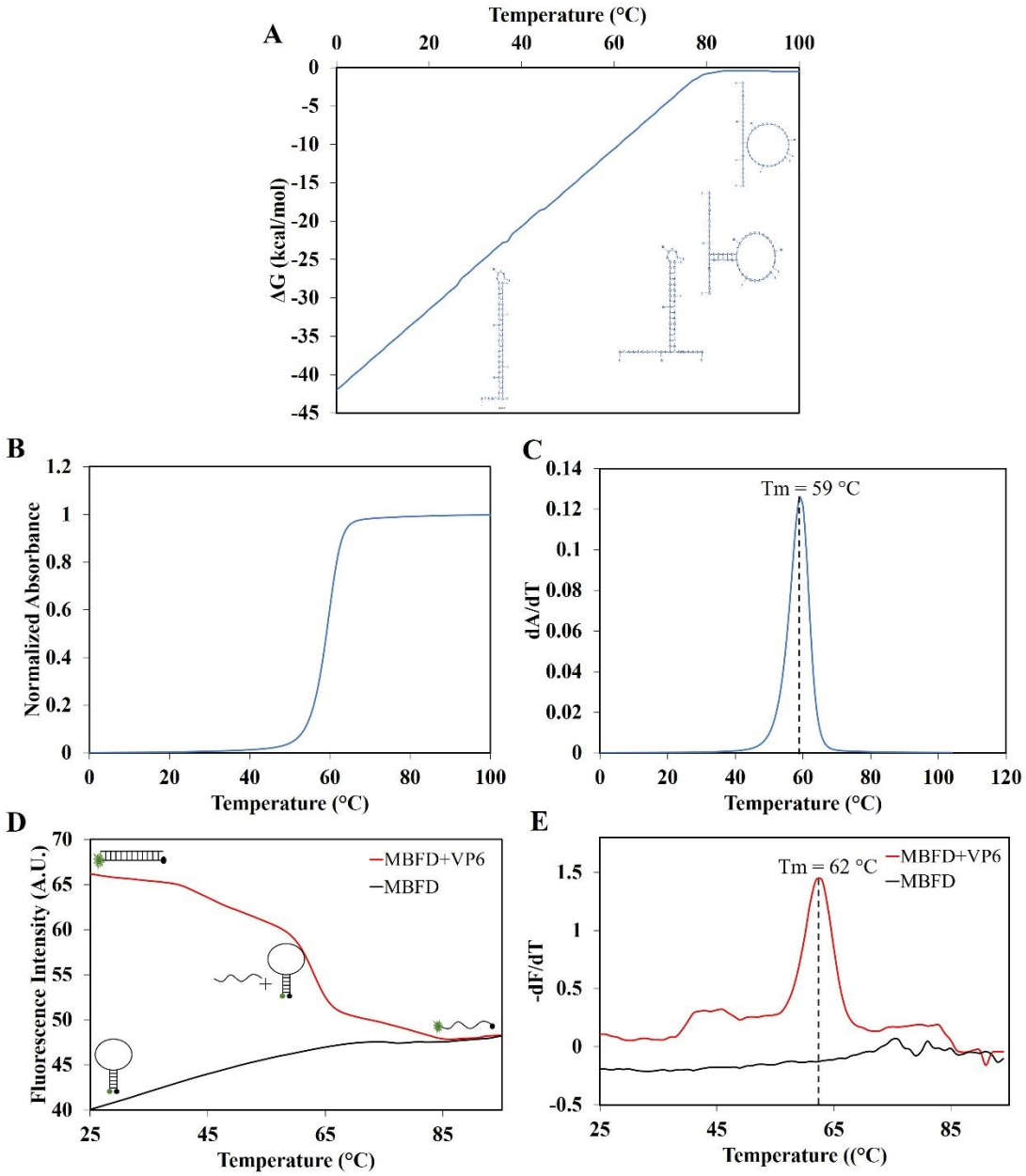


Figure 4.2. Thermal stability analysis of molecular beacon-target complex. A. Temperature vs. Gibbs free energy change (ΔG) plot. Melting curve analysis of the molecular beacon-target complex, showing B. normalized absorbance at $\lambda_{260\ nm}$ versus temperature from 0°C to 100°C and C. the derivative of absorbance (dA/dT) versus temperature from 0°C to 100°C. FRET melting assay of the hybridized complex of molecular beacon with fluorescein-dabcyl pair with the target. D. Fluorescence versus temperature graph from 25°C to 95°C; E. Derivative of fluorescence (dF/dT) versus temperature from 25°C to 95°C.

Further validation was achieved through the analysis of the derivative of absorbance (dA/dT) versus temperature, referred to as the derivative melting curve, presented in **Figure 4.2.C**. This graph revealed a sharp peak at around 59°C, signifying a high degree of cooperativity in the melting transition and indicating a well-structured and stable molecular beacon. The pronounced peak, devoid of broad or multiple peaks, supports the uniform melting behaviour of the sample. According to this data, the molecular beacon-target complex can be utilized at molecular detection assay temperatures below its T_m of 59°C. This ensures that the beacon remains in its stable, closed stem-loop structure or in a double-stranded form hybridized with the target and delivers reliable signal detection. Temperatures exceeding 59°C will cause the complex to transition to a single-stranded state, which could result in erroneous signal generation.

To experimentally evaluate the thermal stability of the molecular beacon-target complex, fluorescence melting curve analysis was performed. The fluorescence melting curve shown in **Figure 4.2.D**, which plots fluorescence intensity versus temperature, revealed that in the absence of the target, the molecular beacon remains in a closed stem-loop configuration at lower temperatures. In this conformation, the fluorophore and quencher are in close proximity in the Fluorescein-molecular beacon-Dabcyl assembly, resulting in minimal fluorescence emission due to FRET, where the quencher efficiently absorbs the fluorescence emitted by the fluorophore. As the temperature increases, the stem-loop structure melts, leading to the separation of the fluorophore and quencher and a corresponding gradual increase in fluorescence intensity as the FRET effect diminishes. Conversely, when the molecular beacon is hybridized with its target, fluorescence is detectable at lower temperatures, indicating the formation of the molecular beacon-target complex and the disruption of the stem-loop structure. This hybridization increases the distance between the fluorophore and quencher, leading to an increase in fluorescence intensity due to reduced FRET. As the temperature rises further, the fluorescence intensity decreases, reflecting the dissociation of the molecular beacon-target complex. This decrease is due to the reformation of the stem-loop structure in the molecular beacon, which brings the fluorophore and quencher back into close proximity, reducing fluorescence emission through FRET. However, at higher temperatures, the beacons denature completely, no longer forming a stem-loop structure, and begin to emit a certain level of fluorescence as they transition to a single-stranded state. This marginal increase in fluorescence intensity at higher temperatures is due to the absence of the stem-loop structure and the resultant separation of the

fluorophore and quencher. Cartoons in the figure illustrate the various molecular beacon states at specific temperatures, depicting the presence or absence of the target and the associated changes in fluorescence. The T_m of the molecular beacon-target complex is presented in **Figure 4.2.E**. The T_m was observed to be 62°C. This indicates that the molecular beacon-target complex is stable and exhibits reliable fluorescence signalling below this temperature, making it suitable for molecular detection assays conducted at or below 62°C. Below this temperature, free molecular beacons remain non-fluorescent due to the close proximity of the fluorophore and quencher, while the probe-target hybrids form and are stable, ensuring effective assay performance.

4.3.4 Spectral Analysis of Fluorescein, Dabcyl, and Molecular Beacon-Target Interaction

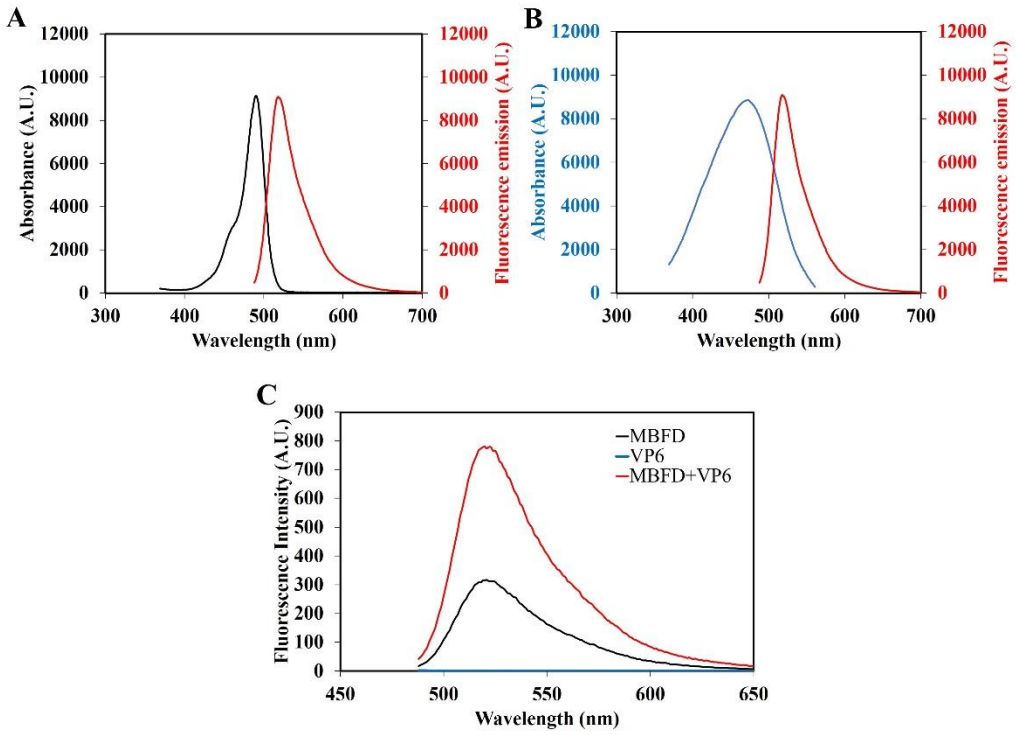


Figure 4.3. Spectral analysis of fluorescein, Dabcyl, and molecular beacon with fluorescein-dabcyl pair (MBFD)-target interaction: A. Fluorescein absorption and emission spectra; B. Dabcyl absorbance and fluorescein emission spectra; C. FRET analysis of the molecular beacon in the presence and absence of the target.

The absorbance spectrum of fluorescein in 1x PBS displayed a broad excitation profile with a peak at $\lambda_{490\ nm}$, while its fluorescence emission spectrum extended from $\lambda_{480\ nm}$ to $\lambda_{600\ nm}$, with a peak at $\lambda_{517\ nm}$ (**Figure 4.3.A**). The absorbance spectrum of dabcyl in 1x PBS spanned $\lambda_{370\ nm}$ to $\lambda_{560\ nm}$, peaking at $\lambda_{470\ nm}$ (**Figure 4.3.B**). Analysis of the

spectral overlap demonstrated a substantial intersection between the emission spectrum of fluorescein ($\lambda_{480\ nm}$ to $\lambda_{600\ nm}$) and the absorbance spectrum of dabcyl ($\lambda_{370\ nm}$ to $\lambda_{560\ nm}$). This overlap indicates efficient FRET from fluorescein to dabcyl, confirming dabcyl's efficacy as a quencher of fluorescein fluorescence.

FRET modulation as a result of molecular beacon-target interaction is represented in **Figure 4.3.C**. Fluorescence measurements of the molecular beacon alone revealed basal fluorescence across the range of $\lambda_{480\ nm}$ to $\lambda_{650\ nm}$, attributable to effective quenching by Dabcyl. This quenching is a result of the close proximity between fluorescein and Dabcyl within the stem-loop configuration of the molecular beacon. The synthetic *VP6* target alone did not exhibit any fluorescence. Upon hybridization of the molecular beacon with the synthetic *VP6* target, a notable increase in fluorescence intensity was observed. This enhancement indicates the disruption of the stem-loop structure and the consequent separation of the fluorophore from the quencher, leading to a reduction in FRET efficiency. The substantial increase in fluorescence of the molecular beacon-target complex relative to the molecular beacon alone corroborates successful target hybridization and underscores the molecular beacon's capability for effective fluorescence-based detection of the *VP6* target sequence.

4.3.5 Optimization of Excitation Wavelength and Incubation Time for Fluorescence Assays with Molecular Beacon

In the fluorescence intensity measurements of the molecular beacon without (**Figure 4.4.A**) and with its target (**Figure 4.4.B**), the peak intensities varied across different excitation wavelengths. This data reveals a general trend of decreasing fluorescence intensity with increasing excitation wavelength. The highest intensity was recorded at $\lambda_{450\ nm}$, with a gradual decline observed as the excitation wavelength increased. Based on these findings, an excitation wavelength of 470 nm was selected for the assays due to its optimal balance between emission intensity and the decreasing trend observed at longer wavelengths.

Fluorescence intensity measurements of the molecular beacon-target complex at various time intervals demonstrated that the signal reached a steady state after 10 minutes of incubation (**Figure 4.4.C**). Subsequent measurements showed no significant changes in fluorescence intensity beyond this period, indicating that longer incubation

times do not alter the fluorescence signal. This reflects that a 10-minute incubation is sufficient to achieve reliable fluorescence measurements, indicating a short response time offered by the genosensor.

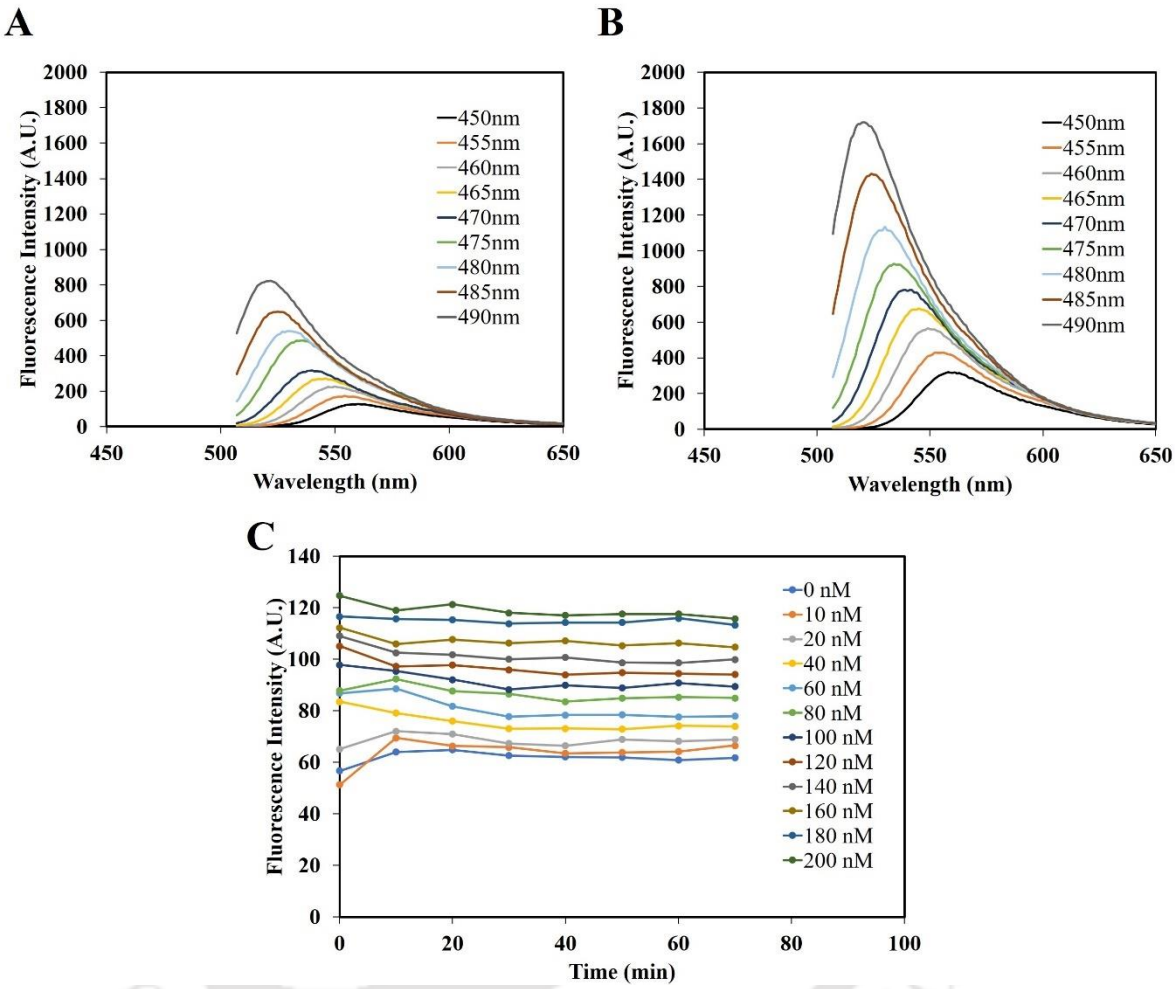


Figure 4.4. Optimization of excitation wavelength and incubation time for the molecular beacon-based assay. Fluorescence spectra of the molecular beacon at different excitation wavelengths in the A. absence of the VP6 target and B. presence of the VP6 target. C. Time optimization of molecular beacon fluorescence readings, with target concentrations ranging from 0 to 200 nM, monitored from 0 to 70 minutes.

4.3.6 Response and Specificity Assessment of the Genosensor

Fluorescence intensity measurements were taken for various concentrations of the *VP6* target mixed with a constant concentration of the molecular beacon (1 µM), demonstrated in **Figure 4.5.A**. The intensity increased linearly with the target concentration, resulting in a calibration curve described by the equation y=0.3834x+59.131, where 'y' represents fluorescence intensity and 'x' represents target

concentration (**Figure 4.5.B**). The R^2 value of 0.9983 indicates a strong correlation between fluorescence intensity and target concentration, confirming the effectiveness of the molecular beacon in detecting the *VP6* target. The LOD was calculated to be 18.8 nM, demonstrating the beacon's ability to reliably detect low concentrations of the *VP6* target. The assay exhibited a linear response up to the tested 200 nM target concentration, with the probe concentration held at 1 µM, ensuring it was in excess relative to the target. This approach prevents the probe from being a limiting factor and avoids saturation, ensuring that the observed signal accurately reflects the target concentration. The probe concentration can be adjusted to fix the dynamic range according to specific application requirements.

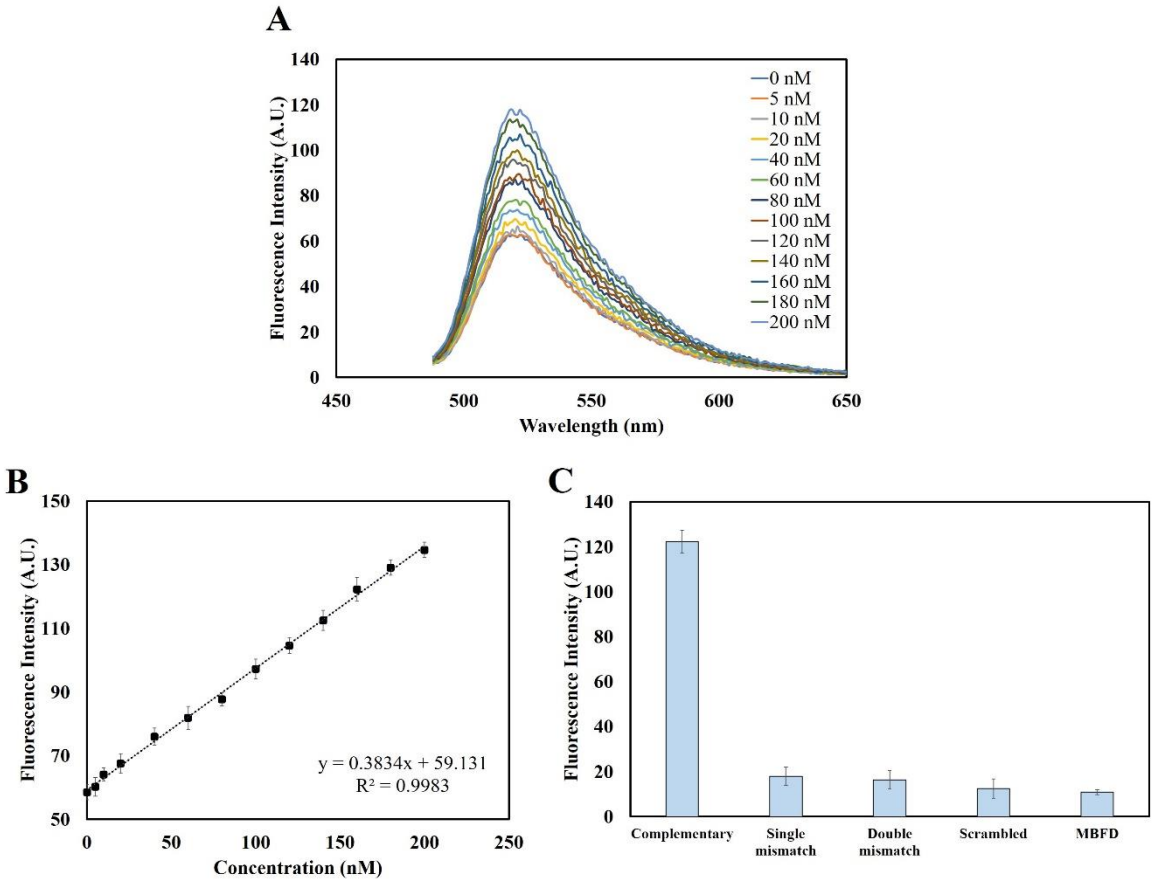


Figure 4.5. Calibration curve and specificity assessment of the molecular beacon for the VP6 target. A. Fluorescence intensity of molecular beacon with fluorescein-dabcyl pair after incubation with different concentrations of the target RVA synthetic VP6 ssDNA target. B. Calibration plot of the peak fluorescence intensity at λ_{521nm} (with λ_{470nm} excitation wavelength) against the concentration of the target RVA VP6 ssDNA in 1X PBS buffer (pH 7.4). The error bars represent the standard deviations of three repeated measurements. C. Specificity of the molecular beacon biosensor for its target

sequence, comparing the biosensor's response after incubation with 200 nM RVA VP6 target, single mismatch, and scrambled sequences at similar concentrations.

The specificity of the molecular beacon was assessed through an interference study, comparing its fluorescence response to the *VP6* target with responses to various non-target sequences, including single mismatch, double mismatch, and scrambled sequences (**Figure 4.5.C**). Percentage differences in fluorescence intensity between the *VP6* target and each non-target sequence were: 85.34% for single mismatch, 86.64% for double mismatch, 89.78% for scrambled sequence, and 91.12% for the background signal of the molecular beacon without any target. These high percent differences indicate that the molecular beacon has a strong ability to discriminate between the *VP6* target and non-target sequences, demonstrating specificity.

Additionally, the signal-to-noise ratio (SNR) was calculated to evaluate how effectively the molecular beacon distinguishes between the target and non-target sequences. For the *VP6* target, the SNR was 37.09, which reflects a strong and clear fluorescence signal relative to the background noise, indicating effective detection of the target. Conversely, the SNRs for single mismatch, double mismatch, and scrambled DNA were lower (5.44, 4.95, and 3.79, respectively), showing that these sequences produce much weaker signals. This lower SNR for non-target sequences confirms that the molecular beacon specifically recognizes the *VP6* target and does not produce false positives for mismatched or scrambled sequences.

4.4 Discussion

The rapid detection of RVA remains a critical challenge, with limited advancements in biosensor development for this pathogen. Rotavirus detection can be accomplished by targeting various components of the virus, including its genetic material, antigenic capsid proteins, or entire virus particles. Few biosensors with different platforms for rotavirus detection have been reported in the literature, which have been summarized in **Table 1.4**. However, majority of the biosensors developed so far for rotavirus uses antibody as biorecognition elements. As discussed previously, the antibody-based recognition elements are prone to environmental sensitivity, potential cross-reactivity, and levied high production costs. Moreover, the low stability and its linked poor shelf-life of this protein-based bioreceptor is a major concern that prompted the current

biosensors research to incline towards the non-protein biorecognition system including naturally stable DNA-based recognition molecules. An unconventional recognition system, utilizing silver nanoparticles, has also been proposed for the direct detection of rotavirus (Biswas et al., 2023). However, the proof-of-concept for the specificity mechanism of this device has not yet been adequately elucidated. Moreover, the sensing mechanism relies on sophisticated instrument such as Raman spectroscopy, which is not yet adequately developed for its deployment in point-of-care (POC) settings. In another effort, a complementary DNA probe-based biosensor utilizing F0F1-ATPase has been reported for rotavirus detection(Zhang & Zhao, 2015). However, its development involves a complex and cumbersome fabrication process, which limits its practical and widespread application.

Nucleic acid-based bio-receptors including molecular beacons, present a more robust alternative. One key advantage of molecular beacons is their potential for multiplexing, which can be achieved by designing multiple molecular beacons, each labelled with distinct fluorophore-quencher pairs and optimized for different excitation and emission wavelengths. By coupling these uniquely designed molecular beacons, each targeting specific rotavirus sequences, researchers can simultaneously detect different human rotavirus groups and strains within a single assay. This approach is applicable in real-time PCR assays, where distinct excitation wavelengths allow differentiation of various beacons, as well as in RT-qPCR for quantifying RNA after conversion to cDNA. Molecular beacons can also be used in FRET (Belon & Frick, 2008) (Krusiński et al., 2010) to monitor real-time interactions and dynamics of rotavirus RNA. Beyond these techniques, molecular beacons are applicable in microarrays for high-throughput detection, FISH (Zheng et al., 2015) for spatial gene expression analysis, and cellular imaging for monitoring gene expression within live cells (Chen et al., 2017). They can also be integrated into point-of-care devices and biosensors for rapid, on-site testing (Hua et al., 2022). The ability to multiplex and accurately detect different rotavirus strains not only enhances diagnostic capabilities but also supports epidemiological and surveillance studies by providing comprehensive insights into the distribution, prevalence, and diversity of rotavirus infections across different populations and regions. Their ease of modification and cost-effective production make them an attractive option for designing versatile and sensitive biosensors (Kuri et al., 2020). Notably, only one study has utilized molecular beacons for the purpose that specifically used for cell culture imaging of rotavirus (Bertol & Gatti, 2016). This approach is

however, mired by lengthy processing times, high costs associated with specialized instrumentation, and the need for skilled personnel. Thus, the application of molecular beacons in rotavirus detection needs to be aggressively explored. Therefore, our research focuses on advancing the use of molecular beacons, from design to experimental validation.

4.5 Conclusion

In this study, we successfully designed, developed and optimized a molecular beacon-based fluorescence biosensor for the detection of RVA *VP6* DNA fragment. Our findings demonstrate that the molecular-beacon biosensor exhibits a high degree of selectivity and sensitivity for the *VP6* target sequence, with a clear fluorescence response that correlates well with the concentration of the target ssDNA. The FRET melting assay confirmed effective hybridization between the molecular beacon and its target, and the calibration curve indicated a strong linear relationship between fluorescence intensity and target concentration. The biosensor achieved sensitivity in lower nano molar range, highlighting its capability for detecting very low concentrations of the target. Additionally, the biosensor is specific to the target sequence as it does not produce significant signal with non-target sequences, underscoring its potential for precise detection in diagnostic applications. These results highlight the potential of the molecular beacon to enhance diagnostic capabilities by providing a rapid, reliable, and cost-effective alternative to traditional methods. Its high sensitivity and specificity make it a promising tool for laboratory and field applications, representing a significant advancement in detecting rotavirus VP6 target. The next level of study should integrate the developed genosensor in a suitable platform with a hand-held fluorimeter to analyze clinical samples for its POC deployment.

Chapter 5

Conclusion and Future Outlook

5.1. Conclusion

5.1.1. Molecular Epidemiology and Surveillance

The comprehensive sequence data mining and phylogenetic analysis conducted in this study provide a critical baseline for ongoing molecular epidemiological surveillance of RVA in India. The identification of prevalent genotypes and their evolutionary relationships enhances our understanding of the genetic diversity, evolution, and transmission dynamics of RVA strains circulating within the country. These findings are essential for tracking the emergence of novel variants and support public health strategies aimed at controlling RVA outbreaks.

5.1.2. Multiepitope Subunit Vaccine

The immunoinformatics-driven identification of multiple RVA epitopes offers a strategic approach to developing multiepitope subunit vaccines targeting highly conserved and immunogenic regions of the virus. This approach holds promise for eliciting broad-spectrum immune responses against diverse RVA strains. By focusing on the most relevant epitopes, the study provides a foundational framework for future vaccine development efforts that can be tailored to regional and strain-specific variations of RVA, potentially enhancing vaccine efficacy and coverage. Furthermore, the subunit vaccine mitigates the risk of intussusception since it is not derived from live components.

5.1.3. Protein Expression and Structural Studies

The successful optimization of recombinant RVA VP6 protein expression in *Escherichia coli* and the detailed analysis of its aggregation-prone regions provide crucial insights into the structural and functional attributes of this key viral protein. The findings support further efforts to engineer VP6 variants with enhanced solubility and stability, which are vital for both structural studies and therapeutic applications. Understanding the structural determinants of aggregation will also inform the design of VP6-based diagnostics and antiviral strategies targeting essential viral processes.

5.1.4. Genosensor Development for RVA Detection

The development of a fluorometric genosensor utilizing a molecular beacon with a fluorescein-dabcyl pair represents a significant advancement in the rapid and sensitive detection of RVA. This innovative diagnostic tool offers the potential for real-time monitoring of RVA in clinical and environmental samples, providing critical data for outbreak surveillance and public health interventions. The genosensor's high selectivity and sensitivity compared to conventional methods underscore its potential utility in clinical diagnostics and epidemiological studies.

5.2. Future Outlook

5.2.1. Molecular Epidemiology and Surveillance

Future research should prioritize expanding the existing RVA sequence database by incorporating more recent and geographically diverse samples from across India and other regions. This expansion will facilitate a more detailed assessment of the spatial and temporal distribution of RVA genotypes, thereby enhancing the resolution of molecular epidemiological studies. Additionally, integrating advanced computational modelling and phylogeographic analysis with molecular data could elucidate the factors influencing RVA transmission dynamics, including host-pathogen interactions, environmental influences, and the impact of vaccination programs. Real-world validation of findings through clinical sample testing from multiple geographic locations is crucial to assess genotype prevalence and monitor viral evolution in response to selective pressures.

5.2.2. Multiepitope Subunit Vaccine

The identified epitopes require thorough validation through in vitro and in vivo studies to confirm their immunogenicity, safety, and ability to confer protective immunity against RVA infection. Advanced immunological assays, including ELISA, neutralization tests, and cytokine profiling, should be employed to characterize the immune response elicited by the multiepitope subunit vaccine candidates. Further research should explore adjuvant formulations and delivery platforms, such as nanoparticle carriers or liposomes, to enhance the immunogenicity of the vaccine. Epidemiological surveillance data should be leveraged to update the selection of

epitopes, ensuring that the vaccine remains relevant against circulating strains and guiding the inclusion of epitopes in peptide vaccine formulations tailored to regional RVA diversity.

The genetic diversity of rotavirus, driven by frequent genomic reassortment and interspecies transmission, poses a challenge to the long-term efficacy of vaccines. This thesis has focused on designing a multi-epitope vaccine targeting conserved regions of VP4, VP6, and VP7, aiming to provide broad-spectrum protection against prevalent strains, particularly those circulating in India. However, the emergence of new genotypes or variants not covered by the current epitope design remains a possibility.

To address this, future strategies should incorporate:

Ongoing Genomic Surveillance: Regular monitoring of circulating and emerging rotavirus strains using global databases such as NCBI Virus Variation Resource will be critical to identify novel mutations and genotypes that may influence vaccine efficacy.

Epitope Refinement: Computational tools can be leveraged to update the epitope repertoire by identifying newly conserved regions across evolving genotypes. These updates could be integrated into vaccine formulations using modular approaches to ensure continuous relevance.

Broadly Neutralizing Epitope Selection: By expanding the selection criteria to include epitopes that are less likely to undergo mutations due to functional constraints, the vaccine could achieve enhanced durability against genetic drift.

Adaptive Vaccine Platforms: Employing platforms like virus-like particles (VLPs) or mRNA-based vaccines could allow rapid adaptation to emerging strains by substituting or modifying epitope-encoding sequences without redesigning the entire construct.

Efficacy Testing Against Diverse Strains: Experimental validation of vaccine efficacy against a panel of globally and regionally relevant strains, including potential reassortants, will provide critical insights into its real-world applicability.

These measures would enable the proposed multi-epitope vaccine to remain effective in addressing the dynamic genetic landscape of rotavirus, ensuring its utility in diverse geographical settings and across future variants.

5.2.3. RVA VP6 Expression and Structural Studies

Future structural studies should focus on generating engineered variants of VP6 suitable for prokaryotic expression and employ techniques such as X-ray crystallography, cryo-electron microscopy, and nuclear magnetic resonance spectroscopy to resolve the three-dimensional structure of VP6 and its engineered variants. Detailed structural analysis will facilitate the identification of key interaction sites, guide the design of inhibitors, and inform modifications to enhance protein solubility and reduce aggregation. Moreover, in vitro functional assays can be utilized to investigate the interactions of VP6 with host cellular factors, providing insights into viral assembly, replication, and pathogenesis. The validation of these findings in animal models and clinical samples will be essential for translating the structural insights into practical diagnostic and therapeutic applications.

5.2.4. Genosensor Development for RVA Detection

The application of the developed genosensor should be extended to clinical validation studies involving a broad spectrum of patient-derived samples, such as stool specimens, to evaluate its diagnostic accuracy and robustness. Comparative studies against gold-standard diagnostic methods, including ELISA and RT-PCR, will help establish the genosensor's sensitivity, specificity, and potential advantages in terms of cost, speed, and ease of use. Further refinement of the genosensor, including optimization of probe sequences, reaction conditions, stability, cross-reactivity and multiplexing for detection of multiples rotavirus strains in a sample, and in biological sample matrices or clinical samples could enhance its performance for point-of-care applications. Additionally, integrating the genosensor into portable and user-friendly diagnostic platforms could facilitate its use in resource-limited settings, supporting large-scale surveillance efforts and outbreak management.

Bibliography

A. Z. Kapikian & Y. Hoshino, R. M. C. (2001). *Reoviridae* (H. Fields, Knipe (ed.); 4th ed.). Lippincott Williams & Wilkins.

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindah, E. (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001

Afchangi, A., Jalilvand, S., Mohajel, N., Marashi, S. M., & Shoja, Z. (2019). Rotavirus VP6 as a potential vaccine candidate. *Reviews in Medical Virology*, *29*(2), 29(2), e2027. https://doi.org/10.1002/rmv.2027

Afchangi, A., Arashkia, A., Shahosseini, Z., Jalilvand, S., Marashi, S. M., Roohvand, F., Mohajel, N., & Shoja, Z. (2017). Immunization of Mice by Rotavirus NSP4-VP6 Fusion Protein Elicited Stronger Responses Compared to VP6 Alone. *Viral Immunology*, *31*(3), 233–241. https://doi.org/10.1089/vim.2017.0075

Ahmed, A. B., & Kajava, A. V. (2013). Breaking the amyloidogenicity code : Methods to predict amyloids from amino acid sequence. *FEBS Letters*, *587*(8), 1089–1095.

Aliabadi, N., Antoni, S., Mwenda, J. M., Weldegebriel, G., Biey, J. N. M., Cheikh, D., Fahmy, K., Teleb, N., Ashmony, H. A., Ahmed, H., Daniels, D. S., Videbaek, D., Wasley, A., Singh, S., de Oliveira, L. H., Rey-Benito, G., Sanwogou, N. J., Wijesinghe, P. R., Liyanage, J. B. L. & Cohen, A. L. (2019). Global impact of rotavirus vaccine introduction on rotavirus hospitalisations among children under 5 years of age, 2008–16: findings from the Global Rotavirus Surveillance Network. *The Lancet Global Health*, *7*(7), e893–e903. https://doi.org/10.1016/S2214-109X(19)30207-4

Ami, D., Natalello, A., Gatti-Lafranconi, P., Lotti, M., & Doglia, S. M. (2005). Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy. *FEBS Letters*, *579*(16), 3433–3436. https://doi.org/10.1016/j.febslet.2005.04.085

Angel, J., Franco, M. A., & Greenberg, H. B. (2007). Rotavirus vaccines: Recent

developments and future considerations. *Nature Reviews Microbiology*, *5*(7), 529–539. https://doi.org/10.1038/nrmicro1692

Arai, R., Ueda, H., Kitayama, A., Kamiya, N., & Nagamune, T. (2001). Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Engineering*, *14*(8), 529–532. https://doi.org/10.1093/protein/14.8.529

Arias, C. F., Dector, M. A., Segovia, L., López, T., Camacho, M., Isa, P., Espinosa, R., & López, S. (2004). RNA silencing of rotavirus gene expression. *Virus Research*, *102*(1), 43–51. https://doi.org/10.1016/j.virusres.2004.01.014

Armstrong, R. E., & Strouse, G. F. (2014). Rationally manipulating aptamer binding affinities in a stem-loop molecular beacon. *Bioconjugate Chemistry*, *25*(10), 1769–1776. https://doi.org/10.1021/bc500286r

Artika, I. M., Dewi, Y. P., Nainggolan, I. M., Siregar, J. E., & Antonjaya, U. (2022). Real-Time Polymerase Chain Reaction: Current Techniques, Applications, and Role in COVID-19 Diagnosis. *Genes*, *13*(12), 1–22. https://doi.org/10.3390/genes13122387

Babaei, A., Rafiee, N., Taheri, B., Sohrabi, H., & Mokhtarzadeh, A. (2022). Recent Advances in Early Diagnosis of Viruses Associated with Gastroenteritis by Biosensors. *Biosensors*, *12*(7). https://doi.org/10.3390/bios12070499

Bachu, V., Barman, K., & Goswami, P. (2023). Analysis on the in-silico ensemble methods for 3D modelling of ssDNA aptamers. *Biophysical Chemistry*, *303*(September), 107111. https://doi.org/10.1016/j.bpc.2023.107111

Bányai, Krisztián, Gábor Kemenesi C, Ivana Budinski B, Fanni Földes B, Brigitta Zana A, Szilvia Marton A, Renáta Varga-Kugler B, Miklós Oldal B & Kornélia Kurucz B, F. J. (2017). Candidate New Rotavirus Species in Schreiber's Bats, Serbia. *Infection, Genetics and Evolution*, *48*, 19–26.

Belon, C. A., & Frick, D. N. (2008). Monitoring helicase activity with molecular beacons. *BioTechniques*, *45*(4), 433–442. https://doi.org/10.2144/000112834

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, *28*(1), 235–242.

Bertol, J. W., & Gatti, M. S. V. (2016). Rapid detection of infectious rotavirus group A using a molecular beacon assay. *Journal of Virological Methods*, *234*, 156–159. https://doi.org/10.1016/j.jviromet.2016.04.017

Besse, S., Poujol, R., & Hussin, J. G. (2022). Comparative Study of Protein Aggregation Propensity and Mutation Tolerance Between Naked Mole-Rat and Mouse. *Genome Biology and Evolution*, *14*(5).

Bidar, N., Amini, M., Oroojalian, F., Baradaran, B., Hosseini, S. S., Shahbazi, M. A., Hashemzaei, M., Mokhtarzadeh, A., Hamblin, M. R., & de la Guardia, M. (2021). Molecular beacon strategies for sensing purpose. *TrAC - Trends in Analytical Chemistry*, *134*, 116143. https://doi.org/10.1016/j.trac.2020.116143

Bishop, R. F., Barnes, G. L., Cipriani, E., & Lund, J. S. (1983). Clinical immunity after neonatal rotavirus infection. A prospective longitudinal study in young children. *The New England Journal of Medicine*, *309*(2), 72–76.

Bishop, R. F.; Davidson, G. P.; Holmes, I. H. & Ruck, B. J. (1973). Virus particles in epithelial cells of duodenal mucosa from children with acute non-bacterial gastroenteritis. *Lancet*, *302*(7841), 1281−1283.

Biswas, S., Devi, Y. D., Sarma, D., Hatiboruah, D., Chamuah, N., Namsa, N. D., & Nath, P. (2023). Detection and analysis of rotavirus in clinical stool samples using silver nanoparticle functionalized paper as SERS substrate. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, *295*(November 2022), 122610. https://doi.org/10.1016/j.saa.2023.122610

Blazevic, V., Lappalainen, S., Nurminen, K., Huhti, L., & Vesikari, T. (2011). Norovirus VLPs and Rotavirus VP6 Protein as Combined Vaccine for Childhood Gastroenteritis. *Vaccine*, *29*(45), 8126–8133. https://doi.org/10.1016/j.vaccine.2011.08.026

Bledzka K, Pesho MM & Ma YQ, P. E. (2012). *Integrin alpha IIb beta 3* (M. A (ed.)). Elsevier Science.

Bredell, H., Smith, J. J., Prins, W. A., Görgens, J. F., & van Zyl, W. H. (2016). Expression of Rotavirus VP6 Protein : a Comparison Amongst Escherichia coli , Pichia pastoris and Hansenula polymorpha. *FEMS Yeast Research*, *16*(2). https://doi.org/10.1093/femsyr/fow001

Buck, P. M., Kumar, S., Wang, X., Agrawal, N. J., Trout, B. L., & Singh, S. K. (n.d.). Computational Methods to Predict Therapeutic Protein Aggregation. In V. V. & C. J. A. (Ed.), *Therapeutic Proteins* (pp. 425–451), 2012. Springer. https://doi.org/10.1007/978-1-61779-921-1

Bugli, F., Caprettini, V., Cacaci, M., Martini, C., Paroni Sterbini, F., Torelli, R., Della Longa, S., Papi, M., Palmieri, V., Giardina, B., Posteraro, B., Sanguinetti, M., & Arcovito, A. (2014). Synthesis and Characterization of Different Immunogenic Viral Nanoconstructs from Rotavirus VP6 Inner Capsid Protein. *International Journal of Nanomedicine*, *9*, 2727–2739.

Bui, H. H., Sidney, J., Peters, B., Sathiamurthy, M., Sinichi, A., Purton, K. A., Mothé, B. R., Chisari, F. V., Watkins, D. I., & Sette, A. (2005). Automated generation and evaluation of specific MHC binding predictive tools: ARB matrix applications. *Immunogenetics*, *57*(5), 304–314. https://doi.org/10.1007/s00251-005-0798-y

Burke, R. M., Tate, J. E., Kirkwood, C. D., Steele, A. D., & Parashar, U. D. (2019). Current and New Rotavirus Vaccines. *Current Opinion in Infectious Diseases*, *32*(2), 435. https://doi.org/10.1097/QCO.0000000000000572

Byzova, T. V., Rabbani, R., D'Souza, S. E., & Plow, E. F. (1998). Role of integrin alpha(v)beta3 in vascular biology. *Thrombosis and Haemostasis*, *80*(5), 726–734.

Carvalho, M. F., & Gill, D. (2019). Rotavirus vaccine efficacy: current status and areas for improvement. *Human Vaccines and Immunotherapeutics*, *15*(6), 1237–1250. https://doi.org/10.1080/21645515.2018.1520583

Centre for Disease Control and Prevention (CDC). (1999). Intussusception among Recipients of Rotavirus Vaccine—United States, 1998-1999. *Mortality and Morbidity Weekly Reports*, *48*, 577–581.

Charpilienne, A., Lepault, J., Rey, F., & Cohen, J. (2002). Identification of Rotavirus VP6 Residues Located at the Interface with VP2 That Are Essential for Capsid Assembly and Transcriptase Activity. *Journal of Virology*, *76*(15), 7822–7831. https://doi.org/10.1128/JVI.76.15.7822

Chen, X., Zaro, J. L., & Shen, W. C. (2013). Fusion Protein Linkers: Property, Design and Functionality. *Advanced Drug Delivery Reviews*, *65*(10), 1357–1369.

Chen, M., Ma, Z., Wu, X., Mao, S., Yang, Y., Tan, J., Krueger, C. J., & Chen, A. K. (2017). A molecular beacon-based approach for live-cell imaging of RNA transcripts with minimal target engineering at the single-molecule level. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-01740-1

Chen, S. C., Jones, D. H., Fynan, E. F., Farrar, G. H., Clegg, J. C. S., & Greenberg, H. B. (1998). Protective Immunity Induced by Oral Immunization with a Rotavirus DNA Vaccine Encapsulated in Microparticles. *Journal of Virology*, *72*(7), 5757–5761.

Chiti, F., & Dobson, C. M. (2006). Protein Misfolding , Functional Amyloid , and Human Disease. *Annual Review of Biochemistry*, *75*, 333–366. https://doi.org/10.1146/annurev.biochem.75.101304.123901

Choi, A. H., Basu, M., Neal, M. M. M. C., Clements, J. D., & Ward, R. L. (1999). Antibody-Independent Protection against Rotavirus Infection of Mice Stimulated by Intranasal Immunization with Chimeric VP4 or VP6 Protein. *Journal of Virology*, *73*(9), 7574–7581.

Chung, I. S., Kim, C. H., Kim, K. Il, Hong, S. H., Park, J. H., Kim, J. K., & Kim, W. Y. (2000). Production of Recombinant Rotavirus VP6 from a Suspension Culture of Transgenic Tomato (Lycopersicon esculentum Mill.) Cells. *Biotechnology Letters*, *22*, 251–255.

Colovos, C., & Yeates, T. (1993). Verification of protein structures : Patterns of nonbonded atomic interactions. *Protein Science*, *2*(9), 1511–1519.

Comeau, S. R., Gatchell, D. W., Vajda, S., & Camacho, C. J. (2004). ClusPro: A fully automated algorithm for protein-protein docking. *Nucleic Acids Research*, *32*(WEB SERVER ISS.), 96–99. https://doi.org/10.1093/nar/gkh354

Conchillo-solé, O., Groot, N. S. De, Avilés, F. X., Vendrell, J., Daura, X., & Ventura, S. (2007). AGGRESCAN : a Server for the Prediction and Evaluation of " Hot Spots " of Aggregation in Polypeptides. *BMC Bioinformatics*, *8*(65). https://doi.org/10.1186/1471-2105-8-65

Crawford, S. E., Ramani, S., Tate, J. E., Parashar, U. D., Svensson, L., Hagbom, M., Franco, M. A., Greenberg, H. B., O'ryan, M., Kang, G., Desselberger, U., Estes, M. K., Troeger, C., Khalil, I. A., Rao, P. C., Cao, S., Blacker, B. F., Ahmed, T.,

Armah, G. & Bernstein, D. I. (2018). Rotavirus infection HHS Public Access. *JAMA Pediatrics*, *172*(SUPPL. 3), 50–53. https://doi.org/10.1038/nrdp.2017.83.Rotavirus

Crawford, S. E., Mukherjee, S. K., Estes, M. K., Lawton, J. A., Shaw, A. L., Ramig, R. F., & Prasad, B. V. V. (2001). Trypsin Cleavage Stabilizes the Rotavirus VP4 Spike. *Journal of Virology*, *75*(13), 6052–6061. https://doi.org/10.1128/JVI.75.13.6052

Crooks, G., Hon, G., Chandonia, J., & Brenner, S. (2004). NCBI GenBank FTP Site\nWebLogo: a sequence logo generator. *Genome Research*, *14*, 1188–1190. https://doi.org/10.1101/gr.849004.1

D Eisenberg & R Lüthy, J. U. B. (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods in Enzymology*, *277*, 396–404.

da Silva Junior, H. C., da Silva E Mouta Junior, S., de Mendonça, M. C., de Souza Pereira, M. C., da Rocha Nogueira, A., de Azevedo, M. L., Leite, J. P., & de Moraes, M. T. (2012). Comparison of two Eukaryotic Systems for the Expression of VP6 Protein of Rotavirus Species A : Transient Gene Expression in HEK293-T Cells and Insect Cell-Baculovirus System. *Biotechnology Letters*, *34*(9), 1623–1627. https://doi.org/10.1007/s10529-012-0946-z

Davis, M. W., & Jorgensen, E. M. (2022). ApE, A Plasmid Editor: A Freely Available DNA Manipulation and Visualization Program. *Frontiers in Bioinformatics*, *2*(February), 1–15. https://doi.org/10.3389/fbinf.2022.818619

Desselberger, U. (2014). Rotaviruses. *Virus Research*, *190*, 75–96. https://doi.org/10.1016/j.virusres.2014.06.016

Desselberger, U. (2017). Differences of rotavirus vaccine effectiveness by country: Likely causes and contributing factors. *Pathogens*, *6*(4), 1–13. https://doi.org/10.3390/pathogens6040065

Devi, Y. D., Devi, A., Gogoi, H., Dehingia, B., Doley, R., Buragohain, A. K., Singh, C. S., Borah, P. P., Rao, C. D., Ray, P., Varghese, G. M., Kumar, S., & Namsa, N. D. (2020). Exploring rotavirus proteome to identify potential B- and T-cell epitope using computational immunoinformatics. *Heliyon*, *6*(12), e05760. https://doi.org/10.1016/j.heliyon.2020.e05760

Dimitrov, I., Bangov, I., Flower, D. R., & Doytchinova, I. (2014). AllerTOP v.2 - A server for in silico prediction of allergens. *Journal of Molecular Modeling*, *20*(6). https://doi.org/10.1007/s00894-014-2278-5

Dimitrov, I., Naneva, L., Doytchinova, I., & Bangov, I. (2014). AllergenFP: Allergenicity prediction by descriptor fingerprints. *Bioinformatics*, *30*(6), 846–851. https://doi.org/10.1093/bioinformatics/btt619

Doytchinova, I. A., & Flower, D. R. (2007). VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, *8*, 1–7. https://doi.org/10.1186/1471-2105-8-4

Duan, T., Du, Y., Xing, C., Wang, H. Y., & Wang, R. F. (2022). Toll-Like Receptor Signaling and Its Role in Cell-Mediated Immunity. *Frontiers in Immunology*, *13*(March), 1–22. https://doi.org/10.3389/fimmu.2022.812774

Dutta, K. (2020). Multi valent DNA vaccine against group A human rotavirus: an in-silico investigation. *BioRxiv*, 2020.01.13.903781. https://www.biorxiv.org/content/10.1101/2020.01.13.903781v2%0Ahttps://www.biorxiv.org/content/10.1101/2020.01.13.903781v2.abstract

Elskens, J., & Madder, A. (2021). Crosslinker-modified nucleic acid probes for improved target identification and biomarker detection. *RSC Chemical Biology*, *2*(2), 410–422. https://doi.org/10.1039/d0cb00236d

Estes, M. K. & Greenberg, H. B. (2013). *Rotaviruses, In Field's Virology* (P. M. Knipe, D. M., Howley (Ed.); 6th ed.). Lippincott: Williams & Wilkins.

Estes, M. (1996). Rotaviruses and their Replication. In P. M. Fields, B. N., Knipe, D. M., Howley (Ed.), *Lippincott Raven* (3rd ed., Vol. 2).

Estes, M. K. (2007). *Rotaviruses* (D. M. Knipe (Ed.); 5th ed.). Lippincott Williams & Wilkins, Philadelphia.

Estes, M. K., Crawford, S. E., Penaranda, M. E., Petrie, B. L., Burns, J. W., Chan, W., Ericson, B., Smith, G. E., & Summers, M. A. X. D. (1987). Synthesis and Immunogenicity of the Rotavirus Major Capsid Antigen Using a Baculovirus Expression System. *Journal of Virology*, *61*(5), 1488–1494.

Eszter Mihalov-Kovács, Á. G., Szilvia Marton, Szilvia L. Farkas, E. F., Miklós Oldal,

Ferenc Jakab, V. M., & Bányai, K. (2015). Candidate New Rotavirus Species in Sheltered Dogs, Hungary. *Emerging Infectious Diseases*, *21*(4), 660–663.

Feng, H., Li, X., Song, W., Duan, M., Chen, H., & Wang, T. (2017). Oral Administration of a Seed-based Bivalent Rotavirus Vaccine Containing VP6 and NSP4 Induces Specific Immune Responses in Mice. *Frontiers in Plant Science*, *8*, 910. https://doi.org/10.3389/fpls.2017.00910

Fischer Walker, C. L., & Black, R. E. (2010). Zinc for the treatment of diarrhoea: Effect on diarrhoea morbidity, mortality and incidence of future episodes. *International Journal of Epidemiology*, *39*(SUPPL. 1), 63–69. https://doi.org/10.1093/ije/dyq023

Flewett, T. H.; Bryden, A. S. & Davies, H. (1973). Virus particles in gastroenteritis. *Lancet*, *302*(7844), 1497.

Frolova, L., Drastich, P., Rossmann, P., Klimesova, K., & Tlaskalova-Hogenova, H. (2008). Expression of Toll-like receptor 2 (TLR2), TLR4, and CD14 in biopsy samples of patients with inflammatory bowel diseases: Upregulated expression of TLR2 in terminal ileum of patients with ulcerative colitis. *Journal of Histochemistry and Cytochemistry*, *56*(3), 267–274. https://doi.org/10.1369/jhc.7A7303.2007

Garbuzynskiy, S. O., Lobanov, M. Y., & Galzitskaya, O. V. (2009). FoldAmyloid: A method of prediction of amyloidogenic regions from protein sequence. *Bioinformatics*, *26*(3), 326–332. https://doi.org/10.1093/bioinformatics/btp691

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the Expasy Server. In J. M. Walker (Ed.), *The Proteomics Protocols Handbook*. Humana Press. https://doi.org/10.1385/1592598900

Gautam, R., Lyde, F., Esona, M. D., Quaye, O., Michael, D., & Viruses, R. (2015). *Detection of Rotavirus Antigen in Stool Specimens*. *58*(1), 292–294. https://doi.org/10.1016/j.jcv.2013.06.022.Comparison

Gentsch, J. R., Laird, A. R., Bielfelt, B., Griffin, D. D., Bányai, K., Ramachandran, M., Jain, V., Cunliffe, N. A., Nakagomi, O., Kirkwood, C. D., Fischer, T. K., Parashar, U. D., Bresee, J. S., Jiang, B., & Glass, R. I. (2005). Serotype diversity and

reassortment between human and animal rotavirus strains: Implications for rotavirus vaccine programs. *Journal of Infectious Diseases*, *192*(SUPPL. 1). https://doi.org/10.1086/431499

Glass, R. I., Parashar, U. D., & Estes, M. K. (2009). Norovirus Gastroenteritis. *New England Journal of Medicine*, *361*(18), 1776–1785. https://doi.org/10.1056/nejmra0804575

Gour, S., & Yadav, J. K. (2021). Aggregation hot spots in the SARS-CoV-2 proteome may constitute potential therapeutic targets for the suppression of the viral replication and multiplication. *Journal of Proteins and Proteomics*, *12*(1), 1–13. https://doi.org/10.1007/s42485-021-00057-y

Grote, A., Hiller, K., Scheer, M., Münch, R., Nörtemann, B., Hempel, D. C., & Jahn, D. (2005). JCat: A novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Research*, *33*(SUPPL. 2), 526–531. https://doi.org/10.1093/nar/gki376

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R., & Hofacker, I. L. (2008). The Vienna RNA websuite. *Nucleic Acids Research*, *36*(Web Server issue), 70–74. https://doi.org/10.1093/nar/gkn188

Guerrero, C. A., Méndez, E., Zárate, S., Isa, P., López, S., & Arias, C. F. (2000). Integrin αvβ3 mediates rotavirus cell entry. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(26), 14644–14649. https://doi.org/10.1073/pnas.250299897

Guliy, O. I., Zaitsev, B. D., Larionova, O. S., & Borodina, I. A. (2019). Virus Detection Methods and Biosensor Technologies. *Biophysics (Russian Federation)*, *64*(6), 890–897. https://doi.org/10.1134/S0006350919060095

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., & Raghava, G. P. S. (2013). In Silico Approach for Predicting Toxicity of Peptides and Proteins. *PLoS ONE*, *8*(9). https://doi.org/10.1371/journal.pone.0073957

Hall, T. A. (1999). BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, *41*, 95–98.

Han, S. X., Jia, X., Ma, J. L., & Zhu, Q. (2013). Molecular beacons: A novel optical diagnostic tool. *Archivum Immunologiae et Therapiae Experimentalis*, *61*(2), 139–148. https://doi.org/10.1007/s00005-012-0209-7

Hatcher, E. L., Zhdanov, S. A., Bao, Y., Blinkova, O., Nawrocki, E. P., Ostapchuck, Y., Schaffer, A. A., & Rodney Brister, J. (2017). Virus Variation Resource-improved response to emergent viral outbreaks. *Nucleic Acids Research*, *45*(D1), D482–D490. https://doi.org/10.1093/nar/gkw1065

Hebditch, M., Carballo-Amador, M. A., Charonis, S., Curtis, R., & Warwicker, J. (2017). Protein-Sol: A web tool for predicting protein solubility from sequence. *Bioinformatics*, *33*(19), 3098–3100. https://doi.org/10.1093/bioinformatics/btx345

Heo, L., Park, H., & Seok, C. (2013). GalaxyRefine: Protein structure refinement driven by side-chain repacking. *Nucleic Acids Research*, *41*(Web Server issue), 384–388. https://doi.org/10.1093/nar/gkt458

Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A Linear Constraint Solver for molecular simulations. *Journal of Computational Chemistry*, *18*(12), 1463–1472. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H

Hoshino, Y., & Kapikian, A. Z. (2000). Rotavirus serotypes: classification and importance in epidemiology, immunity, and vaccine development. *Journal of Health, Population, and Nutrition*, *18*(1), 5–14.

Hua, Y., Ma, J., Li, D., & Wang, R. (2022). DNA-Based Biosensors for the Biochemical Analysis: A Review. *Biosensors*, *12*(3). https://doi.org/10.3390/bios12030183

Hynes, R. O. (1987). Integrins: A family of cell surface receptors. *Cell*, *48*, 549–554. https://doi.org/10.1016/0092-8674(87)90233-9

International Vaccine Access Center (IVAC). Johns Hopkins Bloomberg School of Public Health. (n.d.). Retrieved September 21, 2022, from www.view-hub.org

Iturriza-Gómara, M., Dallman, T., Bányai, K., Böttiger, B., Buesa, J., Diedrich, S., Fiore, L., Johansen, K., Koopmans, M., Korsun, N., Koukou, D., Kroneman, A., László, B., Lappalainen, M., Maunula, L., Marques, A. M., Matthijnssens, J.,

Midgley, S., Mladenova, Z., & Gray, J. (2011). Rotavirus genotypes co-circulating in Europe between 2006 and 2009 as determined by EuroRotaNet, a pan-European collaborative strain surveillance network. *Epidemiology and Infection*, *139*(6), 895–909. https://doi.org/10.1017/S0950268810001810

Jafarpour, S., Ayat, H., & Ahadi, A. M. (2015). Design and Antigenic Epitopes Prediction of a New Trial Recombinant Multiepitopic Rotaviral Vaccine: In Silico Analyses. *Viral Immunology*, *28*(6), 325–330. https://doi.org/10.1089/vim.2014.0152

Jahn, T. R., & Radford, S. E. (2005). The Yin and Yang of protein folding. *Teh FEBS Journal*, *272*, 5962–5970. https://doi.org/10.1111/j.1742-4658.2005.05021.x

Jalilvand, S., Mahdi, S., & Shoja, Z. (2015). Rotavirus VP6 Preparations as a Non-Replicating Vaccine Candidates. *Vaccine*, *33*(29), 3281–3287. https://doi.org/10.1016/j.vaccine.2015.05.026

Janssen, K. P. F., Knez, K., Spasic, D., & Lammertyn, J. (2013). Nucleic acids for ultra-sensitive protein detection. In *Sensors (Switzerland)* (Vol. 13, Issue 1). https://doi.org/10.3390/s130101353

Jayaram, H., Estes, M. K., & Prasad, B. V. V. (2004). Emerging Themes in Rotavirus Cell Entry , Genome Organization , Transcription and Replication. *Virus Research*, *101*(1), 67–81. https://doi.org/10.1016/j.virusres.2003.12.007

Jelle Matthijnssens, Peter H Otto, Max Ciarlet, Ulrich Desselberger, & Marc Van Ranst, R. J. (2012). VP6-Sequence-Based Cutoff Values as a Criterion for Rotavirus Species Demarcation. *Archives of Virology*, *157*(6), 1177–1182. https://doi.org/10.1007/s00705-012-1273-3

Jiang, B., Wang, Y., & Glass, R. I. (2013). Does a monovalent inactivated human rotavirus vaccine induce heterotypic immunity? Evidence from animal studies. *Human Vaccines and Immunotherapeutics*, *9*(8), 1634–1637. https://doi.org/10.4161/hv.24958

Jiménez-Zaragoza, M., Yubero, M. P. L., Martín-Forero, E., Castón, J. R., Reguera, D., Luque, D., de Pablo, P. J., & Rodríguez, J. M. (2018). Biophysical properties of single rotavirus particles account for the functions of protein shells in a multilayered virus. *ELife*, *7*, 1–23. https://doi.org/10.7554/eLife.37295

Jonesteller, C. L., Burnett, E., Yen, C., Tate, J. E., & Parashar, U. D. (2017). *Effectiveness of Rotavirus Vaccination : A Systematic Review of the First Decade of Global Postlicensure Data ,. 65*, 2006–2016. https://doi.org/10.1093/cid/cix369

Joung, I. S., & Cheatham, T. E. (2008). Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *Journal of Physical Chemistry B*, *112*(30), 9020–9041. https://doi.org/10.1021/jp8001614

Judith A.V. & David A.C. (1996). Integrins and cancer. *Current Opinion in Cell Biology*, *8*, 724–730.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Ready K.F. & Sabara. M. (1987). In vitro Assembly of Bovine Rotavirus Nucleocapsid Protein. *Virology*, *157*(1), 189–198.

Kaplon, J., Fremy, C., Pillet, S., Martins, M., Ambert-balay, K., & Aho, S. L. (2015). Diagnostic Accuracy of Seven Commercial Assays for Rapid Detection of Group A Rotavirus Antigens. *Journal of Clinical Microbiology*, *53*(11), 3670–3673. https://doi.org/10.1128/JCM.01984-15.Editor

Kato, T., Kakuta, T., Yonezuka, A., Sekiguchi, T., Machida, Y., & Xu, J. (2023). Expression and Purification of Porcine Rotavirus Structural Proteins in Silkworm Larvae as a Vaccine Candidate. *Molecular Biotechnology*, *65*(3), 401–409. https://doi.org/10.1007/s12033-022-00548-3

Kim, J., Yang, Y. L., Jang, S. H., & Jang, Y. S. (2018). Human β-defensin 2 plays a regulatory role in innate antiviral immunity and is capable of potentiating the induction of antigen-specific immunity. *Virology Journal*, *15*(1), 1–12. https://doi.org/10.1186/s12985-018-1035-2

King, A. M. Q., Adams, M. J., Carstens, E. B., & Lefkowitz, E. J. (2012). *Virus Taxonomy Classification and Nomenclature of Viruses Ninth Report of the International Committee on Taxonomy of Viruses*. www.macmillansolutions.com

Kirienko, N. V., Lepikhov, K. A., Zheleznaya, L. A., & Matvienko, N. I. (2004). Significance of codon usage and irregularities of rare codon distribution in genes for expression of BspLU11III methyltransferases. *Biokhimiya*, *69*(5), 647–657.

Kirkwood, C. D., Ma, L., Carey, M. E., & Steele, A. D. (2019). The Rotavirus Vaccine Development Pipeline. *Vaccine*, *37*(50), 7328–7335. https://doi.org/10.1016/j.vaccine.2017.03.076

Krusiński, T., Ozyhar, A., & Dobryszycki, P. (2010). Dual FRET assay for detecting receptor protein interaction with DNA. *Nucleic Acids Research*, *38*(9). https://doi.org/10.1093/nar/gkq049

Kulkarni, P. S., Desai, S., Tewari, T., Kawade, A., Goyal, N., Garg, B. S., Kumar, D., Kanungo, S., Kamat, V., Kang, G., Bavdekar, A., Babji, S., Juvekar, S., Manna, B., Dutta, S., Angurana, R., Dewan, D., Dharmadhikari, A., Zade, J. K., & Raj, C. V. (2017). A randomized Phase III clinical trial to assess the efficacy of a bovine-human reassortant pentavalent rotavirus vaccine in Indian infants. *Vaccine*, *35*(45), 6228–6237. https://doi.org/10.1016/j.vaccine.2017.09.014

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. https://doi.org/10.1093/MOLBEV/MSW054

Kuri, P. R., Das, P., & Goswami, P. (2020). Fundamentals of Biosensors. *Advanced Materials and Techniques for Biosensors and Bioanalytical Applications*, 1–28. https://doi.org/10.1201/9781003083856-1

Kuri, P. R., & Goswami, P. (2023). Current Update on Rotavirus in-Silico Multiepitope Vaccine Design. *ACS Omega*, *8*(1), 190–207. https://doi.org/10.1021/acsomega.2c07213

Lappalainen, S., Pastor, A. R., Malm, M., López-Guerrero, V., Esquivel-Guadarrama, F., Palomares, L. A., Vesikari, T., & Blazevic, V. (2015). Protection Against Live Rotavirus Challenge in Mice Induced by Parenteral and Mucosal Delivery of VP6 Subunit Rotavirus Vaccine. *Archives of Virology*, *160*(8), 2075–2078. https://doi.org/10.1007/s00705-015-2461-8

Lappalainen, S., Vesikari, T., & Blazevic, V. (2016). Simple and Efficient Ultrafiltration Method for Purification of Rotavirus VP6 Oligomeric Proteins.

Archives of Virology, *161*(11), 3219–3223. https://doi.org/10.1007/s00705-016-2991-8

Larsen, M. V., Lundegaard, C., Lamberth, K., Buus, S., Lund, O., & Nielsen, M. (2007). Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics*, *8*, 1–12. https://doi.org/10.1186/1471-2105-8-424

Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S., & Thornton, J. M. (2018). PDBsum: Structural summaries of PDB entries. *Protein Science*, *27*(1), 129–134. https://doi.org/10.1002/pro.3289

Lee, B., & Lee, B. (2021). Update on Rotavirus Vaccine Underperformance in Low-To Middle-Income Countries and Next-Generation Vaccines. *Human Vaccines & Immunotherapeutics*, *17*(6), 1787–1802. https://doi.org/10.1080/21645515.2020.1844525

Lepault, J., Petitpas, I., Erk, I., Navaza, J., Bigot, D., Dona, M., Vachette, P., Cohen, J., & Rey, F. A. (2001). Structural Polymorphism of the Major Capsid Protein of Rotavirus. *The EMBO Journal*, *20*(7), 1498–1507.

Lester, S. N., & Li, K. (2014). Toll-like receptors in antiviral innate immunity. *Journal of Molecular Biology*, *426*(6), 1246–1264. https://doi.org/10.1016/j.jmb.2013.11.024

Leung, A. K. C., & Barankin, B. (2006). Oral rehydration therapy and early refeeding in the management of childhood gastroenteritis. *Paediatrics and Child Health*, *11*(8), 1–6. https://doi.org/10.1093/pch/11.8.527

Li, X., Guo, L., Kong, M., Su, X., Yang, D., Zou, M., Liu, Y., & Lu, L. (2016). Design and Evaluation of a Multi-Epitope Peptide of Human Metapneumovirus. *Intervirology*, *58*(6), 403–412. https://doi.org/10.1159/000445059

Liu, Z., Wang, F., & Chen, X. (2008). Integrin alphaV-beta3-targeted cancer therapy. *Drug Development Research*, *69*(6), 329–339. https://doi.org/10.1002/ddr.20265.Integrin

López-Blanco, J. R., Aliaga, J. I., Quintana-Ortí, E. S., & Chacón, P. (2014). IMODS: Internal coordinates normal mode analysis server. *Nucleic Acids Research*, *42*(W1), 271–276. https://doi.org/10.1093/nar/gku339

Lopéz-Blanco, J. R., Garzón, J. I., & Chacón, P. (2011). iMod: Multipurpose normal mode analysis in internal coordinates. *Bioinformatics*, *27*(20), 2843–2850. https://doi.org/10.1093/bioinformatics/btr497

Ma, L., & Liu, J. (2020). Catalytic Nucleic Acids: Biochemistry, Chemical Biology, Biosensors, and Nanotechnology. *iScience*, *23*(1), 100815. https://doi.org/10.1016/j.isci.2019.100815

Mahapatro, S., Mahilary, N., Satapathy, A. K., & Das, R. R. (2017). Nitazoxanide in Acute Rotavirus Diarrhea: A Randomized Control Trial from a Developing Country. *Journal of Tropical Medicine*, *2017*. https://doi.org/10.1155/2017/7942515

Malik, Y. S., Bhat, S., Dar, P. S., Sircar, S., Dhama, K., & Singh, R. K. (2020). Evolving Rotaviruses, Interspecies Transmission and Zoonoses. *The Open Virology Journal*, *14*(1), 1–6. https://doi.org/10.2174/1874357902014010001

Manouana, G. P., Niendorf, S., Tomazatos, A., Mbong Ngwese, M., Nzamba Maloum, M., Nguema Moure, P. A., Bingoulou Matsougou, G., Ategbo, S., Rossatanga, E. G., Bock, C. T., Borrmann, S., Mordmüller, B., Eibach, D., Kremsner, P. G., Velavan, T. P., & Adegnika, A. A. (2021). Molecular surveillance and genetic divergence of rotavirus A antigenic epitopes in Gabonese children with acute gastroenteritis. *EBioMedicine*, *73*. https://doi.org/10.1016/j.ebiom.2021.103648

Martella, V., Bányai, K., Matthijnssens, J., Buonavoglia, C., & Ciarlet, M. (2010). Zoonotic aspects of rotaviruses. *Veterinary Microbiology*, *140*(3–4), 246–255. https://doi.org/10.1016/j.vetmic.2009.08.028

Matthijnssens, J., Bilcke, J., Ciarlet, M., Martella, V., Bányai, K., Rahman, M., Zeller, M., Beutels, P., Van Damme, P., & Van Ranst, M. (2009). Rotavirus disease and vaccination: impact on genotype diversity. *Future Microbiology*, *4*(10), 1303–1316.

Matthijnssens, J., Ciarlet, M., Heiman, E., Arijs, I., Delbeke, T., McDonald, S. M., Palombo, E. A., Iturriza-Gómara, M., Maes, P., Patton, J. T., Rahman, M., & Van Ranst, M. (2008). Full Genome-Based Classification of Rotaviruses Reveals a Common Origin between Human Wa-Like and Porcine Rotavirus Strains and Human DS-1-Like and Bovine Rotavirus Strains. *Journal of Virology*, *82*(7),

3204–3219. https://doi.org/10.1128/jvi.02257-07

Matthijnssens, J., Heylen, E., Zeller, M., Rahman, M., Lemey, P., & Van Ranst, M. (2010). Phylodynamic analyses of rotavirus genotypes G9 and G12 underscore their potential for swift global spread. *Molecular Biology and Evolution*, *27*(10), 2431–2436. https://doi.org/10.1093/molbev/msq137

McClain, B., Settembre, E., Temple, B. R. S., Bellamy, A. R., & Harrison, S. C. (2010). X-ray Crystal Structure of the Rotavirus Inner Capsid Particle at 3.8 Å Resolution. *Journal of Molecular Biology*, *397*(2), 587–599. https://doi.org/10.1016/j.jmb.2010.01.055

McGinnis, S., & Madden, T. L. (2004). BLAST: At the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research*, *32*(WEB SERVER ISS.), 20–25. https://doi.org/10.1093/nar/gkh435

McGuffin, L. J., Bryson, K., & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics*, *16*(4), 404–405. https://doi.org/10.1093/bioinformatics/16.4.404

Md. Sadikur Rahman Shuvo, S. K. M. and F. A. (2018). In Silico screening of T-cell and B-cell Epitopes of Rotavirus VP7 and VP4 proteins for Effective Vaccine Design. *Bangladesh Journal of Microbiology*, *35*(1), 45–55.

Meric, G., Robinson, A. S., & Roberts, C. J. (2017). Driving Forces for Nonnative Protein Aggregation and Approaches to Predict Aggregation-Prone Regions. *Annual Review of Chemical and Biomolecular Engineering*, *8*, 139–159.

Mili, M., Bachu, V., Kuri, P. R., Singh, N. K., & Goswami, P. (2024). Improving synthesis and binding affinities of nucleic acid aptamers and their therapeutics and diagnostic applications. *Biophysical Chemistry*, *309*(March), 107218. https://doi.org/10.1016/j.bpc.2024.107218

Moyle, P. M., & Toth, I. (2013). Modern Subunit Vaccines: Development, Components, and Research Opportunities. *ChemMedChem*, *8*(3), 360–376. https://doi.org/10.1002/cmdc.201200487

Naghavi, M., Abajobir, A. A., Abbafati, C., Abbas, K. M., Abd-Allah, F., Abera, S. F., Aboyans, V., Adetokunboh, O., Ärnlöv, J., Afshin, A., Agrawal, A., Kiadaliri, A.

A., Ahmadi, A., Ahmed, M. B., Aichour, A. N., Aichour, I., Aichour, M. T. E., Aiyar, S., Al-Eyadhy, A., … Murray, C. J. L. (2017). Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, *390*(10100), 1151–1210. https://doi.org/10.1016/S0140-6736(17)32152-9

Nanobiotechnol, J., Li, Z., Cui, K., Wang, H., Liu, F., Huang, K., Duan, Z., & Wang, F. (2019). A Milk - Based Self - Assemble Rotavirus VP6 – Ferritin Nanoparticle Vaccine Elicited Protection Against the Viral Infection. *Journal of Nanobiotechnology*, 1–13. https://doi.org/10.1186/s12951-019-0446-6

Narayanan, S., Parren, M., Ventura, S., Chiti, F., Giannoni, E., Dobson, C. M., Aviles, F. X., & Serrano, L. (2004). Short Amino Acid Stretches Can Mediate Amyloid Formation in Globular Proteins : The Src Homology 3 (SH3) Case. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 7258–7263. https://doi.org/10.1073/pnas.0308249101

Nielsen, M., Lundegaard, C., & Lund, O. (2007). Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinformatics*, *8*, 1–12. https://doi.org/10.1186/1471-2105-8-238

O'Brien GJ, Bryant CJ, Voogd C, Greenberg HB, Gardner RC, B. A. (2000). Rotavirus VP6 Expressed by PVX Vectors in Nicotiana benthamiana Coats PVX Rods and Also Assembles into Viruslike Particles. *Virology*, *270*(2), 444–453. https://doi.org/10.1006/viro.2000.0314

O'Mahony, D. S., Pham, U., Iyer, R., Hawn, T. R., & Liles, W. C. (2008). Differential constitutive and cytokine-modulated expression of human Toll-like receptors in primary neutrophils, monocytes, and macrophages. *International Journal of Medical Sciences*, *5*(1), 1–8. https://doi.org/10.7150/ijms.5.1

O'Ryan, M. (2017). Rotavirus Vaccines: A story of success with challenges ahead. *F1000Research*, *6*, 1–9. https://doi.org/10.12688/f1000research.11912.1

Oliveberg, M. (2010). Waltz , an exciting new move in amyloid prediction aMAZe-ing tools for mosaic analysis in zebrafish. *Nature Methods*, *7*(3), 187–188. https://doi.org/10.1038/nmeth0310-187

Omatola C A, & Odaibo G. (2016). Rotavirus Infection Among Under Five Children

Presenting with Gastroenteritis in Ibadan, Nigeria. *Arch. Bas. App. Med*, *3*, 3–8. www.archivesbamui.com

Ong, H., & Chandran, V. (2005). Identification of gastroenteric viruses by electron microscopy using higher order spectral features. *Journal of Clinical Virology*, *34*(3), 195–206. https://doi.org/10.1016/j.jcv.2005.04.001

Organizations, W. H., Record, W. E., States, M., Group, S. A., & Africa, S. (2013). Rotavirus vaccines WHO position paper: January 2013 - Recommendations. *Vaccine*, *31*(52), 6170–6171. https://doi.org/10.1016/j.vaccine.2013.05.037

Palumbo, E., Zhao, B., Xue, B., Uversky, V. N., & Davé, V. (2020). Analyzing aggregation propensities of clinically relevant PTEN mutants: a new culprit in pathogenesis of cancer and other PTENopathies. *Journal of Biomolecular Structure and Dynamics*, *38*(8), 2253–2266. https://doi.org/10.1080/07391102.2019.1630005

Papa, G., Borodavka, A., & Desselberger, U. (2021). Viroplasms: Assembly and functions of rotavirus replication factories. *Viruses*, *13*(7), 1–11. https://doi.org/10.3390/v13071349

Park, M. S., Kim, J. Il, Lee, I., Park, S., Bae, J. Y., & Park, M. S. (2018). Towards the application of human defensins as antivirals. *Biomolecules and Therapeutics*, *26*(3), 242–254. https://doi.org/10.4062/biomolther.2017.172

Parker, E. P. K., Ramani, S., Lopman, B. A., Church, J. A., Iturriza-Gómara, M., Prendergast, A. J., & Grassly, N. C. (2018). Causes of impaired oral vaccine efficacy in developing countries. *Future Microbiology*, *13*(1), 97–118. https://doi.org/10.2217/fmb-2017-0128

Pasare, C., & Medzhitov, R. (2005). Toll-like receptors: linking innate and adaptive immunity. In *Advances in experimental medicine and biology* (pp. 560, 11–18).

Peruski, A. H., & Peruski, L. F. (2003). Immunological methods for detection and identification of infectious disease and biological warfare agents. *Clinical and Diagnostic Laboratory Immunology*, *10*(4), 506–513. https://doi.org/10.1128/CDLI.10.4.506-513.2003

Peternel, Š., Grdadolnik, J., Gaberc-Porekar, V., & Komel, R. (2008). Engineering

inclusion bodies for non denaturing extraction of functional proteins. *Microbial Cell Factories*, *7*, 1–9. https://doi.org/10.1186/1475-2859-7-34

Prabakaran, R., Rawat, P., Kumar, S., & Michael Gromiha, M. (2021). ANuPP: A Versatile Tool to Predict Aggregation Nucleating Regions in Peptides and Proteins. *Journal of Molecular Biology*, *433*(11), 166707. https://doi.org/10.1016/j.jmb.2020.11.006

R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. T. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(283–291).

Radford, T. R. J. and S. E. (2008). Folding Versus Aggregation: Polypeptide Conformations on Competing Pathways. *Archives of Biochemistry and Biophysics*, *469*(1), 100–117. https://doi.org/10.1016/j.abb.2007.05.015

Rapin, N., Lund, O., Bernaschi, M., & Castiglione, F. (2010). Computational immunology meets bioinformatics: The use of prediction tools for molecular binding in the simulation of the immune system. *PLoS ONE*, *5*(4). https://doi.org/10.1371/journal.pone.0009862

Roberto, C., Muñoz, C., Buesa, J., & Gozalbo-rovira, R. (2021). The Rotavirus Vaccine Landscape , an Update. *Pathogens*, *10*(5), 520.

Rojas, M., Dias, H. G., Gonçalves, J. L. S., Manchego, A., Rosadio, R., Pezo, D., & Santos, N. (2019). Genetic diversity and zoonotic potential of rotavirus A strains in the southern Andean highlands, Peru. *Transboundary and Emerging Diseases*, *66*(4), 1718–1726. https://doi.org/10.1111/tbed.13207

Rotavirus: Current Vaccine Intro Status. https://View-Hub.Org/Vaccine/Rota.

Rotavirus Classification Working Group: *RCWG*. (n.d.). https://rega.kuleuven.be/cev/viralmetagenomics/virus-classification/rcwg, 2017

Rousseau, F., Schymkowitz, J., & Serrano, L. (2004). Prediction of Sequence-Dependent and Mutational Effects on the Aggregation of Peptides and Proteins. *Nature Biotechnology*, *22*(10), 1302–1306. https://doi.org/10.1038/nbt1012

Rovida, F., Campanini, G., Sarasini, A., Adzasehoun, K. M. G., Piralla, A., & Baldanti, F. (2013). Comparison of immunologic and molecular assays for the diagnosis of

gastrointestinal viral infections. *Diagnostic Microbiology and Infectious Disease*, *75*(1), 110–111. https://doi.org/10.1016/j.diagmicrobio.2012.09.016

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: A unified platform for automated protein structure and function prediction. *Nature Protocols*, *5*(4), 725–738. https://doi.org/10.1038/nprot.2010.5

Saha.S and Raghava G.P.S. (2004). BcePred:Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-chemical Properties. *In G.Nicosia, V.Cutello, P.J. Bentley and J.Timis (Eds.) ICARIS, LNCS 3239, Springer*, 197–204.

Sambrook J, R. D. (2001). *Molecular Cloning a Laboratory Manual.* (3rd ed.). Cold Spring Harbor Laboratory Press.

Santos, N.; Hoshino, Y. (2005). Global distribution of rotavirus serotypes/genotypes and its implication for the development and implementation of an effective rotavirus vaccine. *Reviews in Medical Virology*, *15*(1), 29−56.

Sathiamoorthy, S., & Shin, J. A. (2012). Boundaries of the Origin of Replication: Creation of a pET-28a-Derived Vector with p15A Copy Control Allowing Compatible Coexistence with pET Vectors. *PLoS ONE*, *7*(10). https://doi.org/10.1371/journal.pone.0047259

Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y., & Sherry, S. T. (2022). Database resources of the national center for biotechnology information. *Nucleic Acids Research*, *50*(D1), D20–D26. https://doi.org/10.1093/nar/gkab1112

Schr€odinger, L. (2015). *The PyMOL Molecular Graphics System. Version 2.3.3.*

Schwartz-cornil, I., Benureau, Y., Greenberg, H., Hendrickson, B. A., & Cohen, J. (2002). Heterologous Protection Induced by the Inner Capsid Proteins of Rotavirus Requires Transcytosis of Mucosal Immunoglobulins. *Journal of Virology*, *76*(16), 8110–8117. https://doi.org/10.1128/JVI.76.16.8110

Settembre, E. C., Chen, J. Z., Dormitzer, P. R., Grigorieff, N., & Harrison, S. C. (2010).

Atomic Model of an Infectious Rotavirus Particle. *The EMBO Journal*, *30*(2), 408–416. https://doi.org/10.1038/emboj.2010.322

Shahbazi Dastjerdeh, M., Shokrgozar, M. A., Rahimi, H., & Golkar, M. (2022). Potential aggregation hot spots in recombinant human keratinocyte growth factor: a computational study. *Journal of Biomolecular Structure and Dynamics*, *40*(18), 8169–8184. https://doi.org/10.1080/07391102.2021.1908912

Sharma, A. Das, Grewal, R. K., Gorle, S., Cuspoca, A. F., Kaushik, V., Rajjak Shaikh, A., Cavallo, L., & Chawla, M. (2023). T cell epitope based vaccine design while targeting outer capsid proteins of rotavirus strains infecting neonates: an immunoinformatics approach. *Journal of Biomolecular Structure and Dynamics*, *0*(0), 1–19. https://doi.org/10.1080/07391102.2023.2226721

Sharma, S., Byrne, H., & O'Kennedy, R. J. (2016). Antibodies and antibody-derived analytical biosensors. *Essays in Biochemistry*, *60*(1), 9–18. https://doi.org/10.1042/EBC20150002

Shoja, Z., Jalilvand, S., Latifi, T., & Roohvand, F. (2022). Rotavirus VP6 : Involvement in Immunogenicity , Adjuvant Activity , and Use as a Vector for Heterologous Peptides , Drug Delivery , and Production of Nano - Biomaterials. *Archives of Virology*, *167*(4), 1013–1023. https://doi.org/10.1007/s00705-022-05407-9

Sormanni, P., Aprile, F. A., & Vendruscolo, M. (2015). The CamSol Method of Rational Design of Protein Mutants with Enhanced Solubility. *Journal of Molecular Biology*, *427*, 478–490. https://doi.org/10.1016/j.jmb.2014.09.026

Steele, A. D., Neuzil, K. M., Cunliffe, N. A., Madhi, S. A., Bos, P., Ngwira, B., Witte, D., Todd, S., Louw, C., Kirsten, M., Aspinall, S., Van Doorn, L. J., Bouckenooghe, A., Suryakiran, P. V., & Han, H. H. (2012). Human rotavirus vaccine Rotarix™ provides protection against diverse circulating rotavirus strains in African infants: a randomized controlled trial. *BMC Infectious Diseases*, *12*(1), 1. https://doi.org/10.1186/1471-2334-12-213

Stobiecka, M., & Chałupa, A. (2015). Biosensors based on molecular beacons. *Chemical Papers*, *69*(1), 62–76. https://doi.org/10.1515/chempap-2015-0026

Sturniolo, T., Bono, E., Ding, J., Raddrizzani, L., Tuereci, O., Sahin, U., Braxenthaler, M., Gallazzi, F., Protti, M. P., Sinigaglia, F., & Hammer, J. (1999). Generation of

tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nature Biotechnology*, *17*(6), 555–561. https://doi.org/10.1038/9858

Sahdev, S., & Khattar, K. S. (2008). Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and Cellular Biochemistry*, *307*(1–2), 249–264.

Tamminen, K., Lappalainen, S., Huhti, L., Vesikari, T., & Blazevic, V. (2013). Trivalent Combination Vaccine Induces Broad Heterologous Immune Responses to Norovirus and Rotavirus in Mice. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0070409

Tartaglia, G. G., & Vendruscolo, M. (2008). The Zyggregator method for predicting protein aggregation propensities. *Chemical Society Reviews*, *37*(7), 1395–1401. https://doi.org/10.1039/b706784b

Tate, J. E., Burton, A. H., Boschi-Pinto, C., Parashar, U. D., Agocs, M., Serhan, F., De Oliveira, L., Mwenda, J. M., Mihigo, R., Ranjan Wijesinghe, P., Abeysinghe, N., Fox, K., & Paladin, F. (2016). Global, Regional, and National Estimates of Rotavirus Mortality in Children <5 Years of Age, 2000-2013. *Clinical Infectious Diseases*, *62*(Suppl 2), S96–S105. https://doi.org/10.1093/cid/civ1013

Troeger, C., Forouzanfar, M., Rao, P. C., Khalil, I., Brown, A., Reiner, R. C., Fullman, N., Thompson, R. L., Abajobir, A., Ahmed, M., Alemayohu, M. A., Alvis-Guzman, N., Amare, A. T., Antonio, C. A., Asayesh, H., Avokpaho, E., Awasthi, A., Bacha, U., Barac, A., & Mokdad, A. H. (2017). Estimates of global, regional, and national morbidity, mortality, and aetiologies of diarrhoeal diseases: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet Infectious Diseases*, *17*(9), 909–948. https://doi.org/10.1016/S1473-3099(17)30276-1

Troeger, C., Khalil, I. A., Rao, P. C., Cao, S., Blacker, B. F., Ahmed, T., Armah, G., Bines, J. E., Brewer, T. G., Colombara, D. V., Kang, G., Kirkpatrick, B. D., Kirkwood, C. D., Mwenda, J. M., Parashar, U. D., Petri, W. A., Riddle, M. S., Steele, A. D., Thompson, R. L., & Reiner, R. C. (2018). Rotavirus Vaccination and the Global Burden of Rotavirus Diarrhea among Children Younger Than 5

Years. *JAMA Pediatrics*, *172*(10), 958–965. https://doi.org/10.1001/jamapediatrics.2018.1960

Tsolis, A. C., Papandreou, N. C., Iconomidou, V. A., & Hamodrakas, S. J. (2013). A Consensus Method for the Prediction of "Aggregation-Prone" Peptides in Globular Proteins. *PLoS ONE*, *8*(1), 1–6. https://doi.org/10.1371/journal.pone.0054175

Tun, X., Wang, E. J., Gao, Z., Lundberg, K., Xu, R., & Hu, D. (2023). Integrin β3-Mediated Cell Senescence Associates with Gut Inflammation and Intestinal Degeneration in Models of Alzheimer's Disease. *International Journal of Molecular Sciences*, *24*(6). https://doi.org/10.3390/ijms24065697

Upadhyay, A. K., Murmu, A., Singh, A., & Panda, A. K. (2012). Kinetics of inclusion body formation and its correlation with the characteristics of protein aggregates in escherichia coli. *PLoS ONE*, *7*(3). https://doi.org/10.1371/journal.pone.0033951

Usman, M., Ayub, A., Habib, S., Rana, M. S., Rehman, Z., Zohaib, A., Jamal, S. B., Jaiswal, A. K., Andrade, B. S., de Carvalho Azevedo, V., Faheem, M., & Javed, A. (2023). Vaccinomics Approach for Multi-Epitope Vaccine Design against Group A Rotavirus Using VP4 and VP7 Proteins. *Vaccines*, *11*(4). https://doi.org/10.3390/vaccines11040726

Van Durme, J., De Baets, G., Van Der Kant, R., Ramakers, M., Ganesan, A., Wilkinson, H., Gallardo, R., Rousseau, F., & Schymkowitz, J. (2016). Solubis: A webserver to reduce protein aggregation through mutation. *Protein Engineering, Design and Selection*, *29*(8), 285–289. https://doi.org/10.1093/protein/gzw019

Varghese, T., Kang, G., & Steele, A. D. (2022). Understanding Rotavirus Vaccine Efficacy and Effectiveness in Countries with High Child Mortality. *Vaccines*, *10*(3), 1–15. https://doi.org/10.3390/vaccines10030346

Varghese, T., Khakha, S. A., Giri, S., Nair, N. P., Badur, M., Gathwala, G., Chaudhury, S., Kaushik, S., Dash, M., Mohakud, N. K., Ray, R. K., Mohanty, P., Kumar, C. P. G., Venkatasubramanian, S., Arora, R., Mohan, V. R., Tate, J. E., Parashar, U. D., & Kang, G. (2021). Rotavirus strain distribution before and after introducing rotavirus vaccine in india. *Pathogens*, *10*(4), 1–9. https://doi.org/10.3390/pathogens10040416

Vartak, A., & Sucheck, S. J. (2016). Recent advances in subunit vaccine carriers.

Vaccines, *4*(2), 1–18. https://doi.org/10.3390/vaccines4020012

Vega, C. G., Bok, M., Vlasova, A. N., Chattha, K. S., Ferna, F. M., Parren, V. G., & Saif, L. J. (2012). IgY Antibodies Protect against Human Rotavirus Induced Diarrhea in the Neonatal Gnotobiotic Piglet Disease Model. *PLoS ONE*, *7*(8), e42788. https://doi.org/10.1371/journal.pone.0042788

Velázquez, F. R., Matson, D. O., Calva, J. J., Guerrero, L., Morrow, A. L., Carter-Campbell, S., Glass, R. I., Estes, M. K., Pickering, L. K., & Ruiz-Palacios, G. M. (1996). Rotavirus infection in infants as protection against subsequent infections. *The New England Journal of Medicine*, *335*(14), 1022–1028.

Ventura, S. (2005). Sequence Determinants of Protein Aggregation : Tools to Increase Protein Solubility. *Microbial Cell Factories*, *4*(11). https://doi.org/10.1186/1475-2859-4-11

Ventura, S., & Villaverde, A. (2006). Protein Quality in Bacterial Inclusion Bodies. *Trends in Biotechnology*, *24*(4), 179–185. https://doi.org/10.1016/j.tibtech.2006.02.007

Vita, R., Mahajan, S., Overton, J. A., Dhanda, S. K., Martini, S., Cantrell, J. R., Wheeler, D. K., Sette, A., & Peters, B. (2019). The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Research*, *47*(D1), D339–D343. https://doi.org/10.1093/nar/gky1006

Wang, H., Naghavi, M., Allen, C., Barber, R. M., Carter, A., Casey, D. C., Charlson, F. J., Chen, A. Z., Coates, M. M., Coggeshall, M., Dandona, L., Dicker, D. J., Erskine, H. E., Haagsma, J. A., Fitzmaurice, C., Foreman, K., Forouzanfar, M. H., Fraser, M. S., Fullman, N., & Zuhlke, L. J. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, *388*(10053), 1459–1544. https://doi.org/10.1016/S0140-6736(16)31012-1

Wang, L., Maji, S. K., Sawaya, M. R., Eisenberg, D., & Riek, R. (2008). Bacterial Inclusion Bodies Contain Amyloid-Like Structure. *Plos Biology*, *6*(8), e195. https://doi.org/10.1371/journal.pbio.0060195

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer,

F. T., Beer, T. A. P. De, Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL : homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*, 296–303. https://doi.org/10.1093/nar/gky427

Wei, L., Ye, X., Sakurai, T., Mu, Z., & Wei, L. (2022). ToxIBTL: Prediction of peptide toxicity based on information bottleneck and transfer learning. *Bioinformatics*, *38*(6), 1514–1524. https://doi.org/10.1093/bioinformatics/btac006

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web : interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, *35*, 407–410. https://doi.org/10.1093/nar/gkm290

Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., & Hochstrasser, D. F. (1999). Protein Identification and Analysis Tools in the ExPASy Server. In: Link, A.J. (eds) 2-D Proteome Analysis Protocols. Methods in Molecular Biology. In *In: Link, A.J. (eds) 2-D Proteome Analysis Protocols. Methods in Molecular Biology, Humana Press. https://doi.org/10.1385/1-59259-584-7:531* (p. vol 112).

Willbold, D., Strodel, B., Schröder, G. F., Hoyer, W., & Heise, H. (2021). Amyloid-type Protein Aggregation and Prion-like Properties of Amyloids. *Chemical Reviews*, *121*(13), 8285–8307. https://doi.org/10.1021/acs.chemrev.1c00196

Won, J. H., Choi, J. S., & Jun, J. Il. (2022). CCN1 interacts with integrins to regulate intestinal stem cell proliferation and differentiation. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-30851-1

Xi, H., Jiang, H., Juhas, M., & Zhang, Y. (2021). Multiplex Biosensing for Simultaneous Detection of Mutations in SARS-CoV-2. *ACS Omega*, *6*(40), 25846–25859. https://doi.org/10.1021/acsomega.1c04024

Xiang, W., Peng, Z., Xu, J., Shen, H., & Li, W. (2020). Evaluation of a commercial latex agglutination test for detecting rotavirus A and human adenovirus in children's stool specimens. *Journal of Clinical Laboratory Analysis*, *34*(5), 1–4. https://doi.org/10.1002/jcla.23208

Yang, Y., Sun, W., Guo, J., Zhao, G., Sun, S., Yu, H., Guo, Y., Li, J., Jin, X., Du, L., Jiang, S., Kou, Z., & Zhou, Y. (2015). In silico design of a DNA-based HIV-1

multi-epitope vaccine for Chinese populations. *Human Vaccines and Immunotherapeutics*, *11*(3), 795–805. https://doi.org/10.1080/21645515.2015.1012017

Yano, A., Onozuka, A., Asahi-Ozaki, Y., Imai, S., Hanada, N., Miwa, Y., & Nisizawa, T. (2005). An ingenious design for peptide vaccines. *Vaccine*, *23*(17–18), 2322–2326. https://doi.org/10.1016/j.vaccine.2005.01.031

Yiduo Xiong, Yi Zhang, & Jian Wang, Y. X. (2023). Using 3dRNA/DNA for RNA and DNA 3D Structure Prediction and Evaluation. *Current Protocols*, *3*(5). https://doi.org/10.1002/cpz1.770

Yin, J., Li, G., Ren, X., & Herrler, G. (2007). Select What You Need : A Comparative Evaluation of the Advantages and Limitations of Frequently Used Expression Systems for Foreign Genes. *Journal of Biotechnology*, *127*(3), 335–347. https://doi.org/10.1016/j.jbiotec.2006.07.012

Zárate, S., Romero, P., Espinosa, R., Arias, C. F., & López, S. (2004). VP7 Mediates the Interaction of Rotaviruses with Integrin αvβ3 through a Novel Integrin-Binding Site. *Journal of Virology*, *78*(20), 10839–10847. https://doi.org/10.1128/jvi.78.20.10839-10847.2004

Zhang, J., & Zhao, Z. (2015). The Establishment of Biosensor Technology Based on F0F1-ATPase Molecular Motor for Detection of Rotavirus and Hepatitis A Virus. *Biosensors Journal*, *s4*(2), 1–5. https://doi.org/10.4172/2090-4967.1000121

Zhao, Q., Chen, W., Chen, Y., Zhang, L., Zhang, J., & Zhang, Z. (2011). Self-Assembled Virus-Like Particles from Rotavirus Structural Protein VP6 for Targeted Drug Delivery. *Bioconjugate Chemistry*, *22*(3), 346–352.

Zheng, J., Yang, R., Shi, M., Wu, C., Fang, X., Li, Y., Li, J., & Tan, W. (2015). Rationally designed molecular beacons for bioanalytical and biomedical applications. *Chemical Society Reviews*, *44*(10), 3036–3055. https://doi.org/10.1039/c5cs00020c

Zhou, B., Zhang, Y., Wang, X., Dong, J., Wang, B., Han, C., Yu, J., & Li, D. (2010). Oral Administration of Plant-Based Rotavirus VP6 Induces Antigen-Specific IgAs , IgGs and Passive Protection in Mice. *Vaccine*, *28*(37), 6021–6027. https://doi.org/10.1016/j.vaccine.2010.06.094

Zhou, R., Liu, L., & Wang, Y. (2021). Viral proteins recognized by different TLRs. *Journal of Medical Virology*, *93*(11), 6116–6123. https://doi.org/10.1002/jmv.27265

Appendix A: Supporting Tables and Figures

Table A.1. Full length RVA VP7 sequence details from NCBI Virus Variation Resource for the host: *Homo sapiens* and region: India.

Si. no.	Nucleotide Accession	Nucleotide Length (bases)	Protein Accession	Protein length (amino acid residues)	Segment / Protein	Genotype	Isolation source	Collection date
1	FJ948838	1062	ACZ04418	326	VP7	G1	feces	2006
2	FJ948845	1031	ACZ04425	326	VP7	G1	feces	2006
3	FJ948843	1031	ACZ04423	326	VP7	G1	feces	2006
4	FJ948854	1062	ACZ04434	326	VP7	G1	feces	2006
5	FJ948839	1062	ACZ04419	326	VP7	G1	feces	2006
6	FJ948833	1062	ACZ04413	326	VP7	G1	feces	2006
7	FJ948851	1062	ACZ04431	326	VP7	G1	feces	2006
8	FJ948849	1062	ACZ04429	326	VP7	G1	feces	2006
9	FJ948837	1062	ACZ04417	326	VP7	G1	feces	2006
10	FJ948830	1062	ACZ04410	326	VP7	G1	feces	2006
11	FJ948853	1031	ACZ04433	326	VP7	G1	feces	2006
12	FJ948831	1062	ACZ04411	326	VP7	G1	feces	2006
13	FJ948829	1062	ACZ04409	326	VP7	G1	feces	2006
14	FJ948832	1062	ACZ04412	326	VP7	G1	feces	2006
15	FJ948855	1062	ACZ04435	326	VP7	G1	feces	2006
16	FJ948847	1062	ACZ04427	326	VP7	G1	feces	2006
17	FJ948836	1062	ACZ04416	326	VP7	G1	feces	2006
18	FJ948848	1062	ACZ04428	326	VP7	G1	feces	2006
19	KP793024	1019	AKT71999	326	VP7	G1	feces	2014
20	JF742651	1048	AFD33531	326	VP7	G3	feces	2009
21	EU200798	1052	ABY64686	326	VP7	G10	feces	2009
22	EU984109	1062	ACL93332	326	VP7	G1	feces	2006
23	MK043948	1050	QAR17782	326	VP7	G3	feces	2017
24	MK043947	1022	QAR17781	326	VP7	G3	feces	2017
25	DQ478581	1062	ABF19729	326	VP7	G2	feces	2006
26	KX536667	1062	APG80718	326	VP7	G9	feces	2009
27	KX638509	1062	AST25338	326	VP7	G1	feces	2009
28	KX638510	1023	AST25339	326	VP7	G1	feces	2009
29	KX574239	1062	ASR91763	326	VP7	G9	feces	2009
30	KX574253	1062	ASR91777	326	VP7	G2	feces	2009

31	KX638511	1023	AST25340	326	VP7	G1	feces	2009
32	KX638512	1011	AST25341	326	VP7	G1	feces	2009
33	KX574276	1062	ASR91800	326	VP7	G12	feces	2009
34	KX574277	1062	ASR91801	326	VP7	G12	feces	2009
35	KX638513	1062	AST25342	326	VP7	G1	feces	2009
36	KX638538	1062	AST25367	326	VP7	G1	feces	2009
37	KX574236	1062	ASR91760	326	VP7	G9	feces	2009
38	KX574237	1062	ASR91761	326	VP7	G9	feces	2009
39	KX574278	1062	ASR91802	326	VP7	G12	feces	2009
40	KX638536	1016	AST25365	326	VP7	G1	feces	2009
41	KX638514	1011	AST25343	326	VP7	G1	feces	2009
42	KX638515	1062	AST25344	326	VP7	G1	feces	2009
43	KX574238	1062	ASR91762	326	VP7	G9	feces	2009
44	KX638516	1062	AST25345	326	VP7	G1	feces	2009
45	KX638517	1062	AST25346	326	VP7	G1	feces	2009
46	KX536668	1062	APG80719	326	VP7	G9	feces	2010
47	KX574240	1062	ASR91764	326	VP7	G9	feces	2010
48	KX638518	1062	AST25347	326	VP7	G1	feces	2010
49	KX638519	1062	AST25348	326	VP7	G1	feces	2010
50	KX638520	1062	AST25349	326	VP7	G1	feces	2010
51	KX574241	1062	ASR91765	326	VP7	G9	feces	2010
52	KX638521	1062	AST25350	326	VP7	G1	feces	2010
53	KX638522	1029	AST25351	326	VP7	G1	feces	2010
54	KX638523	1062	AST25352	326	VP7	G1	feces	2010
55	KX638524	1062	AST25353	326	VP7	G1	feces	2010
56	KX638525	1062	AST25354	326	VP7	G1	feces	2010
57	KX574254	1062	ASR91778	326	VP7	G2	feces	2010
58	KX574255	1062	ASR91779	326	VP7	G2	feces	2010
59	KX574242	1062	ASR91766	326	VP7	G9	feces	2010
60	KX574279	1062	ASR91803	326	VP7	G12	feces	2010
61	KX574232	1062	ASR91756	326	VP7	G9	feces	2010
62	KX574243	1062	ASR91767	326	VP7	G9	feces	2010
63	KX638526	1023	AST25355	326	VP7	G1	feces	2010
64	KX638527	1062	AST25356	326	VP7	G1	feces	2010
65	KX638537	1062	AST25366	326	VP7	G1	feces	2010
66	KX536669	1062	APG80720	326	VP7	G9	feces	2011
67	KX638528	981	AST25357	326	VP7	G1	feces	2011
68	KX638539	1062	AST25368	326	VP7	G1	feces	2011
69	KX638529	1062	AST25358	326	VP7	G1	feces	2011
70	KX638530	1062	AST25359	326	VP7	G1	feces	2011
71	KX638531	1062	AST25360	326	VP7	G1	feces	2011
72	KX638532	1062	AST25361	326	VP7	G1	feces	2011
73	KX574269	997	ASR91793	326	VP7	G2	feces	2011
74	KX574256	1023	ASR91780	326	VP7	G2	feces	2011
75	KX574257	1023	ASR91781	326	VP7	G2	feces	2011
76	KX574258	1048	ASR91782	326	VP7	G2	feces	2011

77	KX574259	1017	ASR91783	326	VP7	G2	feces	2011
78	KX574260	1017	ASR91784	326	VP7	G2	feces	2011
79	KX574244	1062	ASR91768	326	VP7	G9	feces	2011
80	KX574252	981	ASR91776	326	VP7	G9	feces	2011
81	KX574271	1023	ASR91795	326	VP7	G12	feces	2011
82	KX638533	981	AST25362	326	VP7	G1	feces	2011
83	KX574245	1062	ASR91769	326	VP7	G9	feces	2011
84	KX574246	1062	ASR91770	326	VP7	G9	feces	2011
85	KX638534	1062	AST25363	326	VP7	G1	feces	2011
86	KX638540	1062	AST25369	326	VP7	G1	feces	2011
87	KX638541	1062	AST25370	326	VP7	G1	feces	2012
88	KX638542	1062	AST25371	326	VP7	G1	feces	2012
89	KX638543	1062	AST25372	326	VP7	G1	feces	2012
90	KX638544	1062	AST25373	326	VP7	G1	feces	2012
91	KX638545	1062	AST25374	326	VP7	G1	feces	2012
92	KX574261	1052	ASR91785	326	VP7	G2	feces	2012
93	KX574262	1062	ASR91786	326	VP7	G2	feces	2012
94	KX574263	1062	ASR91787	326	VP7	G2	feces	2012
95	KX638546	1062	AST25375	326	VP7	G1	feces	2012
96	KX574247	1062	ASR91771	326	VP7	G9	feces	2012
97	KX574248	1062	ASR91772	326	VP7	G9	feces	2012
98	KX574272	1044	ASR91796	326	VP7	G12	feces	2012
99	KX574273	1057	ASR91797	326	VP7	G12	feces	2012
100	KX574280	1062	ASR91804	326	VP7	G12	feces	2012
101	KX574249	1062	ASR91773	326	VP7	G9	feces	2012
102	KX638547	1062	AST25376	326	VP7	G1	feces	2012
103	KX574233	1062	ASR91757	326	VP7	G9	feces	2012
104	KX574274	1044	ASR91798	326	VP7	G12	feces	2012
105	KX574275	1057	ASR91799	326	VP7	G12	feces	2012
106	KX574270	1062	ASR91794	326	VP7	G2	feces	2012
107	KX638548	1062	AST25377	326	VP7	G1	feces	2013
108	KX638549	1062	AST25378	326	VP7	G1	feces	2013
109	KX638535	1062	AST25364	326	VP7	G1	feces	2013
110	KX638550	1062	AST25379	326	VP7	G1	feces	2013
111	KX638551	1062	AST25380	326	VP7	G1	feces	2013
112	KX574264	1062	ASR91788	326	VP7	G2	feces	2013
113	KX574265	1062	ASR91789	326	VP7	G2	feces	2013
114	KX574266	1062	ASR91790	326	VP7	G2	feces	2013
115	KX574267	1062	ASR91791	326	VP7	G2	feces	2013
116	KX574268	1062	ASR91792	326	VP7	G2	feces	2013
117	KX574234	1062	ASR91758	326	VP7	G9	feces	2013
118	KX574235	1062	ASR91759	326	VP7	G9	feces	2013
119	KX574250	1062	ASR91774	326	VP7	G9	feces	2013
120	KX574251	1062	ASR91775	326	VP7	G9	feces	2013
121	KX574281	1062	ASR91805	326	VP7	G12	feces	2013
122	KX574282	1062	ASR91806	326	VP7	G12	feces	2013

123	KX574283	1062	ASR91807	326	VP7	G12	feces	2013
124	KX638552	1062	AST25381	326	VP7	G1	feces	2013
125	KX638553	1062	AST25382	326	VP7	G1	feces	2013
126	JX411969	964	AGH88811	297	VP7	G1	feces	2011
127	JX411970	967	AGH88812	297	VP7	G1	feces	2011
128	JX411968	1061	AGH88810	327	VP7	G9	feces	2011

Table A.2. Full length RVA VP4 sequence details from NCBI Virus Variation Resource for the host: *Homo sapiens* and region: India.

Si. no.	Nucleotide Accession	Nucleotide Length	Protein Accession	Protein Length	Segment / Protein	Genotype	Isolation source	Collection date
1	EU200796	2352	ABY64684	772	VP4	P[11]	feces	2009
2	HQ881574	2359	AFA52130	775	VP4	P[8]	feces	2007
3	HQ881575	2359	AFA52131	775	VP4	P[8]	feces	2008
4	EU984107	2359	ACL93330	775	VP4	P[8]	feces	2006
5	EU984106	2359	ACL93329	775	VP4	P[8]	feces	2006
6	HM467806	2359	ADW85797	775	VP4	P[8]	feces	2006
7	HM467803	2359	ADW85793	775	VP4	P[8]	feces	2006
8	HM467808	2359	ADW85796	775	VP4	P[8]	feces	2007
9	HM467805	2359	ADW85795	775	VP4	P[8]	feces	2007
10	HM467807	2359	ADW85798	775	VP4	P[8]	feces	2006
11	HM467804	2359	ADW85794	775	VP4	P[8]	feces	2006
12	KU738597	2328	AQT31696	775	VP4	P[8]	feces	2009
13	KU738599	2328	AQT31698	775	VP4	P[8]	feces	2011
14	KX536661	2362	APG80712	775	VP4	P[4]	feces	2009
15	KX646551	2328	ASU62251	775	VP4	P[8]	feces	2009
16	KX646552	2333	ASU62252	775	VP4	P[8]	feces	2009
17	KX646629	2359	ASU62329	775	VP4	P[4]	feces	2009
18	KX646613	2359	ASU62313	775	VP4	P[4]	feces	2009
19	KX646553	2359	ASU62253	775	VP4	P[8]	feces	2009
20	KX646554	2359	ASU62254	775	VP4	P[8]	feces	2009
21	KX646630	2362	ASU62330	775	VP4	P[4]	feces	2009
22	KX646605	2359	ASU62305	775	VP4	P[8]	feces	2009
23	KX646606	2359	ASU62306	775	VP4	P[8]	feces	2009
24	KX646555	2359	ASU62255	775	VP4	P[8]	feces	2009
25	KX646556	2335	ASU62256	775	VP4	P[8]	feces	2009
26	KX646592	2359	ASU62292	775	VP4	P[8]	feces	2009
27	KX646595	2359	ASU62295	775	VP4	P[8]	feces	2009
28	KX646607	2359	ASU62307	775	VP4	P[8]	feces	2009
29	KX646642	2359	ASU62342	775	VP4	P[6]	feces	2009
30	KX646557	2359	ASU62257	775	VP4	P[8]	feces	2009
31	KX646558	2359	ASU62258	775	VP4	P[8]	feces	2009
32	KX646596	2359	ASU62296	775	VP4	P[8]	feces	2009
33	KX646559	2359	ASU62259	775	VP4	P[8]	feces	2009
34	KX646560	2359	ASU62260	775	VP4	P[8]	feces	2009

35	KX536662	2362	APG80713	775	VP4	P[4]	feces	2010
36	KX646597	2359	ASU62297	775	VP4	P[8]	feces	2010
37	KX646561	2359	ASU62261	775	VP4	P[8]	feces	2010
38	KX646562	2359	ASU62262	775	VP4	P[8]	feces	2010
39	KX646563	2359	ASU62263	775	VP4	P[8]	feces	2010
40	KX646598	2359	ASU62298	775	VP4	P[8]	feces	2010
41	KX646564	2359	ASU62264	775	VP4	P[8]	feces	2010
42	KX646565	2359	ASU62265	775	VP4	P[8]	feces	2010
43	KX646566	2359	ASU62266	775	VP4	P[8]	feces	2010
44	KX646567	2359	ASU62267	775	VP4	P[8]	feces	2010
45	KX646568	2359	ASU62268	775	VP4	P[8]	feces	2010
46	KX646614	2359	ASU62314	775	VP4	P[4]	feces	2010
47	KX646615	2359	ASU62315	775	VP4	P[4]	feces	2010
48	KX646599	2359	ASU62299	775	VP4	P[8]	feces	2010
49	KX646608	2359	ASU62308	775	VP4	P[8]	feces	2010
50	KX646600	2359	ASU62300	775	VP4	P[8]	feces	2010
51	KX646631	2359	ASU62331	775	VP4	P[4]	feces	2010
52	KX646632	2359	ASU62332	775	VP4	P[4]	feces	2010
53	KX646569	2359	ASU62269	775	VP4	P[8]	feces	2010
54	KX646570	2359	ASU62270	775	VP4	P[8]	feces	2010
55	KX536663	2362	APG80714	775	VP4	P[4]	feces	2011
56	KX646571	2359	ASU62271	775	VP4	P[8]	feces	2011
57	KX646572	2335	ASU62272	775	VP4	P[8]	feces	2011
58	KX646573	2359	ASU62273	775	VP4	P[8]	feces	2011
59	KX646574	2359	ASU62274	775	VP4	P[8]	feces	2011
60	KX646575	2359	ASU62275	775	VP4	P[8]	feces	2011
61	KX646576	2359	ASU62276	775	VP4	P[8]	feces	2011
62	KX646612	2359	ASU62312	775	VP4	P[8]	feces	2011
63	KX646616	2359	ASU62316	775	VP4	P[4]	feces	2011
64	KX646617	2333	ASU62317	775	VP4	P[4]	feces	2011
65	KX646618	2359	ASU62318	775	VP4	P[4]	feces	2011
66	KX646619	2359	ASU62319	775	VP4	P[4]	feces	2011
67	KX646620	2359	ASU62320	775	VP4	P[4]	feces	2011
68	KX646633	2359	ASU62333	775	VP4	P[4]	feces	2011
69	KX646601	2359	ASU62301	775	VP4	P[8]	feces	2011
70	KX646645	2359	ASU62345	775	VP4	P[6]	feces	2011
71	KX646577	2359	ASU62277	775	VP4	P[8]	feces	2011
72	KX646634	2359	ASU62334	775	VP4	P[4]	feces	2011
73	KX646635	2359	ASU62335	775	VP4	P[4]	feces	2011
74	KX646578	2359	ASU62278	775	VP4	P[8]	feces	2011
75	KX646579	2359	ASU62279	775	VP4	P[8]	feces	2011
76	KX646636	2359	ASU62336	775	VP4	P[4]	feces	2011
77	KX646580	2359	ASU62280	775	VP4	P[8]	feces	2012
78	KX646581	2359	ASU62281	775	VP4	P[8]	feces	2012
79	KX646582	2359	ASU62282	775	VP4	P[8]	feces	2012
80	KX646583	2359	ASU62283	775	VP4	P[8]	feces	2012

81	KX646584	2359	ASU62284	775	VP4	P[8]	feces	2012
82	KX646621	2359	ASU62321	775	VP4	P[4]	feces	2012
83	KX646622	2359	ASU62322	775	VP4	P[4]	feces	2012
84	KX646623	2347	ASU62323	775	VP4	P[4]	feces	2012
85	KX646585	2359	ASU62285	775	VP4	P[8]	feces	2012
86	KX646637	2347	ASU62337	775	VP4	P[4]	feces	2012
87	KX646638	2359	ASU62338	775	VP4	P[4]	feces	2012
88	KX646646	2359	ASU62346	775	VP4	P[6]	feces	2012
89	KX646647	2359	ASU62347	775	VP4	P[6]	feces	2012
90	KX646648	2359	ASU62348	775	VP4	P[6]	feces	2012
91	KX646639	2359	ASU62339	775	VP4	P[4]	feces	2012
92	KX646586	2359	ASU62286	775	VP4	P[8]	feces	2012
93	KX646602	2359	ASU62302	775	VP4	P[8]	feces	2012
94	KX646649	2346	ASU62349	775	VP4	P[6]	feces	2012
95	KX646650	2359	ASU62350	775	VP4	P[6]	feces	2012
96	KX646644	2359	ASU62344	775	VP4	P[6]	feces	2012
97	KX646587	2359	ASU62287	775	VP4	P[8]	feces	2013
98	KX646588	2359	ASU62288	775	VP4	P[8]	feces	2013
99	KX646589	2359	ASU62289	775	VP4	P[8]	feces	2013
100	KX646590	2359	ASU62290	775	VP4	P[8]	feces	2013
101	KX646591	2359	ASU62291	775	VP4	P[8]	feces	2013
102	KX646624	2359	ASU62324	775	VP4	P[4]	feces	2013
103	KX646625	2359	ASU62325	775	VP4	P[4]	feces	2013
104	KX646626	2359	ASU62326	775	VP4	P[4]	feces	2013
105	KX646627	2359	ASU62327	775	VP4	P[4]	feces	2013
106	KX646628	2359	ASU62328	775	VP4	P[4]	feces	2013
107	KX646603	2359	ASU62303	775	VP4	P[8]	feces	2013
108	KX646604	2359	ASU62304	775	VP4	P[8]	feces	2013
109	KX646640	2358	ASU62340	775	VP4	P[4]	feces	2013
110	KX646641	2359	ASU62341	775	VP4	P[4]	feces	2013
111	KX646609	2359	ASU62309	775	VP4	P[8]	feces	2013
112	KX646610	2359	ASU62310	775	VP4	P[8]	feces	2013
113	KX646611	2359	ASU62311	775	VP4	P[8]	feces	2013
114	KX646643	2359	ASU62343	775	VP4	P[6]	feces	2013
115	KX646593	2359	ASU62293	775	VP4	P[8]	feces	2013
116	KX646594	2359	ASU62294	775	VP4	P[8]	feces	2013

Table A.3. Full length RVA VP6 gene sequences from NCBI Virus Variation Resource for the host: *Homo sapiens* and region: India.

Si. no.	Nculeotide Accession	Nucleotide Length	Protein Accession	Protein Length	Segment / Protein	Genotype	Isolation source	Collection date
1	HQ609563	1356	ADW41817	397	VP6	I1	feces	2006
2	HQ609562	1356	ADW41816	397	VP6	I1	feces	2006
3	HQ609561	1356	ADW41815	397	VP6	I1	feces	2006
4	EU984108	1356	ACL93331	397	VP6	I1	feces	2006
5	FJ685614	1356	ACV81679	397	VP6	I1	feces	1992
6	KF740531	1331	AIC82074	397	VP6	I12	feces	2008
7	KX536664	1356	APG80715	397	VP6	I2	feces	2009
8	KX638554	1353	AST25383	397	VP6	I1	feces	2009
9	KX638555	1320	AST25384	397	VP6	I1	feces	2009
10	KX638641	1356	AST25470	397	VP6	I2	feces	2009
11	KX638624	1356	AST25453	397	VP6	I2	feces	2009
12	KX638556	1356	AST25385	397	VP6	I1	feces	2009
13	KX638557	1356	AST25386	397	VP6	I1	feces	2009
14	KX638642	1356	AST25471	397	VP6	I2	feces	2009
15	KX638611	1356	AST25440	397	VP6	I1	feces	2009
16	KX638612	1356	AST25441	397	VP6	I1	feces	2009
17	KX638558	1356	AST25387	397	VP6	I1	feces	2009
18	KX638559	1356	AST25388	397	VP6	I1	feces	2009
19	KX638600	1356	AST25429	397	VP6	I1	feces	2009
20	KX638601	1356	AST25430	397	VP6	I1	feces	2009
21	KX638613	1356	AST25442	397	VP6	I1	feces	2009
22	KX638598	1356	AST25427	397	VP6	I1	feces	2009
23	KX638560	1356	AST25389	397	VP6	I1	feces	2009
24	KX638561	1356	AST25390	397	VP6	I1	feces	2009
25	KX638602	1356	AST25431	397	VP6	I1	feces	2009
26	KX638562	1356	AST25391	397	VP6	I1	feces	2009
27	KX638563	1353	AST25392	397	VP6	I1	feces	2009
28	KX536665	1356	APG80716	397	VP6	I2	feces	2010
29	KX638603	1356	AST25432	397	VP6	I1	feces	2010
30	KX638564	1353	AST25393	397	VP6	I1	feces	2010
31	KX638565	1356	AST25394	397	VP6	I1	feces	2010
32	KX638566	1353	AST25395	397	VP6	I1	feces	2010
33	KX638604	1356	AST25433	397	VP6	I1	feces	2010
34	KX638567	1356	AST25396	397	VP6	I1	feces	2010
35	KX638568	1356	AST25397	397	VP6	I1	feces	2010
36	KX638569	1356	AST25398	397	VP6	I1	feces	2010
37	KX638570	1353	AST25399	397	VP6	I1	feces	2010
38	KX638571	1353	AST25400	397	VP6	I1	feces	2010
39	KX638625	1356	AST25454	397	VP6	I2	feces	2010
40	KX638626	1356	AST25455	397	VP6	I2	feces	2010
41	KX638605	1356	AST25434	397	VP6	I1	feces	2010

42	KX638614	1353	AST25443	397	VP6	I1	feces	2010
43	KX638606	1353	AST25435	397	VP6	I1	feces	2010
44	KX638643	1356	AST25472	397	VP6	I2	feces	2010
45	KX638644	1356	AST25473	397	VP6	I2	feces	2010
46	KX638572	1353	AST25401	397	VP6	I1	feces	2010
47	KX638573	1356	AST25402	397	VP6	I1	feces	2010
48	KX638599	1356	AST25428	397	VP6	I1	feces	2010
49	KX536666	1356	APG80717	397	VP6	I2	feces	2011
50	KX638574	1320	AST25403	397	VP6	I1	feces	2011
51	KX638575	1356	AST25404	397	VP6	I1	feces	2011
52	KX638576	1353	AST25405	397	VP6	I1	feces	2011
53	KX638577	1353	AST25406	397	VP6	I1	feces	2011
54	KX638578	1353	AST25407	397	VP6	I1	feces	2011
55	KX638579	1353	AST25408	397	VP6	I1	feces	2011
56	KX638595	1353	AST25424	397	VP6	I1	feces	2011
57	KX638627	1356	AST25456	397	VP6	I2	feces	2011
58	KX638628	1356	AST25457	397	VP6	I2	feces	2011
59	KX638629	1299	AST25458	397	VP6	I2	feces	2011
60	KX638630	1356	AST25459	397	VP6	I2	feces	2011
61	KX638631	1356	AST25460	397	VP6	I2	feces	2011
62	KX638645	1356	AST25474	397	VP6	I2	feces	2011
63	KX638607	1356	AST25436	397	VP6	I1	feces	2011
64	KX638618	1317	AST25447	397	VP6	I1	feces	2011
65	KX638580	1353	AST25409	397	VP6	I1	feces	2011
66	KX638646	1356	AST25475	397	VP6	I2	feces	2011
67	KX638647	1356	AST25476	397	VP6	I2	feces	2011
68	KX638581	1353	AST25410	397	VP6	I1	feces	2011
69	KX638582	1356	AST25411	397	VP6	I1	feces	2011
70	KX638648	1356	AST25477	397	VP6	I2	feces	2011
71	KX638583	1356	AST25412	397	VP6	I1	feces	2012
72	KX638584	1356	AST25413	397	VP6	I1	feces	2012
73	KX638585	1356	AST25414	397	VP6	I1	feces	2012
74	KX638586	1356	AST25415	397	VP6	I1	feces	2012
75	KX638587	1356	AST25416	397	VP6	I1	feces	2012
76	KX638632	1356	AST25461	397	VP6	I2	feces	2012
77	KX638633	1356	AST25462	397	VP6	I2	feces	2012
78	KX638634	1356	AST25463	397	VP6	I2	feces	2012
79	KX638588	1356	AST25417	397	VP6	I1	feces	2012
80	KX638649	1351	AST25478	397	VP6	I2	feces	2012
81	KX638650	1356	AST25479	397	VP6	I2	feces	2012
82	KX638619	1356	AST25448	397	VP6	I1	feces	2012
83	KX638620	1353	AST25449	397	VP6	I1	feces	2012
84	KX638621	1356	AST25450	397	VP6	I1	feces	2012
85	KX638651	1356	AST25480	397	VP6	I2	feces	2012
86	KX638589	1356	AST25418	397	VP6	I1	feces	2012
87	KX638608	1356	AST25437	397	VP6	I1	feces	2012

88	KX638622	1356	AST25451	397	VP6	I1	feces	2012
89	KX638623	1194	AST25452	397	VP6	I1	feces	2012
90	KX638640	1356	AST25469	397	VP6	I2	feces	2012
91	KX638590	1356	AST25419	397	VP6	I1	feces	2013
92	KX638591	1356	AST25420	397	VP6	I1	feces	2013
93	KX638592	1356	AST25421	397	VP6	I1	feces	2013
94	KX638593	1356	AST25422	397	VP6	I1	feces	2013
95	KX638594	1356	AST25423	397	VP6	I1	feces	2013
96	KX638635	1356	AST25464	397	VP6	I2	feces	2013
97	KX638636	1356	AST25465	397	VP6	I2	feces	2013
98	KX638637	1356	AST25466	397	VP6	I2	feces	2013
99	KX638638	1356	AST25467	397	VP6	I2	feces	2013
100	KX638639	1356	AST25468	397	VP6	I2	feces	2013
101	KX638609	1356	AST25438	397	VP6	I1	feces	2013
102	KX638610	1356	AST25439	397	VP6	I1	feces	2013
103	KX638652	1356	AST25481	397	VP6	I2	feces	2013
104	KX638653	1356	AST25482	397	VP6	I2	feces	2013
105	KX638615	1356	AST25444	397	VP6	I1	feces	2013
106	KX638616	1356	AST25445	397	VP6	I1	feces	2013
107	KX638617	1356	AST25446	397	VP6	I1	feces	2013
108	KX638596	1355	AST25425	397	VP6	I1	feces	2013
109	KX638597	1356	AST25426	397	VP6	I1	feces	2013
110	MG676125	1301	AYM47553	397	VP6	I1	feces	2015
111	MG676126	1301	AYM47554	397	VP6	I1	feces	2015
112	MG676127	1305	AYM47555	397	VP6	I1	feces	2015
113	MG676128	1331	AYM47556	397	VP6	I1	feces	2015
114	MG676121	1301	AYM47549	397	VP6	I1	feces	2014
115	MG676122	1301	AYM47550	397	VP6	I1	feces	2014
116	MG676123	1301	AYM47551	397	VP6	I2	feces	2015
117	MG676124	1301	AYM47552	397	VP6	I1	feces	2015
118	MG676130	1303	AYM47558	397	VP6	I2	feces	2015
119	MG676131	1292	AYM47559	397	VP6	I2	feces	2015
120	MG676132	1301	AYM47560	397	VP6	I1	feces	2015
121	MG676133	1321	AYM47561	397	VP6	I2	feces	2015
122	MG676134	1320	AYM47562	397	VP6	I1	feces	2015
123	MG676135	1321	AYM47563	397	VP6	I1	feces	2015
124	MG676136	1301	AYM47564	397	VP6	I1	feces	2015
125	MG676137	1282	AYM47565	397	VP6	I1	feces	2015
126	MG676138	1297	AYM47566	397	VP6	I1	feces	2015
127	MG676139	1300	AYM47567	397	VP6	I2	feces	2016

Table A.4. Details of 1378 full length human rotavirus A VP6 gene sequences used for the design of molecular beacon.

Si. no.	Accession ID	Region	Year
1	KP645271	Australia	2010
2	KX632247	Uganda	2013
3	GU390437	Russia	2005
4	HM066155	Brazil	2009
5	KU048630	Italy	2016
6	KX655499	Uganda	2013
7	JQ230070	Russia	2009
8	GQ468562	Russia	2004
9	GQ477107	Russia	2007
10	MH060092	Thailand	2013
11	JX841153	Russia	2011
12	KX638581	India	2011
13	KX655510	Uganda	2013
14	KC139784	China	2000
15	KJ126826	China	2009
16	GQ477130	Russia	2009
17	KX638573	India	2010
18	KP882320	Ghana	2008
19	KY657824	China	2016
20	JX867344	Russia	2011
21	KF036172	Russia	2012
22	GU390429	Russia	2005
23	KU248409	Bangladesh	2010
24	KT988297	Italy	2013
25	JN014003	South Africa	2008
26	LC102930	Japan	2014
27	KF036168	Russia	2011
28	GQ468560	Russia	2004
29	JX195089	Italy	2010
30	KX638568	India	2010
31	KX638615	India	2013
32	DQ490555	Bangladesh	2000
33	KX880438	Germany	2014
34	MG676125	India	2015
35	KT988220	Italy	2010
36	LC105316	Japan	2014
37	KY657816	China	2016
38	GQ477093	Russia	2007
39	JQ230102	Russia	2008
40	KJ820873	Brazil	1997
41	HQ657175	South Africa	2009
42	KY657827	China	2016
43	KY657830	China	2016
44	JQ230092	Russia	2010
45	JX567763	Australia	2011
46	KF360215	Brazil	2010
47	MG181781	Malawi	2012
48	AB733132	Japan	1980

49	JQ613169	Russia	2010
50	JN706542	Thailand	2009
51	MG599532	Brazil	2013
52	KX536664	India	2009
53	KX638592	India	2013
54	LC105524	Japan	2014
55	KY657847	China	2017
56	GU390414	Russia	2003
57	MH060102	Thailand	2017
58	JN706532	Thailand	2008
59	KJ721696	Brazil	2011
60	KX638617	India	2013
61	MG181539	Malawi	2012
62	JQ585602	Russia	2010
63	LC105425	Japan	2014
64	JX867354	Russia	2010
65	HQ611011	Russia	2010
66	HM123840	Brazil	2008
67	LC105327	Japan	2014
68	KF036174	Russia	2012
69	HQ611006	Russia	2010
70	MH484057	Russia	2013
71	KX638606	India	2010
72	KP881924	Bangladesh	2008
73	KY657838	China	2016
74	KF036134	Russia	2006
75	GQ477099	Russia	2008
76	GU390457	Russia	2009
77	KT694998	USA	1988
78	KX638611	India	2009
79	KJ748480	Ghana	2008
80	LC105090	Japan	2014
81	JX867364	Russia	2010
82	JQ230060	Russia	2004
83	MG676132	India	2015
84	JX867351	Russia	2010
85	GU390446	Russia	2008
86	HM066131	Brazil	2006
87	HM123834	Brazil	2007
88	KC782523	USA	2010
89	KP882551	Ghana	2008
90	MG181330	Malawi	2009
91	KF036160	Russia	2010
92	HM066127	Brazil	2006
93	HM123822	Brazil	2006
94	KX171562	South Korea	2012
95	LC228308	Japan	2015
96	KX638590	India	2013
97	LC103040	Japan	2014
98	LC086806	Thailand	2013
99	JQ818163	Russia	2004
100	HM066159	Brazil	2009
101	KC890878	Greece	2010

102	KC618426	South Africa	2008
103	LC406790	Kenya	2014
104	KX632324	Uganda	2013
105	HQ445976	Russia	2007
106	GQ477124	Russia	2009
107	KX356111	Japan	2012
108	KX638554	India	2009
109	KX638629	India	2011
110	KC834700	Australia	2004
111	GQ477133	Russia	2009
112	LC169944	Thailand	2014
113	AB796451	Japan	2012
114	KM116042	USA	2013
115	MG214341	Morocco	2011
116	KP882144	Bangladesh	2008
117	MF580851	China	2014
118	MF580853	China	2016
119	MH473479	Russia	2012
120	KX638602	India	2009
121	KF690129	Australia	2012
122	LC105402	Japan	2014
123	GU390415	Russia	2003
124	KJ721701	Brazil	2008
125	KF716338	USA	2011
126	KX638587	India	2012
127	KP883101	Mali	2008
128	KC571496	Australia	2010
129	KX638579	India	2011
130	KJ870915	Democratic Republic of the Congo	2008
131	JX965142	Australia	2010
132	GQ477080	Russia	2006
133	HQ738592	Russia	2008
134	HQ738596	Russia	2008
135	JX195067	Italy	2010
136	KP882760	Mali	2008
137	MG181682	Malawi	2013
138	JX867356	Russia	2010
139	MH473484	Russia	2012
140	KX171568	South Korea	2015
141	LC169933	Thailand	2014
142	JN129110	Nicaragua	2010
143	JN129108	Nicaragua	2010
144	KJ870882	Democratic Republic of the Congo	2010
145	KY657841	China	2016
146	KF812682	South Korea	2010
147	KF035111	Barbados	2012
148	JN591408	Malawi	2012
149	KF812690	South Korea	2012
150	KU363133	Thailand	2013
151	MG181792	Malawi	2012
152	KF812666	South Korea	2009
153	FJ685614	India	1992

154	JN129099	Nicaragua	2010
155	KJ162412	Russia	2011
156	KP882738	Kenya	2009
157	LC105371	Japan	2014
158	JQ230081	Russia	2009
159	KF812670	South Korea	2011
160	KJ820829	Brazil	1997
161	LC169845	Thailand	2013
162	KC178803	Italy	2007
163	LC105035	Japan	2014
164	KF036140	Russia	2010
165	JX867366	Russia	2010
166	KF812708	South Korea	2011
167	HM123829	Brazil	2007
168	KP007183	Philippines	2012
169	KX778586	China	2015
170	KP882034	Bangladesh	2008
171	KP882243	Bangladesh	2008
172	KF812696	South Korea	2007
173	GQ468559	Russia	2004
174	KX356117	Japan	2012
175	JX867349	Russia	2010
176	KX638646	India	2011
177	KP882606	Ghana	2009
178	KY657817	China	2016
179	GQ477094	Russia	2007
180	KU356666	Bangladesh	2013
181	LC169856	Thailand	2013
182	KU048616	Italy	2016
183	HM467948	USA	2011
184	KX638562	India	2009
185	KX638570	India	2010
186	KP941128	Kenya	2008
187	KF036135	Russia	2006
188	GQ477102	Russia	2008
189	KF371671	China	2007
190	KJ721702	Brazil	2007
191	MG181924	Malawi	2012
192	GQ477120	Russia	2009
193	MH473483	Russia	2012
194	KF907300	Brazil	2006
195	KP882122	Bangladesh	2008
196	MG181583	Malawi	2013
197	JQ230097	Russia	2011
198	JX271005	Tunisia	2008
199	JQ230067	Russia	2008
200	HQ611014	Russia	2010
201	KP645260	Australia	2010
202	KP883145	Mali	2009
203	KX632313	Uganda	2012
204	GQ468565	Russia	2005
205	KU356611	Bangladesh	2013
206	JN129106	Nicaragua	2010

207	KX638565	India	2010
208	KX638651	India	2012
209	MG181462	Malawi	2005
210	GQ468566	Russia	2005
211	HM066152	Brazil	2008
212	KX171569	South Korea	2015
213	KU356589	Bangladesh	2010
214	KP882947	Mali	2008
215	KX655443	Uganda	2013
216	LC086740	Thailand	2013
217	KF680463	Russia	2010
218	JX867347	Russia	2010
219	KX638648	India	2011
220	JQ585598	Russia	2006
221	GU390430	Russia	2005
222	HM123828	Brazil	2006
223	FJ747629	Germany	2008
224	KP645293	Australia	2010
225	KT988165	Italy	2009
226	KP882859	Mali	2008
227	LC119097	Nigeria	1999
228	HQ611001	Russia	2010
229	MG676128	India	2015
230	JQ713097	Russia	2006
231	KF036171	Russia	2012
232	HQ611003	Russia	2010
233	KC618421	South Africa	2008
234	KU048628	Italy	2016
235	KX778596	China	2016
236	KX638595	India	2011
237	MG181385	Malawi	2002
238	LC086784	Thailand	2013
239	KY657833	China	2016
240	KF812713	South Korea	2011
241	MH473500	Russia	2012
242	JQ230093	Russia	2010
243	JF460838	Belgium	2009
244	LC438383	Japan	1974
245	KP882793	Mali	2008
246	KF812701	South Korea	2011
247	JQ230061	Russia	2004
248	HM123817	Brazil	2005
249	MH569771	Brazil	2016
250	LC019045	Myanmar	2011
251	JQ230079	Russia	2009
252	KJ126829	China	2011
253	GU390452	Russia	2009
254	GQ477104	Russia	2008
255	JN706533	Thailand	2008
256	KP882628	Ghana	2008
257	KP883167	Mali	2008
258	KR701633	USA	2011
259	LC102974	Japan	2014

260	KC713896	Russia	2012
261	MH484056	Russia	2013
262	KU248432	Bangladesh	2010
263	MG599533	Brazil	2013
264	KX638593	India	2013
265	KX638609	India	2013
266	KF812667	South Korea	2009
267	KU248376	Bangladesh	2010
268	AB796452	Japan	2012
269	KU048621	Italy	2016
270	KJ094939	China	2010
271	GQ477127	Russia	2009
272	KF036139	Russia	2010
273	HQ610998	Russia	2010
274	GQ477114	Russia	2008
275	HM066126	Brazil	2005
276	MG676123	India	2015
277	KP882925	Mali	2008
278	KU356600	Bangladesh	2013
279	HQ702223	USA	2009
280	KX778590	China	2015
281	KX638637	India	2013
282	LC066665	Thailand	2013
283	GU390443	Russia	2008
284	GU592515	Russia	2009
285	GU390449	Russia	2008
286	HM123835	Brazil	2007
287	KM116053	USA	2013
288	KX632269	Uganda	2013
289	LC086773	Thailand	2013
290	AB796454	Japan	2012
291	KX638626	India	2010
292	LC105262	Japan	2014
293	JQ818160	Russia	2003
294	KC618427	South Africa	2008
295	LC105272	Japan	2014
296	MF580850	China	2013
297	KF812677	South Korea	2012
298	HM123825	Brazil	2006
299	JN591410	Malawi	2012
300	KX638639	India	2013
301	KT988231	Italy	2012
302	LC105283	Japan	2014
303	KY657821	China	2016
304	HM066141	Brazil	2008
305	LC228374	Japan	2016
306	JX841142	Russia	2011
307	KX356104	Japan	2012
308	KX257406	Suriname	2013
309	AB669018	Japan	2008
310	GQ477121	Russia	2009
311	GQ477125	Russia	2009
312	HM467946	USA	2011

313	KX638576	India	2011
314	KP882276	Bangladesh	2008
315	KP882749	Mali	2008
316	MG181231	Malawi	1998
317	JQ818166	Russia	2007
318	JX867367	Russia	2010
319	MH473481	Russia	2012
320	KP170807	Brazil	2008
321	JN129103	Nicaragua	2010
322	KX778587	China	2015
323	KX638586	India	2012
324	KP882177	Bangladesh	2008
325	MG181407	Malawi	2003
326	JQ230084	Russia	2010
327	JQ230088	Russia	2010
328	KF812671	South Korea	2011
329	KX171563	South Korea	2012
330	KC834699	Australia	2009
331	KU059781	Australia	2013
332	GQ477083	Russia	2006
333	JQ818153	Russia	2008
334	HQ611032	Russia	2010
335	HQ738597	Russia	2009
336	HM066132	Brazil	2006
337	KF447854	China	2010
338	LC406812	Kenya	2011
339	KP882419	Ghana	2008
340	KX632346	Uganda	2013
341	LC105392	Japan	2014
342	KF812693	South Korea	2011
343	LC386080	Japan	2017
344	KX638628	India	2011
345	KX638622	India	2012
346	KP645315	Australia	2010
347	LC105216	Japan	2014
348	KY657842	China	2016
349	KF036143	Russia	2010
350	KF812683	South Korea	2010
351	JQ230058	Russia	2003
352	JN887819	South Korea	2006
353	LC228319	Japan	2016
354	KX778589	China	2015
355	KP882804	Mali	2008
356	KF812715	South Korea	2011
357	KF812673	South Korea	2012
358	KP007192	Philippines	2012
359	AB714263	Viet Nam	2008
360	KP882518	Ghana	2008
361	KP883189	Mali	2008
362	LC384334	Japan	2016
363	GU390423	Russia	2005
364	KU048629	Italy	2016
365	KX356109	Japan	2012

366	KX638645	India	2011
367	KP882826	Mali	2008
368	LC406834	Kenya	2013
369	MG181726	Malawi	2014
370	MG181737	Malawi	2014
371	JX867348	Russia	2010
372	MH473490	Russia	2012
373	MH473501	Russia	2012
374	HQ611019	Russia	2010
375	KF716341	USA	2011
376	KP882991	Mali	2009
377	JF521477	USA	1980
378	JQ230086	Russia	2010
379	MH473470	Russia	2011
380	HM066153	Brazil	2008
381	MG181616	Malawi	2013
382	KY657849	China	2017
383	JQ230076	Russia	2009
384	JQ248938	Russia	2006
385	HQ738601	Russia	2010
386	KP882661	Ghana	2008
387	MG181253	Malawi	1999
388	LC105338	Japan	2014
389	KX171558	South Korea	2010
390	JN129111	Nicaragua	2010
391	KP883035	Mali	2008
392	JX867346	Russia	2010
393	KX638566	India	2010
394	KX638652	India	2013
395	JQ713098	Russia	2007
396	JX867360	Russia	2010
397	MH484055	Russia	2012
398	KU048627	Italy	2016
399	KU048619	Italy	2016
400	KP645337	Australia	2011
401	MG676134	India	2015
402	HM123837	Brazil	2008
403	JN706534	Thailand	2008
404	KC618430	South Africa	2008
405	HQ702212	USA	2009
406	KP006511	Guatemala	2009
407	KX638643	India	2010
408	KP645282	Australia	2010
409	MG181429	Malawi	2004
410	JQ585600	Russia	2007
411	GU138214	Russia	2007
412	KP007203	Philippines	2012
413	KX638598	India	2009
414	KF812704	South Korea	2012
415	JQ230066	Russia	2007
416	MG599534	Brazil	2013
417	KY657825	China	2016
418	KF036136	Russia	2007

419	KF907296	Brazil	2011
420	KF360213	Brazil	2010
421	KF360211	Brazil	2010
422	KJ748469	Ghana	2008
423	KT988209	Italy	2010
424	KY657851	China	2017
425	KF661429	Russia	2010
426	KC713892	Russia	2012
427	GQ477109	Russia	2007
428	GU138216	Russia	2007
429	HM123818	Brazil	2006
430	JN706544	Thailand	2009
431	JQ230094	Russia	2010
432	JX867352	Russia	2010
433	KY937206	Cambodia	2014
434	KU317470	Egypt	2012
435	KP882155	Bangladesh	2008
436	LC105176	Japan	2014
437	JQ818154	Russia	2008
438	MH473473	Russia	2012
439	JN591411	Malawi	2012
440	LC169977	Thailand	2014
441	KP188812	Italy	2009
442	KP882650	Ghana	2008
443	JF521466	USA	1980
444	JQ818161	Russia	2003
445	JQ230095	Russia	2010
446	KF812711	South Korea	2012
447	GQ477098	Russia	2008
448	KU048624	Italy	2016
449	KC178799	Italy	1996
450	MH704729	Indonesia	2015
451	HQ657142	Malawi	2001
452	LC102952	Japan	2014
453	KF661433	Russia	2010
454	KY497555	Pakistan	2010
455	JN129100	Nicaragua	2010
456	KX638627	India	2011
457	KX638634	India	2012
458	KP882100	Bangladesh	2008
459	KU356655	Bangladesh	2013
460	KC890876	Greece	2008
461	KC618424	South Africa	2008
462	AB605601	Viet Nam	2006
463	LC169889	Thailand	2013
464	KP882914	Mali	2008
465	GU390420	Russia	2004
466	KP170812	Brazil	2009
467	KJ820895	Brazil	1998
468	KX632258	Uganda	2013
469	HQ611002	Russia	2010
470	GU390444	Russia	2007
471	LC367318	Thailand	2015

472	HM123824	Brazil	2006
473	KX778593	China	2015
474	KP882067	Bangladesh	2008
475	JX965144	Australia	2010
476	MF580848	China	2011
477	JQ230073	Russia	2009
478	HM123836	Brazil	2007
479	AB796449	Japan	2012
480	KX638589	India	2012
481	JQ230083	Russia	2009
482	AB605598	Viet Nam	2006
483	KP882111	Bangladesh	2008
484	KY657843	China	2016
485	JQ818167	Russia	2007
486	HQ611024	Russia	2010
487	HQ738584	Russia	2008
488	GQ477113	Russia	2008
489	MH473480	Russia	2012
490	KU048632	Italy	2016
491	KX356103	Japan	2012
492	KP882573	Ghana	2009
493	MG181847	Malawi	2012
494	MG181319	Malawi	2013
495	LC086751	Thailand	2013
496	MH060112	Thailand	2017
497	KT695108	USA	1991
498	KX356113	Japan	2012
499	KC178807	Italy	2010
500	KX638575	India	2011
501	KY657822	China	2016
502	HM123830	Brazil	2007
503	LC190490	Thailand	2014
504	JN014004	South Africa	2008
505	MG781045	Thailand	2011
506	HM066146	Brazil	2008
507	KP882683	Ghana	2008
508	MG181275	Malawi	2000
509	KX632335	Uganda	2012
510	KF812684	South Korea	2010
511	JQ230059	Russia	2004
512	KU597746	China	2014
513	KP645304	Australia	2010
514	KP882452	Ghana	2009
515	GU390419	Russia	2004
516	KF661455	Russia	2012
517	KF812676	South Korea	2012
518	HM467947	USA	2011
519	KC178805	Italy	2008
520	KX638632	India	2012
521	KX638621	India	2012
522	KX632302	Uganda	2012
523	KT936630	Thailand	2012
524	KF036142	Russia	2010

525	KF812664	South Korea	2007
526	HM066138	Brazil	2007
527	KJ162414	Russia	2012
528	KX655477	Uganda	2012
529	LC105360	Japan	2014
530	KF812695	South Korea	2011
531	HM066156	Brazil	2008
532	JX841159	Russia	2011
533	KT281125	USA	2012
534	LC105002	Japan	2014
535	LC102941	Japan	2014
536	KY657835	China	2016
537	KJ721704	Brazil	2006
538	KX638569	India	2010
539	KP882254	Bangladesh	2008
540	KJ411437	USA	2012
541	GU390436	Russia	2005
542	FJ423118	USA	2009
543	KX638620	India	2012
544	MH569769	Brazil	2015
545	KP882727	Kenya	2009
546	KP882892	Mali	2008
547	JQ230071	Russia	2009
548	HM123819	Brazil	2006
549	KP883024	Mali	2008
550	KJ126827	China	2010
551	GU390426	Russia	2005
552	GQ477092	Russia	2007
553	KF907297	Brazil	2011
554	JX841166	Russia	2011
555	DQ490549	Bangladesh	2000
556	KU248387	Bangladesh	2010
557	KF716342	USA	2011
558	KX638618	India	2011
559	LC086718	Thailand	2013
560	JQ230101	Russia	2011
561	GU390434	Russia	2005
562	GQ468568	Russia	2005
563	MH060122	Thailand	2017
564	KJ820862	Brazil	1997
565	JQ613168	Russia	2010
566	HM066150	Brazil	2008
567	KJ721699	Brazil	2010
568	LC336590	Japan	2012
569	KU956011	Honduras	2011
570	LC105257	Japan	2014
571	HM066149	Brazil	2008
572	MG676126	India	2015
573	KF036173	Russia	2012
574	GU390458	Russia	2009
575	JN706541	Thailand	2008
576	KX638601	India	2009
577	EF554086	Belgium	2002

578	MG181627	Malawi	2013
579	KF447865	China	2010
580	KC890881	Greece	2010
581	KX638603	India	2010
582	KU508384	Hungary	2002
583	LC105123	Japan	2014
584	AB669006	Japan	2008
585	KC713888	Russia	2011
586	JX867363	Russia	2010
587	KF661451	Russia	2012
588	MH473474	Russia	2012
589	KU048626	Italy	2016
590	MG676131	India	2015
591	MG676137	India	2015
592	AB861949	Kenya	2010
593	KF036155	Russia	2012
594	HQ611010	Russia	2010
595	KF371681	China	2004
596	KF716346	USA	2011
597	KX638612	India	2009
598	KP881891	Bangladesh	2008
599	KP881935	Bangladesh	2008
600	KT988187	Italy	2009
601	KU714448	Malawi	2000
602	LC105101	Japan	2014
603	JN706547	Thailand	2011
604	KJ820907	Brazil	1999
605	KX638591	India	2013
606	JQ988903	Croatia	2006
607	KY616905	Japan	2014
608	MG181880	Malawi	2012
609	KY657828	China	2016
610	JQ230077	Russia	2009
611	HQ611009	Russia	2010
612	JN706535	Thailand	2008
613	KF360214	Brazil	2010
614	KX638640	India	2012
615	KP881913	Bangladesh	2008
616	KP882694	Kenya	2009
617	KC713901	Russia	2012
618	JX867362	Russia	2010
619	JX867359	Russia	2010
620	KP882617	Ghana	2009
621	HQ738599	Russia	2009
622	HM123831	Brazil	2007
623	LC169922	Thailand	2014
624	KX638559	India	2009
625	MG181935	Malawi	2000
626	HQ445977	Russia	2007
627	HM123821	Brazil	2006
628	JQ818164	Russia	2004
629	KF036165	Russia	2011
630	LC228363	Japan	2016

631	MG181693	Malawi	2014
632	AB669010	Japan	2008
633	MH473476	Russia	2012
634	HM066134	Brazil	2006
635	KM116031	USA	2013
636	MG181572	Malawi	2013
637	KF812697	South Korea	2008
638	KU356578	Bangladesh	2010
639	KC618425	South Africa	2008
640	MF997039	USA	2015
641	KP882078	Bangladesh	2007
642	KP882221	Bangladesh	2008
643	DQ492673	Bangladesh	2003
644	MG181660	Malawi	2013
645	KF812687	South Korea	2012
646	HM123827	Brazil	2006
647	JN706552	Thailand	2011
648	LC169955	Thailand	2014
649	KX638578	India	2011
650	KT988264	Italy	2013
651	MG181770	Malawi	2012
652	LC019089	Myanmar	2011
653	HQ609562	India	2006
654	HQ611021	Russia	2010
655	GU390447	Russia	2008
656	KF716337	USA	2011
657	KX632280	Uganda	2013
658	EU805774	Ecuador	2006
659	GU390450	Russia	2008
660	MH473469	Russia	2011
661	KU199285	Bangladesh	2013
662	MG181869	Malawi	2012
663	JN706550	Thailand	2010
664	GQ414544	Germany	2009
665	KJ721706	Brazil	2005
666	KX778585	China	2015
667	KX638588	India	2012
668	GU390418	Russia	2004
669	LC228330	Japan	2016
670	JN129107	Nicaragua	2010
671	HM467951	USA	2011
672	KC178804	Italy	2008
673	KX638631	India	2011
674	KC834701	Australia	1999
675	HQ611027	Russia	2010
676	HQ611030	Russia	2010
677	JQ248937	Russia	2005
678	LC103007	Japan	2014
679	JQ818159	Russia	2003
680	JQ230080	Russia	2009
681	GQ477085	Russia	2006
682	LC105469	Japan	2014
683	KF036145	Russia	2011

684	KF812709	South Korea	2011
685	KP170813	Brazil	2009
686	KC890879	Greece	2009
687	KX638580	India	2011
688	MH569768	Brazil	2016
689	KP882639	Ghana	2008
690	KT988253	Italy	2013
691	KP883046	Mali	2008
692	KY657844	China	2016
693	KF812685	South Korea	2011
694	HQ611025	Russia	2010
695	MH473489	Russia	2012
696	KX356106	Japan	2012
697	JN605441	South Africa	1999
698	JQ230072	Russia	2009
699	KF812686	South Korea	2011
700	KJ721703	Brazil	2007
701	KU048617	Italy	2016
702	KX638650	India	2012
703	LC105165	Japan	2014
704	HQ611013	Russia	2010
705	GQ468564	Russia	2005
706	KF716343	USA	2011
707	KP883211	Mali	2008
708	LC019100	Myanmar	2011
709	KX265694	Egypt	2012
710	HM467949	USA	2011
711	KX638563	India	2009
712	KP882430	Ghana	2009
713	LC105194	Japan	2014
714	GU390425	Russia	2005
715	KX778595	China	2016
716	KX638596	India	2013
717	LC105046	Japan	2014
718	LC102985	Japan	2014
719	JX841180	Russia	2011
720	GU199521	Bangladesh	2001
721	LC105414	Japan	2014
722	LC228385	Japan	2016
723	KP882287	Bangladesh	2008
724	KP882364	Ghana	2009
725	KP883178	Mali	2008
726	LC105247	Japan	2014
727	MG181506	Malawi	2012
728	JQ230098	Russia	2011
729	GU390433	Russia	2005
730	GU390427	Russia	2005
731	HM066151	Brazil	2008
732	KJ820851	Brazil	1997
733	KX778591	China	2015
734	KX638613	India	2009
735	KP882232	Bangladesh	2009
736	KP882309	Ghana	2008

737	JQ230074	Russia	2009
738	HQ611020	Russia	2010
739	KP881946	Bangladesh	2008
740	MG181517	Malawi	2012
741	LC105143	Japan	2014
742	LC102919	Japan	2014
743	JQ585599	Russia	2007
744	JQ230062	Russia	2004
745	KP882958	Mali	2008
746	KY657832	China	2016
747	HM066148	Brazil	2008
748	KX632291	Uganda	2013
749	LC066643	Thailand	2013
750	LC086762	Thailand	2013
751	KC713889	Russia	2012
752	JX867350	Russia	2010
753	HQ611004	Russia	2010
754	HQ611000	Russia	2010
755	MG676121	India	2014
756	KP883200	Mali	2008
757	JN706540	Thailand	2008
758	KU356633	Bangladesh	2013
759	KF185107	Brazil	2012
760	KC618432	South Africa	2008
761	KU048625	Italy	2016
762	KX638600	India	2009
763	KF036156	Russia	2008
764	JN706536	Thailand	2009
765	MG181286	Malawi	2000
766	LC105305	Japan	2014
767	HM123816	Brazil	2005
768	KP881968	Bangladesh	2008
769	KJ870904	Democratic Republic of the Congo	2008
770	MG676130	India	2015
771	MG676136	India	2015
772	JQ230096	Russia	2010
773	KC713895	Russia	2012
774	KX638647	India	2011
775	KX638608	India	2012
776	MG181418	Malawi	2004
777	KF812714	South Korea	2011
778	GU390455	Russia	2009
779	JN706546	Thailand	2010
780	HQ702256	USA	2009
781	MG181913	Malawi	2012
782	HQ609561	India	2006
783	MH473495	Russia	2012
784	KX638610	India	2013
785	LC105382	Japan	2014
786	GQ477115	Russia	2008
787	HQ702234	USA	2009
788	MG676124	India	2015
789	MG181649	Malawi	2013

790	KU714459	Malawi	2000
791	LC102897	Japan	2014
792	MH473475	Russia	2012
793	HM066129	Brazil	2005
794	HM066133	Brazil	2006
795	HQ611028	Russia	2010
796	KC020043	Russia	2009
797	LC105154	Japan	2014
798	KF812688	South Korea	2012
799	KP170806	Brazil	2008
800	KP170808	Brazil	2008
801	KC618428	South Africa	2008
802	HM467950	USA	2011
803	MG181451	Malawi	2005
804	HM123838	Brazil	2008
805	KX638558	India	2009
806	KX638625	India	2010
807	KF812668	South Korea	2009
808	HQ738600	Russia	2009
809	MH473477	Russia	2012
810	MH060082	Thailand	2013
811	HM123826	Brazil	2006
812	HM066135	Brazil	2006
813	AB796453	Japan	2012
814	KC178801	Italy	2006
815	JQ230087	Russia	2010
816	KP007172	Philippines	2012
817	KX356115	Japan	2012
818	KY616897	Japan	2013
819	GU390442	Russia	2008
820	HQ738590	Russia	2008
821	GQ477086	Russia	2006
822	JN706551	Thailand	2011
823	MG181484	Malawi	2008
824	KX655455	Uganda	2013
825	JN014002	South Africa	2008
826	LC055551	Thailand	2012
827	JQ818158	Russia	2003
828	JQ230085	Russia	2010
829	HQ611033	Russia	2010
830	HM066140	Brazil	2008
831	KT694943	USA	1974
832	JN129102	Nicaragua	2010
833	KX638583	India	2012
834	KX638636	India	2013
835	MG181638	Malawi	2013
836	MG181671	Malawi	2013
837	KF812692	South Korea	2011
838	MH473482	Russia	2012
839	MH473488	Russia	2012
840	KX171564	South Korea	2013
841	KT988176	Italy	2009
842	GU390417	Russia	2004

843	GQ477122	Russia	2009
844	LC105013	Japan	2014
845	LC105480	Japan	2014
846	JQ951795	Russia	2010
847	KJ094938	China	2010
848	GQ477084	Russia	2006
849	HQ738586	Russia	2008
850	LC406823	Kenya	2012
851	KU048634	Italy	2016
852	LC086795	Thailand	2013
853	KY657845	China	2017
854	JQ818152	Russia	2008
855	KF812706	South Korea	2010
856	KF812674	South Korea	2012
857	HQ738598	Russia	2009
858	KX356105	Japan	2012
859	KP882397	Ghana	2009
860	KP882485	Ghana	2009
861	KT988154	Italy	2005
862	KY657837	China	2016
863	MH473491	Russia	2012
864	JN706537	Thailand	2009
865	JQ230099	Russia	2011
866	GU390438	Russia	2005
867	GU390428	Russia	2005
868	JN706549	Thailand	2010
869	KF740531	India	2008
870	KX638599	India	2010
871	KX638582	India	2011
872	MH569767	Brazil	2015
873	MH473502	Russia	2012
874	GQ477131	Russia	2009
875	JF460827	Belgium	2009
876	KX638644	India	2010
877	KX638572	India	2010
878	KJ870893	Democratic Republic of the Congo	2010
879	HQ611018	Russia	2010
880	KP710857	Russia	2014
881	KP882342	Ghana	2009
882	MG676139	India	2016
883	KF371692	China	2005
884	KU199274	Bangladesh	2010
885	KP007150	Philippines	2012
886	MG181605	Malawi	2013
887	KY657846	China	2017
888	HM066154	Brazil	2008
889	KP007161	Philippines	2012
890	KX638616	India	2013
891	KP882045	Bangladesh	2008
892	KY000548	Germany	2016
893	KP882441	Ghana	2009
894	LC105447	Japan	2014
895	JX867345	Russia	2009

896	HQ738591	Russia	2008
897	KX638560	India	2009
898	KP882166	Bangladesh	2008
899	MG181594	Malawi	2013
900	HQ611016	Russia	2010
901	KF716344	USA	2011
902	KF036153	Russia	2012
903	KP645326	Australia	2011
904	MG181891	Malawi	2012
905	MG181759	Malawi	2014
906	GU390456	Russia	2009
907	JN706543	Thailand	2009
908	KF360212	Brazil	2010
909	KC178800	Italy	2004
910	KP882199	Bangladesh	2009
911	MG181495	Malawi	2012
912	KX638605	India	2017
913	MG181528	Malawi	2012
914	AB861960	Kenya	2010
915	JQ585603	Russia	2010
916	GQ477090	Russia	2007
917	GU138213	Russia	2007
918	KP883013	Mali	2008
919	MG181341	Malawi	2002
920	LC105491	Japan	2014
921	LC169900	Thailand	2013
922	KX655466	Uganda	2013
923	GU390440	Russia	2008
924	HQ611007	Russia	2010
925	GQ477106	Russia	2007
926	GQ477108	Russia	2007
927	MG599531	Brazil	2013
928	KC618431	South Africa	2008
929	KP882848	Mali	2008
930	LC105502	Japan	2014
931	KF036175	Russia	2012
932	LC169988	Thailand	2014
933	KX638642	India	2009
934	KC713891	Russia	2012
935	JX867361	Russia	2010
936	HQ611012	Russia	2010
937	JQ230065	Russia	2006
938	GQ477096	Russia	2008
939	MG676133	India	2015
940	KP883156	Mali	2008
941	JQ818162	Russia	2003
942	KF036159	Russia	2010
943	KF036163	Russia	2011
944	KF812699	South Korea	2010
945	KJ094940	China	2011
946	KF447843	China	2010
947	LC406801	Kenya	2014
948	KJ870926	Democratic Republic of the Congo	2009

949	KP883079	Mali	2008
950	JQ818157	Russia	2003
951	KF812702	South Korea	2012
952	JQ230064	Russia	2006
953	MH473478	Russia	2012
954	HM066136	Brazil	2006
955	KU248398	Bangladesh	2010
956	KJ721700	Brazil	2009
957	JN591412	Malawi	2012
958	KX638635	India	2013
959	MG181363	Malawi	2002
960	JX867343	Russia	2007
961	JX185762	Italy	2007
962	LC169878	Thailand	2013
963	KX638624	India	2009
964	KX638557	India	2009
965	KC890877	Greece	2009
966	KC618423	South Africa	2008
967	KP882903	Mali	2008
968	KP883057	Mali	2009
969	MG181858	Malawi	2012
970	GQ477123	Russia	2009
971	MH473498	Russia	2012
972	KY616890	Japan	2013
973	MH484054	Russia	2012
974	HM066128	Brazil	2005
975	HM066143	Brazil	2008
976	KX171559	South Korea	2010
977	KX778594	China	2015
978	KF716339	USA	2011
979	MG181748	Malawi	2014
980	LC105294	Japan	2014
981	MF580849	China	2012
982	HM123833	Brazil	2007
983	KX536665	India	2010
984	HQ657164	South Africa	2009
985	LC105205	Japan	2014
986	KF812689	South Korea	2012
987	GQ477118	Russia	2009
988	KP882133	Bangladesh	2008
989	KT988286	Italy	2013
990	LC103018	Japan	2014
991	KF036169	Russia	2011
992	JX867357	Russia	2010
993	KX171565	South Korea	2013
994	KX638564	India	2010
995	KP882408	Ghana	2007
996	KY657840	China	2016
997	JQ230082	Russia	2009
998	JQ248939	Russia	2010
999	JX867365	Russia	2010
1000	JQ818151	Russia	2008
1001	HQ738593	Russia	2008

1002	HQ611029	Russia	2010
1003	JN129105	Nicaragua	2010
1004	JN129109	Nicaragua	2010
1005	KX638633	India	2012
1006	KU059770	Australia	2013
1007	KP882782	Mali	2008
1008	GQ477089	Russia	2007
1009	KU048633	Italy	2016
1010	HQ445975	Russia	2007
1011	GQ477116	Russia	2009
1012	HM066139	Brazil	2007
1013	JN591409	Malawi	2012
1014	LC228407	Japan	2016
1015	KX356116	Japan	2012
1016	KC178806	Italy	2011
1017	MG181704	Malawi	2014
1018	LC105057	Japan	2014
1019	KF812707	South Korea	2010
1020	KF812675	South Korea	2012
1021	HM066145	Brazil	2008
1022	KF041434	China	2010
1023	KX356108	Japan	2012
1024	MG181473	Malawi	2005
1025	KF812665	South Korea	2008
1026	HQ738589	Russia	2008
1027	GQ477087	Russia	2006
1028	KU048631	Italy	2016
1029	KJ162413	Russia	2012
1030	JN129098	Nicaragua	2010
1031	KP882463	Ghana	2009
1032	KP882584	Ghana	2009
1033	GU390453	Russia	2009
1034	JN706548	Thailand	2011
1035	KP882023	Bangladesh	2008
1036	MG181374	Malawi	2002
1037	JQ230068	Russia	2009
1038	KF812694	South Korea	2011
1039	GQ477132	Russia	2009
1040	KX638561	India	2009
1041	KX638571	India	2010
1042	KF907298	Brazil	2008
1043	MG676138	India	2015
1044	HM066157	Brazil	2009
1045	FJ747617	Germany	2008
1046	KP882331	Ghana	2009
1047	HQ611015	Russia	2010
1048	GQ477103	Russia	2008
1049	KC618422	South Africa	2008
1050	LC169966	Thailand	2014
1051	KF716345	USA	2011
1052	KX638619	India	2012
1053	KP882870	Mali	2008
1054	JQ230100	Russia	2011

1055	KU048615	Italy	2016
1056	KP882375	Ghana	2009
1057	KY657818	China	2017
1058	GQ477095	Russia	2007
1059	LC105068	Japan	2014
1060	GQ468567	Russia	2005
1061	KF036132	Russia	2003
1062	JQ230078	Russia	2009
1063	GU390435	Russia	2005
1064	GU138212	Russia	2005
1065	GU390459	Russia	2009
1066	KP258402	Belgium	2012
1067	MG676127	India	2015
1068	KP883112	Mali	2008
1069	KF812700	South Korea	2010
1070	HQ610999	Russia	2010
1071	KP170810	Brazil	2009
1072	KP882188	Bangladesh	2008
1073	JQ230090	Russia	2010
1074	KF812710	South Korea	2011
1075	KX638649	India	2012
1076	KP882210	Bangladesh	2008
1077	GU947705	China	2008
1078	JQ613167	Russia	2010
1079	KF036170	Russia	2012
1080	KC890880	Greece	2009
1081	KX638555	India	2009
1082	KX638604	India	2010
1083	KP883068	Mali	2008
1084	KC713890	Russia	2012
1085	KJ126828	China	2010
1086	HQ611017	Russia	2010
1087	GQ477105	Russia	2007
1088	HM066158	Brazil	2009
1089	JQ585604	Russia	2010
1090	LC228396	Japan	2016
1091	KP882716	Kenya	2008
1092	KY657834	China	2016
1093	GU390431	Russia	2005
1094	KP881902	Bangladesh	2008
1095	KF812712	South Korea	2011
1096	GU390441	Russia	2008
1097	HQ611008	Russia	2010
1098	EU984108	India	2006
1099	KU048622	Italy	2016
1100	KP882496	Ghana	2008
1101	LC105132	Japan	2014
1102	KY657829	China	2017
1103	KY657836	China	2016
1104	JN706538	Thailand	2009
1105	KX638641	India	2009
1106	KX638594	India	2013
1107	KT988198	Italy	2009

1108	LC105349	Japan	2014
1109	AB741653	Japan	2011
1110	KP882705	Kenya	2009
1111	LC102963	Japan	2014
1112	KF812698	South Korea	2009
1113	GQ477079	Russia	2006
1114	HQ445978	Russia	2008
1115	KU356622	Bangladesh	2013
1116	KX638556	India	2009
1117	JQ818156	Russia	2003
1118	HM123832	Brazil	2007
1119	KF036162	Russia	2010
1120	MH473493	Russia	2012
1121	HM123820	Brazil	2006
1122	LC169867	Thailand	2013
1123	LC169911	Thailand	2013
1124	JN605408	Cameroon	2009
1125	KP882012	Bangladesh	2008
1126	KX880416	Germany	2015
1127	KP882936	Mali	2008
1128	GU138215	Russia	2007
1129	GQ477117	Russia	2009
1130	GQ477119	Russia	2009
1131	AB848006	Japan	2012
1132	MG181561	Malawi	2012
1133	JX867342	Russia	2007
1134	GQ477126	Russia	2009
1135	KX638623	India	2012
1136	KP882298	Ghana	2008
1137	KP882837	Mali	2008
1138	GU390448	Russia	2008
1139	MH484053	Russia	2012
1140	HM066137	Brazil	2007
1141	LC019078	Myanmar	2011
1142	HQ609563	India	2006
1143	KF812678	South Korea	2013
1144	HQ611022	Russia	2010
1145	MH473499	Russia	2012
1146	LC228352	Japan	2016
1147	KX536666	India	2011
1148	KX638638	India	2013
1149	KT988275	Italy	2013
1150	MG181814	Malawi	2012
1151	JX867368	Russia	2011
1152	KF812680	South Korea	2009
1153	GQ468561	Russia	2004
1154	KF371702	China	2004
1155	JN129104	Nicaragua	2010
1156	KX778588	China	2015
1157	KX638630	India	2011
1158	KX638585	India	2012
1159	KP883123	Mali	2008
1160	KY657820	China	2016

1161	JN706553	Thailand	2009
1162	KX638577	India	2011
1163	KP882595	Ghana	2009
1164	HQ738594	Russia	2008
1165	HM066142	Brazil	2008
1166	GU390451	Russia	2008
1167	JX841143	Russia	2011
1168	EF554119	Belgium	2011
1169	MG181550	Malawi	2012
1170	JQ230089	Russia	2010
1171	JX867358	Russia	2010
1172	KF812672	South Korea	2011
1173	KX171566	South Korea	2014
1174	LC019067	Myanmar	2011
1175	JQ818165	Russia	2004
1176	KF036144	Russia	2010
1177	HM123839	Brazil	2008
1178	KX356107	Japan	2012
1179	KP882529	Ghana	2007
1180	GQ468557	Russia	2004
1181	GQ477088	Russia	2006
1182	GQ477128	Russia	2009
1183	MG781035	Thailand	2010
1184	KJ721705	Brazil	2006
1185	JN014006	South Africa	2008
1186	LC105436	Japan	2014
1187	JQ248940	Russia	2010
1188	HM066144	Brazil	2008
1189	KP882056	Bangladesh	2008
1190	KP882507	Ghana	2008
1191	MG181836	Malawi	2012
1192	HQ738588	Russia	2008
1193	MH473486	Russia	2012
1194	KX171560	South Korea	2011
1195	KP882980	Mali	2008
1196	MG181825	Malawi	2012
1197	GQ468563	Russia	2005
1198	GQ477100	Russia	2008
1199	KU248420	Bangladesh	2010
1200	JX841168	Russia	2011
1201	KF716340	USA	2011
1202	KX638614	India	2010
1203	KF661454	Russia	2012
1204	KX356118	Japan	2012
1205	KP883090	Mali	2008
1206	KY657819	China	2016
1207	GU390422	Russia	2004
1208	KT695053	USA	1989
1209	KX655521	Uganda	2013
1210	MH473471	Russia	2012
1211	KU048618	Italy	2016
1212	GU390432	Russia	2005
1213	GU390460	Russia	2009

1214	KJ820884	Brazil	1998
1215	MH933786	Cameroon	2014
1216	KX638567	India	2010
1217	KX638653	India	2013
1218	MG181715	Malawi	2014
1219	MG181352	Malawi	2002
1220	MG181242	Malawi	1999
1221	HQ657153	South Africa	2009
1222	KX171567	South Korea	2014
1223	AB930193	Japan	2014
1224	KP882353	Ghana	2009
1225	MG181803	Malawi	2012
1226	MG181297	Malawi	2000
1227	KR632622	Italy	2012
1228	KY657848	China	2017
1229	KF907299	Brazil	2010
1230	KP882265	Bangladesh	2008
1231	KX880427	Germany	2016
1232	GU947708	China	2008
1233	JQ230069	Russia	2009
1234	KC618429	South Africa	2008
1235	JF460816	USA	2006
1236	EF560707	Bangladesh	2001
1237	KP882969	Mali	2008
1238	MG181902	Malawi	2012
1239	KX655488	Uganda	2013
1240	JQ585601	Russia	2009
1241	KC713893	Russia	2012
1242	KJ721697	Brazil	2011
1243	KP881957	Bangladesh	2008
1244	KY657826	China	2016
1245	JQ230075	Russia	2009
1246	MH484058	Russia	2013
1247	KX638597	India	2013
1248	LC086729	Thailand	2013
1249	HQ611005	Russia	2010
1250	LC386069	Japan	2017
1251	KU048620	Italy	2016
1252	KU128896	Belgium	2013
1253	MH704728	Indonesia	2014
1254	KY657831	China	2016
1255	AB605600	Viet Nam	2006
1256	MG571800	Venezuela	2015
1257	KF036133	Russia	2003
1258	GU390454	Russia	2009
1259	JN706545	Thailand	2010
1260	KP881990	Bangladesh	2008
1261	KP882386	Ghana	2009
1262	MG676122	India	2014
1263	KY657839	China	2017
1264	KC713894	Russia	2012
1265	JN706539	Thailand	2009
1266	KX638607	India	2011

1267	KP881979	Bangladesh	2009
1268	JQ230091	Russia	2010
1269	JX867353	Russia	2010
1270	JQ230063	Russia	2005
1271	KX778617	Dominican Republic	2013
1272	KP882672	Ghana	2007
1273	MG676135	India	2015
1274	KP882881	Mali	2008
1275	LC103029	Japan	2014
1276	KF812703	South Korea	2012
1277	KJ094941	China	2011
1278	JQ613193	Russia	2011
1279	HQ702245	USA	2009
1280	KF716348	USA	2011
1281	LC105513	Japan	2014
1282	KY657850	China	2017
1283	MG181264	Malawi	2000
1284	JX965143	Australia	2010
1285	MF580852	China	2015
1286	HM123823	Brazil	2006
1287	KF716347	USA	2011
1288	KP882562	Ghana	2008
1289	KP882815	Mali	2008
1290	KJ803839	Paraguay	2004
1291	MH473472	Russia	2012
1292	HM066130	Brazil	2006
1293	HM123841	Brazil	2008
1294	KX265706	Egypt	2012
1295	LC228341	Japan	2016
1296	KU048623	Italy	2016
1297	AB861971	Kenya	2010
1298	GQ477129	Russia	2009
1299	KJ820840	Brazil	1997
1300	KU059792	Australia	2014
1301	MG181308	Malawi	2000
1302	LC105458	Japan	2014
1303	KF036137	Russia	2010
1304	GU390416	Russia	2003
1305	KP170809	Brazil	2008
1306	KC178802	Italy	2007
1307	KP882001	Bangladesh	2008
1308	KP882540	Ghana	2008
1309	KF036167	Russia	2011
1310	KF812679	South Korea	2013
1311	KU048636	Italy	2016
1312	MG181440	Malawi	2004
1313	LC105227	Japan	2014
1314	KF812669	South Korea	2010
1315	GU390445	Russia	2007
1316	AB796450	Japan	2012
1317	JN605452	Kenya	2010
1318	JN605430	South Africa	2010
1319	KT988242	Italy	2012

1320	AB669014	Japan	2008
1321	MF580847	China	2010
1322	KP170811	Brazil	2009
1323	KU356644	Bangladesh	2013
1324	KJ721698	Brazil	2010
1325	KP883002	Mali	2008
1326	KF036141	Russia	2010
1327	JQ818168	Russia	2010
1328	KF812705	South Korea	2013
1329	GQ477110	Russia	2008
1330	KX778592	China	2015
1331	KP882771	Mali	2008
1332	MG181396	Malawi	2003
1333	KF500523	USA	2012
1334	LC019056	Myanmar	2011
1335	HQ738585	Russia	2008
1336	KU048635	Italy	2016
1337	HM066147	Brazil	2008
1338	KP882089	Bangladesh	2008
1339	LC066654	Thailand	2013
1340	KY657823	China	2017
1341	KF812691	South Korea	2013
1342	JX195078	Italy	2010
1343	KX638574	India	2011
1344	KT988143	Italy	2005
1345	JN014005	South Africa	2008
1346	LC102996	Japan	2014
1347	KF812681	South Korea	2009
1348	JN129101	Nicaragua	2010
1349	KX356110	Japan	2012
1350	KX638584	India	2012
1351	LC105079	Japan	2014
1352	JX867355	Russia	2010
1353	KF036161	Russia	2010
1354	MH473485	Russia	2012
1355	KX171561	South Korea	2011
1356	KT369175	Italy	2011
1357	JN605419	Zimbabwe	2009
1358	JQ818155	Russia	2010
1359	HQ738595	Russia	2008
1360	JX841144	Russia	2011
1361	KJ094937	China	2009
1362	GQ477081	Russia	2006
1363	GQ468558	Russia	2004
1364	AB605599	Viet Nam	2006
1365	KP882474	Ghana	2008
1366	KP883134	Mali	2008
1367	KF447843	China	2014
1368	KF447854	China	2014
1369	U984108	India	2011
1370	MW384411	China	2022
1371	MW384400	China	2022
1372	MW384389	China	2022

1373	MW384378	China	2022
1374	MW384367	China	2022
1375	MW384356	China	2022
1376	MW384345	China	2022
1377	MW384334	China	2022
1378	MW384323	China	2022

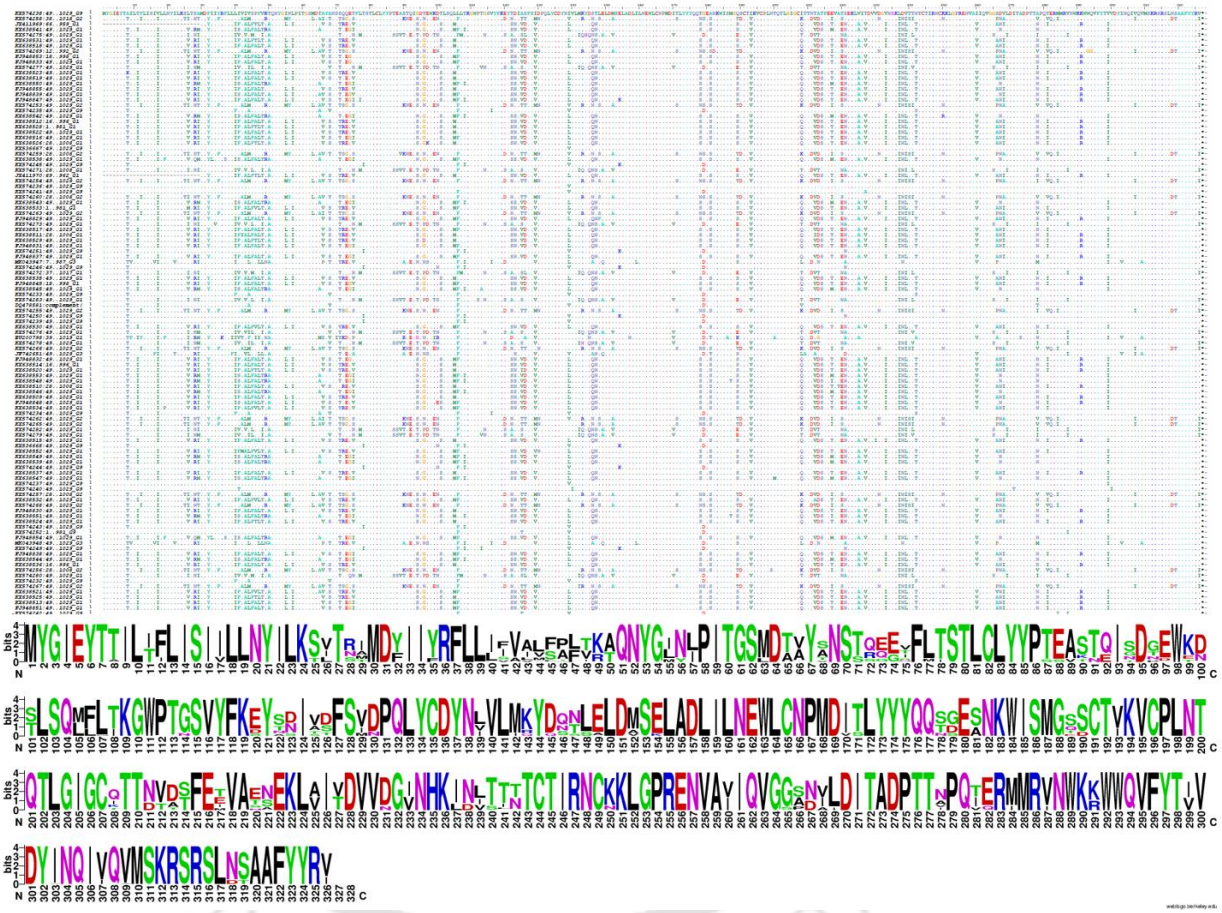


Figure A.1. Visualization of multiple sequence alignment by BioEdit v. 7.0.9 (top) and sequence logo created by WebLogo3 (bottom) for RVA VP7 protein sequences. The conserved sites have been depicted as plotting identities as a dot using BioEdit v. 7.0.9. The standard logo plot stacks characters vertically at each alignment position, reflecting their relative frequencies with proportional heights in WebLogo3.

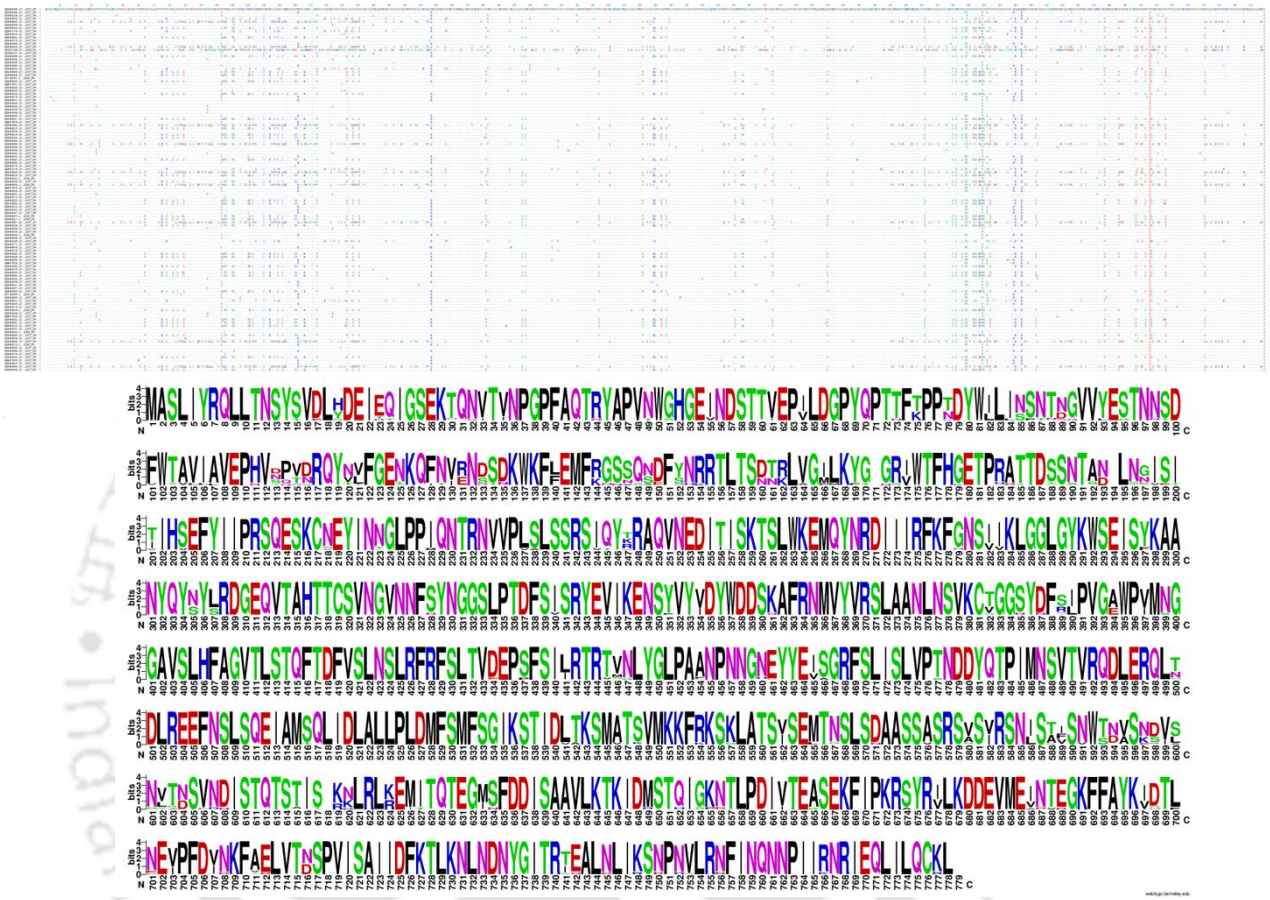


Figure A.2. Visualization of multiple sequence alignment by BioEdit v. 7.0.9 (top) and sequence logo created by WebLogo3 (bottom) for RVA VP4 protein sequences. The conserved sites have been depicted as plotting identities as a dot using BioEdit v. 7.0.9. The standard logo plot stacks characters vertically at each alignment position, reflecting their relative frequencies with proportional heights in WebLogo3.

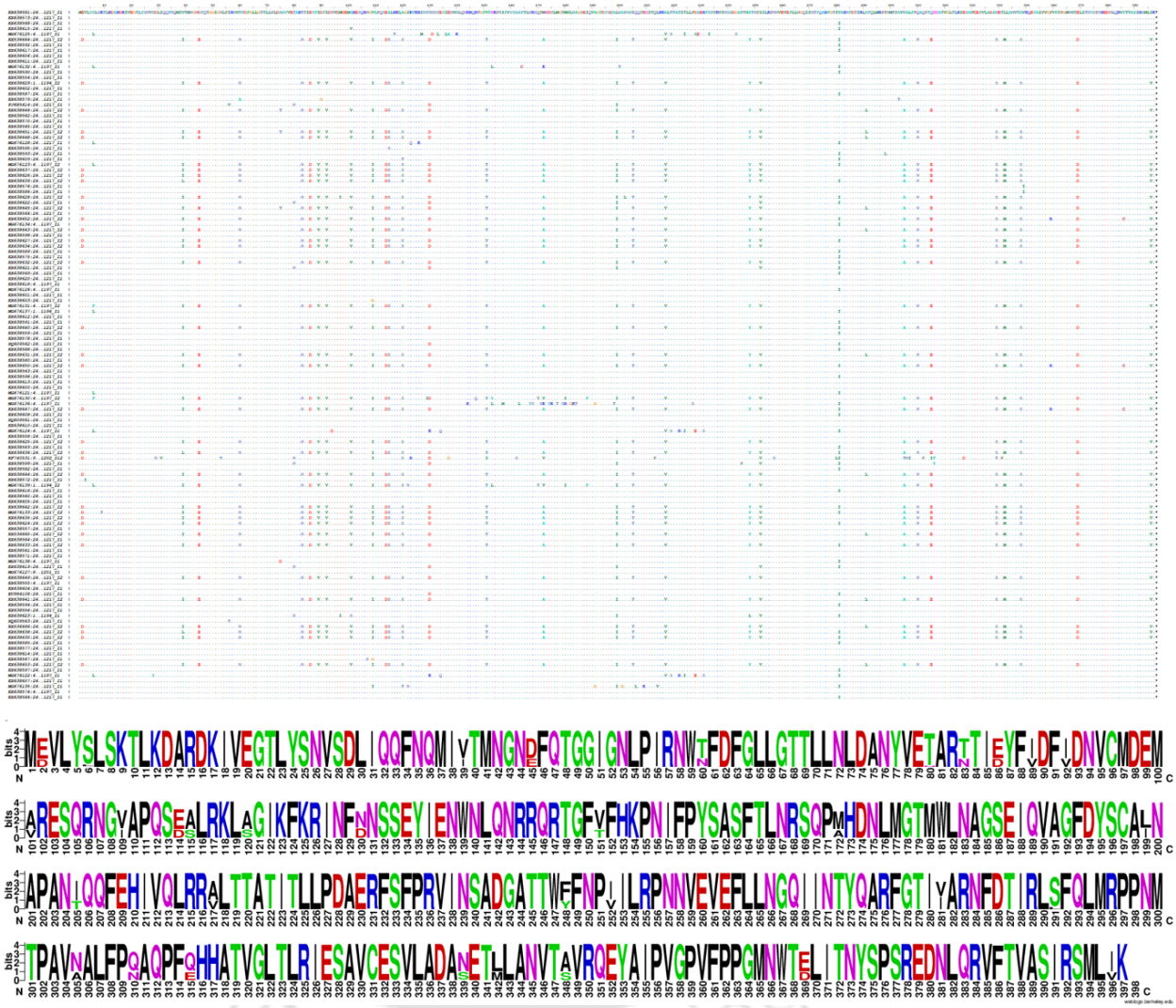


Figure A.3. Visualization of multiple sequence alignment by BioEdit v. 7.0.9 (top) and sequence logo created by WebLogo3 (bottom) for RVA VP6 protein sequences. The conserved sites have been depicted as plotting identities as a dot using BioEdit v. 7.0.9. The standard logo plot stacks characters vertically at each alignment position, reflecting their relative frequencies with proportional heights in WebLogo3.

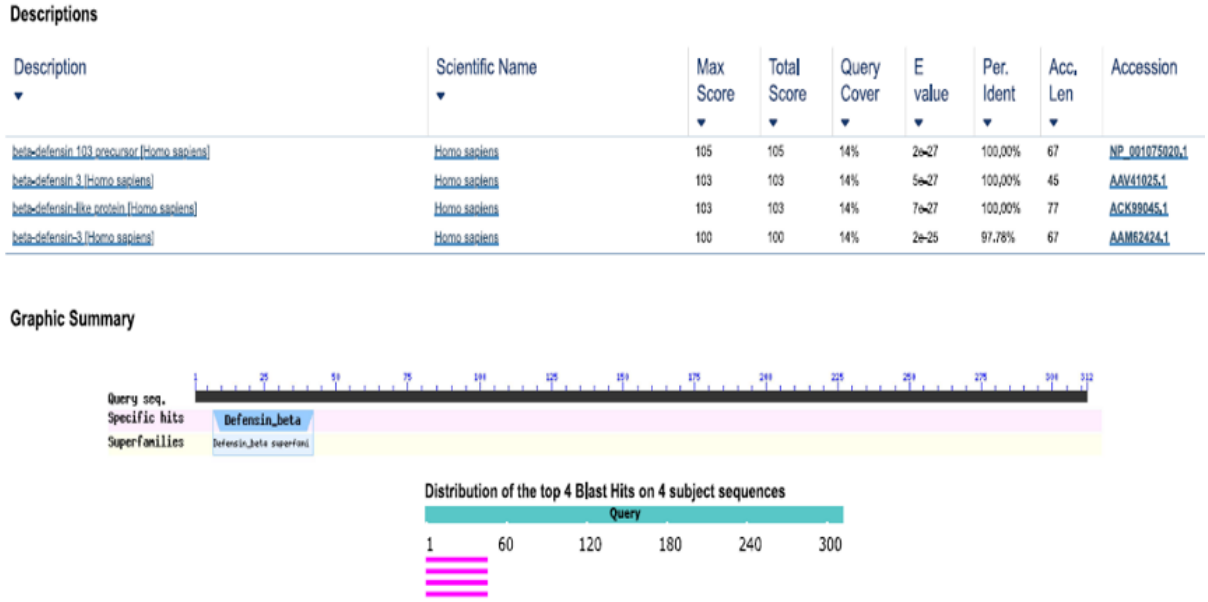


Figure A.4. BlastP analysis: Homology assessment of RVAMEV against *Homo sapiens* (taxid 9606), *Lactobacillus johnsonii* (taxid 33959), *Lactobacillus rhamnosus* (taxid 47715), and the *Lactobacillus casei* group (taxid 655183).

List of Publications

A. In Peer-Reviewed Journals:

1. **Pooja Rani Kuri** and Pranab Goswami. Unravelling Aggregation Propensity of Rotavirus A VP6 Expressed as *E. coli* Inclusion Bodies through In Silico Prediction. ***Scientific Reports.*** 14(1):21464. 2024. doi: 10.1038/s41598-024-69896-1 (Part of Chapter 3).

2. **Pooja Rani Kuri** and Pranab Goswami. Reverse Vaccinology-Based Multi-Epitope Vaccine Design Against Indian Group A Rotavirus Targeting VP7, VP4, and VP6 Proteins. ***Microbial Pathogenesis***, 193, 106775. 2024. doi:10.1016/j.micpath.2024.106775 (Part of Chapter 2).

3. **Pooja Rani Kuri** and Pranab Goswami. Current Update on Rotavirus in-Silico Multiepitope Vaccine Design. ***ACS omega***, 8,1 190-207. 2022. doi:10.1021/acsomega.2c07213 (Part of Chapter 1).

4. Phurpa Dema Thungon, **Pooja Rani Kuri**, Vinay Bachu, Pranab Goswami, Silk-fibroin film as enzyme stabilizing material and optical signal transducer for developing alcohol oxidase-based μPAD methanol biosensor. ***Biosensors and Bioelectronics X***, Volume 11, 2022, 100147, ISSN 2590-1370. 2022. doi:10.1016/j.biosx.2022.100147

5. Malaya Mili, Vinay Bachu, **Pooja Rani Kuri**, Naveen Kumar Singh and Pranab Goswami. Improving synthesis and binding affinities of nucleic acid aptamers and their therapeutics and diagnostic applications. ***Biophysical chemistry***, 309, 107218. 2024. doi:10.1016/j.bpc.2024.107218

B. Manuscript Communicated:

1. **Pooja Rani Kuri** and Pranab Goswami. Development of a Molecular Beacon-Based Genosensor for Detection of Human Rotavirus A. (Under review)

C. Book Chapter

1. **Pooja Rani Kuri**, Priyanki Das, and Pranab Goswami, 2020. "Fundamentals of Biosensors", Advanced Materials and Techniques for Biosensors and Bioanalytical Applications. CRC Press, p. 143-161. (Part of Chapter 1)

D. Abstracts Published in Conferences

1. **Pooja Rani Kuri** and Pranab Goswami. "Development of a Molecular Beacon-Based Genosensor for Detection of Human Rotavirus A", Research and Industrial Conclave Research Conclave, (2024), held during 9 to 11 August at IIT Guwahati.
2. **Pooja Rani Kuri** and Pranab Goswami. "Synthesis, characterization and application of L-glutamate derived carbon dots for the development of rotavirus genosensor", 8th International Conference on Advanced Nanomaterials and Nanotechnology (ICANN 2023), held during 29 November to 1 December, 2023 at IIT Guwahati, India.
3. **Pooja Rani Kuri** and Pranab Goswami. "Immunoinformatics mediated screening of Rotavirus A epitopes for potential peptide vaccine development", International Conference on Biotechnology for Sustainable Bioresources and Biotechnology (BSBB-2022) held during Dec 7-11, 2022 at IIT Guwahati, India.
4. **Pooja Rani Kuri** and Pranab Goswami. "Molecular Epidemiology of Rotavirus A in India: Molecular Diversity and Current Insights", International Conference on Fundamental and Applied Sciences (ICFAS) 2021 organised by Faculty of Science and IQAC from 24th March 2021 to 26th March 2021.
5. **Pooja Rani Kuri** and Pranab Goswami. "Nanomaterials-Mediated Strategies in Biosensing Technology for Virus Detection", International Conference on Advanced Materials Science and Applications (Online) organized by the Department of Physics and IQAC of Ramaiah Institute of Technology (Autonomous Institute, Affiliated to VTU) Bangalore, India during 3-4 September, 2020.
6. **Pooja Rani Kuri** and Pranab Goswami. "Current Trends and Recent Advances in Detection Strategies of Viral Gastroenteritis", 1st International E-Conference

on Recent Developments in Science, Engineering and Information Technology, ICRDSEIT-2020 organized by the Department of Computer Science, Directorate of Distance Education, Madurai Kamaraj University, Madurai – 625021 during 23-25 September, 2020.

7. **Pooja Rani Kuri** and Pranab Goswami. "Engineering of Aptamers for Analytical Applications", online International Conference on Multidisciplinary Innovations in Science and Technology (MIST-2020) on 18-19 September, 2020 organized by The Oxford College of Science, Bangalore- 560102.
8. **Pooja Rani Kuri** and Pranab Goswami. "Colorimetric Biosensing of Pathogenic Viruses and its Potential for Point-of-Care Applications", International e-Conference on Materials Processing & Characterization (ICMP&C-2020) on 18th & 19th September, 2020 organized by Department of Physics, Chaitanya Bharathi Institute of Technology, Hyderabad.
9. **Pooja Rani Kuri** and Pranab Goswami. "Aptamer-Modified Nanomaterials and their Biosensing Applications", International Webinar on Chemistry: A Motivation in Research held on 26.08.20 and 27.08.20 organized by Dept. of Chemistry, Government General Degree College, Keshiary, West Bengal, India in association with IQAC and IT Network Subcommittee.
10. **Pooja Rani Kuri** and Pranab Goswami. "Application of Biomaterials as Stabilization Matrix of Bioreceptor in Biosensor Technology", International E-Conference on Cutting Edge Smart Materials & Nanotechnology (ICCESMN-2020) held on 20-21 August 2020 jointly organized by Department of Physics and IQAC of Sadguru Gadage Maharaj College, Karad, Maharashtra, India.
11. **Pooja Rani Kuri** and Pranab Goswami. "Role of Biomaterials in Biosensor Technology" and awarded as the best poster presentation in the Two-Day International Virtual Conference on "Current Trends in Advanced Chemistry" (CTAC - 2020), organized by the Department of Chemistry, Saranathan College of Engineering, Trichy-12 from 17th to 18th August 2020. (**Best Poster Presentation**)

E. Workshops Attended

1. Advanced Entrepreneurship and Skill Development Program (E-SDP) on "Entrepreneurship Promotion in Areas of Solar, Wind, and Hybrid Energy Systems". Ministry of Micro, Small & Medium Enterprises (MSME), Govt. of India in association with School of Energy Sciences and Engineering, Indian Institute of Technology Guwahati. 29th January – 02th February 2024.
2. Participated in the workshop, "Implementation of MsME innovative scheme-Intellectual Propery Right (IPR)" held at IIT Guwahati on 22.02.2023.
3. Participated in the workshop, "Drug discovery and development" conducted between 29 Oct to 22 Nov 2022 organized by Decode Life.
4. Two Weeks International Online STTP on "Computational Techniques & Applications in Bioinformatics". MVSR Engineering College, Hyderabad, Telangana, India. 13th - 25th September 2021.
5. Participated in an online CeNSE DBT Nano-biotechnology Alliance (C-DNA) Familiarization Workshop conducted from 11th to 13th January, 2021 by Centre for Nano Science and Engineering (CeNSE), Indian Institute of Science, Bangalore.
6. Participated in "5-Days Faculty Development Programme on Trends in Biotechnology" from 15th September to 19th September, 2020 organized by Department of Biotechnology, Maulana Abul Kalam Azad University of Technology, West Bengal, India in association with the Islamic University of Science and Technology, Jammu and Kashmir, India.
7. Participated in one week AICTE sponsored STTP on "Nanotechnology and Functional Materials (NTFM-Phase-III)" held from 14-09-2020 to 19-09-2020, through online mode, organized by Department of Mechanical Engineering, S V College of Engineering, Tirupati, Andhra Pradesh,517507.
8. Participated in the TEQIP-III Sponsored One Week Faculty Development program on Recent Advances on Nanotechnology, Catalysis &Bio-Chemical Engineering. (RANCBE - 2020) held at Veer Surendra Sai University of Technology, Burla, Odisha from 16th to 20th Sep., 2020.
9. Participated in the One Week AICTE sponsored Short Term Training Programme on "Synthesis, Characterization and Applications of Nanomaterials" during 24th to 29th Aug, 2020 organized by Centre for Nano Science and Technology, Institute of Science and Technology, Hyderabad.

10. Participated in one-week Faculty Development Program on “Nano Hybrid Composite Materials Characterization & Applications” organized by Department of Mechanical Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad from 24th to 28th August 2020.
11. Attended the six days online STTP on Internet of Medical Things Enabling Technologies in Health Care sponsored by AICTE organized by Department of Electronics and Communication Engineering, Chennai Institute of Technology from 27.07.2020 to 1.08.2020.
12. Attended 3 days International e-Workshop on “Docking, QSAR and Molecular Dynamics” jointly organized by Department of Biotechnology, Ramaiah Institute of Technology and Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Ramaiah University of Applied Sciences, Bengaluru, Karnataka, India in association with IEEE-EMB MSRIT student chapter and SRIGEN-Society of Biotechnologists from 29 to 31 July 2020.
13. Participated in a 2 Day National Open Online Workshop on “Bioinformatics:Concepts,Tools And Database” Organized by Research Development Committee,K J Somaiya College of Science & Commerce,Mumbai -77 on 27th & 28th July,2020
14. Completed 5 days online workshop on Molecular docking & Biomolecular Interactions held from 06.07.2020 to 10.07.2020 organized by Centre for Bioinformatics, Computational and Systems Biology, Pathfinder Research and Training Foundation, Greater Noida, UP, India.
15. Participated in One Week International Skill Development Program on “Innovations in Natural Product Driven Drug Discovery and Analytical Chemistry” organized by Department of Pharmaocognosy & Pharmaceutical Analysis, Shri Vishnu College of Pharmacy Professionals (APP), Society of Pharmacognosy (ISP) and Mybo group from 6th to 11th July, 2020.

Front Pages of the Papers Published in International Journals

scientific reports

Check for updates

OPEN

Unravelling aggregation propensity of rotavirus A VP6 expressed as *E. coli* inclusion bodies through in silico prediction

Pooja Rani Kuri & Pranab Goswami✉

The inner capsid protein of rotavirus, VP6, emerges as a promising candidate for next-generation vaccines against rotaviruses owing to its abundance in virion particles and high conservation. However, the formation of inclusion bodies during prokaryotic VP6 expression poses a significant hurdle to rotavirus research and applications. Here, we employed experimental and computational approaches to investigate inclusion body formation and aggregation-prone regions (APRs). Heterologous recombinant VP6 expression in *Escherichia coli* BL21(DE3) cells resulted in inclusion body formation, confirmed by transmission electron microscopy revealing amorphous aggregates. Thioflavin T assay demonstrated incubation temperature-dependent aggregation of VP6 inclusion bodies. Computational predictions of APRs in rotavirus A VP6 protein were performed using sequence-based tools (TANGO, AGGRESCAN, Zyggregator, Waltz, FoldAmyloid, ANuPP, Camsol intrinsic) and structure-based tools (SolubiS, CamSol structurally corrected, Aggrescan3D). A total of 24 consensus APRs were identified, with 21 of them being surface-exposed in VP6. All identified APRs display a predominance of hydrophobic amino acids, ranging from 33 to 100%. Computational identification of these APRs corroborates our experimental observation of VP6 inclusion body or aggregate formation. Characterization of VP6's aggregation propensity facilitates understanding of its behaviour during prokaryotic expression and opens avenues for protein engineering of soluble variants, advancing research on rotavirus VP6 in pathology, therapy, and diagnostics.

Keywords Rotavirus, Protein aggregation, Aggregation prone regions, Inclusion bodies

Rotaviruses infect a wide range of vertebrates, including both birds and mammals. The serological classification of the *Rotavirus* genus encompasses ten distinct species (Rotavirus groups A–J) based on the antigenic specificity and genetic variability exhibited by the viral capsid protein VP6[1–3]. Group A rotavirus (RVA) is the primary etiological agent responsible for acute gastroenteritis in humans, predominantly afflicting the under-five population, exerting a global impact[4]. RVA infections have accounted for a substantial annual mortality toll, surpassing half a million cases, preceding rotavirus vaccine implementation[5,6].

Rotavirus VP6 has been acknowledged as a promising non-live next-generation vaccine candidate against rotavirus owing to its attributes of high abundance in the virion particle, high conservancy, and immunogenic nature conferring heterotypic protection[7,8]. Diverse vaccine modalities encoding rotavirus VP6 antigen, encompassing DNA vaccines [9–11], subunit vaccines incorporating recombinant VP6 protein[12] and self-assembled structures have been documented to elicit immune responses or confer protection in animal models[13–16]. VP6 self-assembled structures, particularly VP6 nanotubes/Virus-Like Particles (VLPs), demonstrate superior immunogenicity than that of VP6 monomers or trimers, eliciting strong immune responses[7]. Additionally, they exhibit dual functionality as adjuvants and carriers for immunological applications[17]. Advancements in the production and purification of VP6 nanotubes/VLPs have enabled exploration of their multifaceted potential as immunogens, adjuvants, delivery vehicles, and nano-biomaterials[18]. Furthermore, the rotavirus VP6 antigen is utilized as a detection biomarker for the purpose of rotavirus diagnosis[19,20].

VP6 is a structural protein, present in the middle of the triple layered capsids of the virion. Notably, rotavirus structural proteins govern host specificity, cellular entry, enzymatic machinery for viral transcript synthesis, and harbor immunological epitopes. Situated in the middle layer, 260 VP6 proteins engage in extensive interactions

Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati, Assam 781039, India. ✉email: pgoswami@iitg.ac.in

Microbial Pathogenesis 193 (2024) 106775

Contents lists available at ScienceDirect

Microbial Pathogenesis

journal homepage: www.elsevier.com/locate/micpath

Reverse vaccinology-based multi-epitope vaccine design against Indian group A rotavirus targeting VP7, VP4, and VP6 proteins

Pooja Rani Kuri, Pranab Goswami *

Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Assam, 781039, India

ARTICLE INFO

Keywords:
Rotavirus
Gastroenteritis
Phylogenetic analysis
Reverse vaccinology
Immune simulation

ABSTRACT

Rotavirus, a primary contributor to severe cases of infantile gastroenteritis on a global scale, results in significant morbidity and mortality in the under-five population, particularly in middle to low-income countries, including India. WHO-approved live-attenuated vaccines are linked to a heightened susceptibility to intussusception and exhibit low efficacy, primarily attributed to the high genetic diversity of rotavirus, varying over time and across different geographic regions. Herein, molecular data on Indian rotavirus A (RVA) has been reviewed through phylogenetic analysis, revealing G1P[8] to be the prevalent strain of RVA in India. The conserved capsid protein sequences of VP7, VP4 and VP6 were used to examine helper T lymphocyte, cytotoxic T lymphocyte and linear B-cell epitopes. Twenty epitopes were identified after evaluation of factors such as antigenicity, non-allergenicity, non-toxicity, and stability. These epitopes were then interconnected using suitable linkers and an N-terminal beta defensin adjuvant. The *in silico* designed vaccine exhibited structural stability and interactions with integrins ($\alpha v\beta 3$ and $\alpha IIb\beta 3$) and toll-like receptors (TLR2 and TLR4) indicated by docking and normal mode analyses. The immune simulation profile of the designed RVA multiepitope vaccine exhibited its potential to trigger humoral as well as cell-mediated immunity, indicating that it is a promising immunogen. These computational findings indicate potential efficacy of the designed vaccine against rotavirus infection.

1. Introduction

Globally, rotavirus is an etiologic agent of severe infantile gastroenteritis, affecting the pediatric population under five. It has a multiple host range impacting both avian and mammalian species during their early stages of life and is transmitted through fecal-oral route [1]. The rotavirus virion is nonenveloped and comprises of capsid proteins encasing the genetic material of 11 segments of dsRNA, encoding a total of twelve proteins, including six structural proteins (VP1 to VP4, VP6 and VP7) as well as six non-structural proteins (NSP1 to NSP6) [2]. Structural proteins are responsible for determining host specificity, facilitating viral entry into host cells, and orchestrating enzymatic processes necessary for the generation of viral transcripts, featuring immunological epitopes. Non-structural proteins are synthesized post-host cell entry and are involved in various functions such as genome replication and modulation of innate immune responses [2]. Rotaviruses, members of the Reoviridae family, constitute the *Rotavirus* genus. Within this genus, a serological classification into ten distinct species (Rotavirus groups A–J) has been established based on the genetic diversity and antigenic distinctiveness of the viral major capsid protein VP6. Group A rotavirus is the predominant cause of gastroenteritis in human infants, making it the central focus of the current discussion and research effort. Utilizing a binary classification system, group A rotaviruses are further delineated into P genotypes (representing protease-sensitive VP4) and G genotypes (representing glycoprotein VP7), with distinctions rooted in the sequence disparities in RNA segment 4, encoding VP4 capsid protein and segment 9, encoding VP7 capsid protien. Presently, global scientific records document the existence of 41 G genotypes and 57 P genotypes within the group A rotaviruses [3–5]. Globally, six predominant G types (G1, G2, G3, G4, G9, and G12) and three P types (P[4], P[6], and P[8]) [6–8] trigger human gastroenteritis. The strains of group A rotavirus prevalent worldwide, including G1P[8], G2P[4], G3P[8], G4P[8], G9P[8], and G12P[8], are responsible for over 90 % of diarrhea cases in humans [7–9].

Several live-attenuated rotavirus vaccines have been developed and licensed for use in immunization programs over the years. Rotashield, the inaugural rotavirus vaccine sanctioned in the United States in 1998, faced association with intussusception, leading to its prompt withdrawal

* Corresponding author.
E-mail addresses: pooja176106028@iitg.ac.in (P.R. Kuri), pgoswami@iitg.ac.in (P. Goswami).

https://doi.org/10.1016/j.micpath.2024.106775
Received 4 April 2024; Received in revised form 25 June 2024; Accepted 30 June 2024
Available online 1 July 2024

ACS OMEGA

http://pubs.acs.org/journal/acsodf

Review

Current Update on Rotavirus in-Silico Multiepitope Vaccine Design

Pooja R. Kuri and Pranab Goswami*

Cite This: ACS Omega 2023, 8, 190−207

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Rotavirus gastroenteritis is one of the leading causes of pediatric morbidity and mortality worldwide in infants and under-five populations. The World Health Organization (WHO) recommended global incorporation of the rotavirus vaccine in national immunization programs to alleviate the burden of the disease. Implementation of the rotavirus vaccination in certain regions of the world brought about a significant and consistent reduction of rotavirus-associated hospitalizations. However, the efficacy of licensed vaccines remains suboptimal in low-income countries where the incidences of rotavirus gastroenteritis continue to happen unabated. The problem of low efficacy of currently licensed oral rotavirus vaccines in low-income countries necessitates continuous exploration, design, and development of new rotavirus vaccines. Traditional vaccine development is a complex, expensive, labor-intensive, and time-consuming process. Reverse vaccinology essentially utilizes the genome and proteome information on pathogens and has opened new avenues for in-silico multiepitope vaccine design for a plethora of pathogens, promising time reduction in the complete vaccine development pipeline by complementing the traditional vaccinology approach. A substantial number of reviews on licensed rotavirus vaccines and those under evaluation are already available in the literature. However, a collective account of rotavirus in-silico vaccines is lacking in the literature, and such an account may further fuel the interest of researchers to use reverse vaccinology to expedite the vaccine development process. Therefore, the main focus of this review is to summarize the research endeavors undertaken for the design and development of rotavirus vaccines by the reverse vaccinology approach utilizing the tools of immunoinformatics.

1. INTRODUCTION

Rotavirus gastroenteritis is one of the leading causes of pediatric morbidity and mortality worldwide in infants and children under five years of age. Rotavirus, as an etiologic agent of infantile gastroenteritis, was identified after a breakthrough in 1973 when this virus was discovered in fecal samples from children representing acute diarrhea[1] and duodenal biopsy samples representing severe diarrhea.[2] Since then, rotavirus has been confirmed as a major cause of life-threatening infantile gastroenteritis in under-five globally and in the young of many avian and mammalian species.[3]

Rotavirus infection has been associated with 114 million positive cases, 24 million cases requiring outpatient visits, 2.3 million cases requiring hospitalization for the reported period of 1986 to 2000,[4] and more than 200,000 deaths in under-five globally as estimated in the year 2013.[5] In the year 2015, diarrhea mortality in under-five was estimated to be 500,000 globally,[6,7] of which 29.3% of deaths were accounted for by rotavirus infection.[7] In 2016, more than 248 million cases of under-five-associated rotavirus infection were reported, requiring 1,537,000 hospitalizations with an estimated 128,500 deaths, which accounted for about 28.8% of the total diarrhea deaths in under-five in that year globally. While rotavirus-associated mortality and morbidity rates were found to be the highest in sub-Saharan Africa, Southeast Asia, and South Asia, rotavirus-associated morbidity was also found to be magnified in high-income countries, viz., the United States, with 593,000 episodes in under-five.[8] Although the prevalence of rotavirus infection and associated rotavirus diarrhea hospitalizations is ubiquitous on a global scale, more than 90% of cases presenting fatal rotavirus infection are from low-income countries, with the probable contributing factors being limited access to proper sanitation, health care, hydration therapy, and a greater prevalence of comorbid conditions and others.[9] The introduction of rotavirus vaccination in several countries to alleviate the disease burden of rotavirus diarrhea was reported to avert an estimated 28,800 deaths following vaccination of 27.8% of under-five children in 2016. Under-five diarrhea-associated mortality has reduced by more than 45% since 2005[10] probably due to improvements in water and sanitation, a reduction in comorbid conditions like undernutrition, and the inclusion of rotavirus vaccination in several countries.

In 2009, the World Health Organization (WHO) recommended incorporation of a rotavirus vaccine in national immunization programs globally to alleviate the burden of the

Received: November 9, 2022
Accepted: December 14, 2022
Published: December 30, 2022

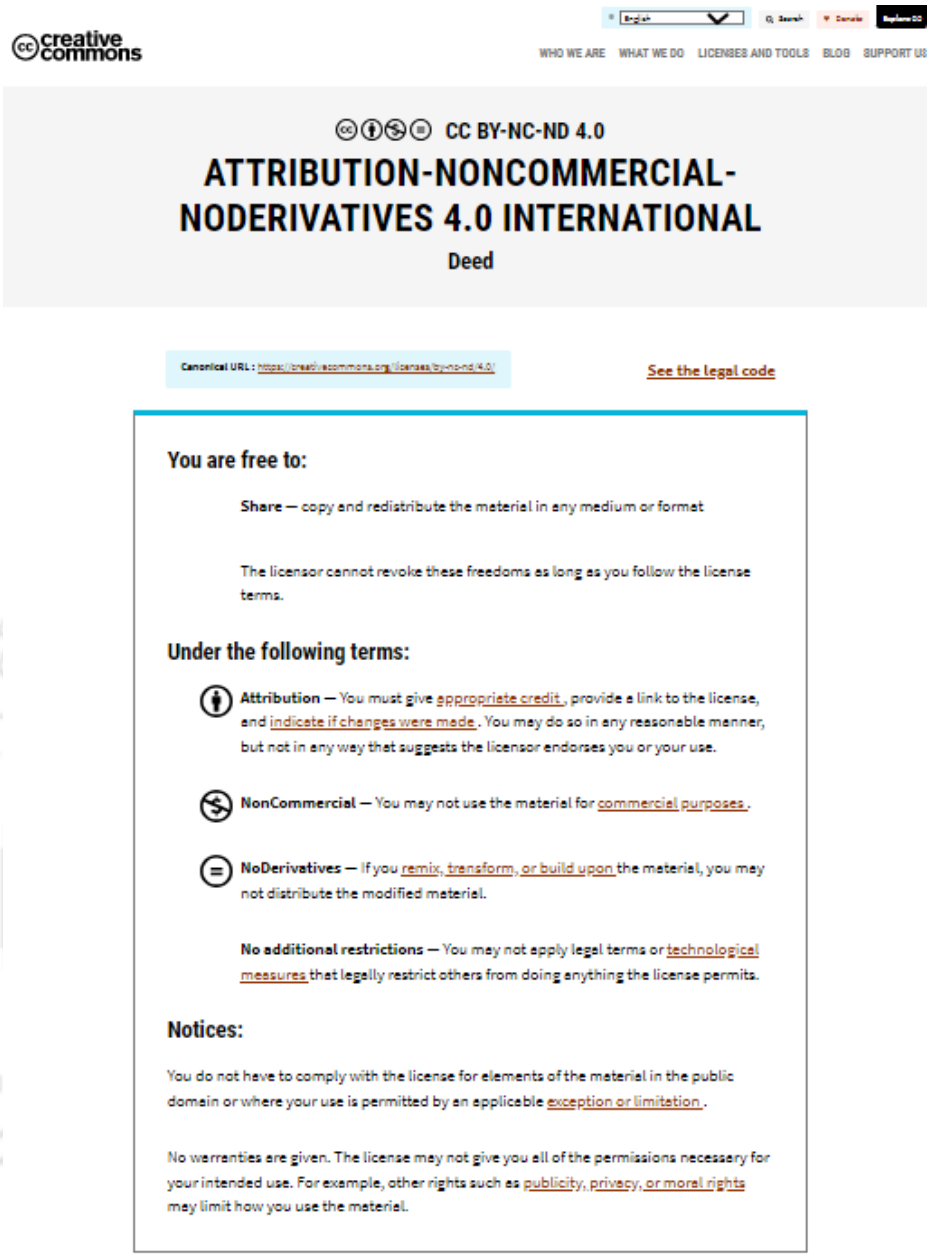
https://doi.org/10.1021/acsomega.2c07213
ACS Omega 2023, 8, 190−207

1 Fundamentals of Biosensors

Pooja Rani Kuri, Priyanki Das* and Pranab Goswami*
Indian Institute of Technology Guwahati, Assam, India (* equal contribution)

CONTENTS

Copyright Permissions



creative commons

WHO WE ARE WHAT WE DO LICENSES AND TOOLS BLOG SUPPORT US

CC BY-NC-ND 4.0

ATTRIBUTION-NONCOMMERCIAL-NODERIVATIVES 4.0 INTERNATIONAL

Deed

Canonical URL: https://creativecommons.org/licenses/by-nc-nd/4.0/

See the legal code

You are free to:

Share — copy and redistribute the material in any medium or format

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit , provide a link to the license, and indicate if changes were made . You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial — You may not use the material for commercial purposes .

NoDerivatives — If you remix, transform, or build upon the material, you may not distribute the modified material.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation .

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.