

**An Integrated Bioinformatics and Computational Biophysics
Approach to Enterovirus Surveillance and Research.**

A thesis submitted in fulfilment of the requirements of
the degree of Doctor of Philosophy

Jason Anthony Roberts

Graduate Certificate (Bioinformatics)

Bachelor of Applied Science (Laboratory Medicine)

**School of Applied Sciences
Health Innovations Research Institute
RMIT University
March 2014**

Declaration

I certify that except where due acknowledgement has been made; the work is that of the candidate alone. This body of work has not been submitted previously, in whole or in part, to qualify for any other academic award. The content of this thesis is the result of work which has been carried out since the official commencement date of the approved research program. Any editorial work, paid or unpaid, carried out by a third party is acknowledged.

Jason Anthony Roberts

Acknowledgments

I would like to thank my loving wife Katrina and my boisterously inquisitive son William for their patience during the research phase of this project and during the formulation of this thesis. Most children grow up asking “Are we there yet?”, William has grown up asking “Have you finished your book yet Daddy?”. Yes William I have, and I dedicate it to you in the hope that it continues to inspire the spark of curiosity I see in you every day.

I would also like to thank Associate Professor Bruce Thorley, Dr Andrew Hung, Dr Mike Kuiper, Professor Peter Smooker and Allen Roberts for their guidance throughout the research project and their positive support and constructive criticism during the thesis preparation.

I thank Dr Mike Catton, Dallas Wilson and Thomas Aiken of the Victorian Infectious Diseases Reference Laboratory for their assistance in the acquisition of the in-house computational technology and kind provision of time providing technical support that has permitted the use of complex software and hardware for this project.

The National Enterovirus Reference Laboratory and WHO Regional Poliomyelitis Reference Laboratory are supported by funding from the Victorian State Government, the Department of Health, Australia and the World Health Organization.

This research was supported by the Victorian Life Sciences Computation Initiative (VLSCI) grant numbers VR0069 and VR0262 on its Peak Computing Facility at the University of Melbourne, an initiative of the Victorian Government, Australia.

I would like to acknowledge and thank nVidia™ for the kind donation of five Tesla General Purpose Graphics Processing Units (GPGPU’s) for the purposes of this study.

Preface

I sought assistance from, and collaborated with the people listed below and owe them a debt of gratitude for their contributions to the results in this thesis.

Chapter 2

Dr Julian Druce, Victorian Infectious Diseases Reference Laboratory (VIDRL): Provided Bovine Viral Diarrhoea Virus material and associated primer sequences for the development of the PCR inhibition detection system used in the development of the Real-time RT-PCR methods.

Chapter 3

Ms Aishah Ibrahim, Ms Yeon Yoon and Ms Sophie Polychronopoulos of VIDRL: assisted with the derivation of sequence data from viruses used in the examination of the molecular epidemiology of circulating enteroviruses identified in the Western Pacific Region in association with Acute Flaccid Paralysis during the period 2004 to 2013.

Sequencing full capsid data and the examination of growth characteristics for selected EV-C96 viruses were performed by Ms. Aishath Hassan as part of a MSc. (Laboratory Medicine) project: “Characterisation of Enterovirus C96 Isolates Exhibiting Variations in Cell Tropism” in 2013.

Clinical samples used in the analysis of Australian enterovirus circulation were derived from the Enterovirus Reference Laboratory Network of Australia. Gina Papadakis and Dr Chris Birch provided additional sequence data from EV-A71 cases for analysis in section 3.4.2, figures 3.6 and 3.7. Samples used in the examination of AFP cases from PNG and the Pacific Island countries were received as part of the WHO Global Polio Eradication Initiative and processed by the National Enterovirus Reference Laboratory, VIDRL, which operates as a Regional Poliomyelitis Reference laboratory and National Poliovirus Reference Laboratory for these countries.

Chapters 4 and 5

Dr Mike Kuiper (Victorian Life Sciences Computation Initiative): Gave specific advice relating to the simulation of complex biomolecules and the parameterisation of small molecules using VMD and NAMD on personal computers and supercomputers. Scripts and associated files where Mike has provided input are indicated in appendix A5.

Cyril Reboul and Grisha Meyer (Monash University) and Matthew Downton (IBM Australia): Assisted with the creation of a Tcl script required to generate rhombic-dodecahedral periodic boundary conditions, used in the atomistic molecular dynamics simulations of enteroviruses.

Publications

- Roberts, J.A.**, Bowden, D.S., Thorley, B. R., Reville, P.A. Chapter: *New Technologies for Viral Diagnosis and Detection, Foodborne Viral Pathogens*, Hansman, G., White, P., (Eds.), CRC Press, 1st Ed. **ISBN-10:** 1466579501. (In Press)
- Roberts, J. A.**, Hobday, L. K., Ibrahim, A., Aitken, T., Thorley, B. R. 2013. Annual Report of the Australian National Enterovirus Reference Laboratory 2012, *Communicable Diseases Intelligence*, 37 (2). [Refer to appendix A8.2.1].
- Roberts, J. A.**, Hobday, L. K., Ibrahim, A., Aitken, T., Thorley, B. R., 2013. Annual Report of the Australian National Enterovirus Reference Laboratory 2010-2011, *Communicable Diseases Intelligence*, 37 (2). [Refer to appendix A8.2.2].
- Roberts, J.**, A. Hung, and B. Thorley. 2013. Application of Bayesian methods to the inference of phylogeny for enterovirus surveillance. *Euro surveillance*: 18(9). [Refer to appendix A8.2.3].
- Roberts, J.A.**, Thorley, B.R., Bruggink, L., Marshall, J. 2013. Electron microscope detection of an endogenous infection of retrovirus-like particles in L20B cells. *Electron Microscopy*, First published online : 16 Jan 2013. [Refer to appendix A8.2.4].
- Kesson, A. M., Ming, C. C., Troedson, C. 2013. Thorley, B. R., **Roberts, J. A.** 2013. Echovirus 19 Associated with a Case of Acute Flaccid Paralysis. *Journal of Paediatrics and Child Health*, First published online: 18 Dec 2012, DOI:10.1111/jpc.12043. [Refer to appendix A8.2.5].
- Roberts, J.A.**, Kuiper, M.J., Thorley, B.R., Smooker P.M., Hung. A. 2012. Investigation of a predicted N-terminal amphipathic α -helix using atomistic molecular dynamics simulation of a complete prototype poliovirus virion. *Journal of Molecular Graphics and Modelling*, 38, September: 165-173. [Refer to appendix A8.2.6].
- Liu, H., **Roberts J. A.**, Moore, D., Anderson, B., Pallansch, M. A., Pevear, D. C., Collett M. S., Oberste, M. S. 2012. Characterization of Poliovirus Variants Selected for Resistance to the Antiviral Compound V-073 *Antimicrobial Agents and Chemotherapy*, 56 (11): 5568-5574. [Refer to appendix A8.2.7].
- Thorley, B. R., **Roberts, J. A.** 2012. The Association of Picornaviruses with Gastroenteritis. *Microbiology Australia*, 33 (2): 79-81. [Refer to appendix A8.2.8].

Table of Contents

<i>Declaration</i>	<i>ii</i>
<i>Acknowledgments</i>	<i>iii</i>
<i>Preface</i>	<i>iv</i>
<i>Publications</i>	<i>v</i>
<i>Table of Contents</i>	<i>vi</i>
<i>List of Figures</i>	<i>xi</i>
<i>List of Tables</i>	<i>xv</i>
<i>Abbreviations</i>	<i>xvii</i>
Thesis Abstract	1
Chapter One	
Literature Review	3
1.1 Introduction	4
1.2 The Viral Pathogen: Enterovirus	6
1.2.1 Enterovirus Taxonomy	8
1.2.2 Enterovirus Structure	10
1.2.3 The Enterovirus Replication Cycle	15
1.2.4 Cultivation and Serotype Identification of Enteroviruses	18
1.2.5 Enterovirus Detection and Characterisation using Molecular Methods ...	22
1.2.6 The Association of Enteroviruses with Human Disease	30
1.2.7 Enteroviruses as a Public Health Consideration	32
1.2.8 The Polio Eradication Initiative and the Evolution of VDPVs	34
1.2.9 Poliovirus Endgame Strategy and the Poliovirus Antiviral Initiative ...	37
1.3 The Determination and Visualisation of Virus Structure	40
1.3.1 X-ray Crystallography	40
1.3.2 Transmission Electron Microscopy	42
1.3.3 Cryogenic Electron Microscopy	44
1.3.4 Nuclear Magnetic Resonance	47
1.4 Virus Structure Prediction Using Comparative Protein Modelling Methods	49
1.5 Molecular Dynamics Simulation	52
1.5.1 Simulation of Virus Components	58
1.5.2 Simulation of Complete Virions	59
1.6 Project Aims	61

Chapter Two

Development of a Novel Enterovirus Detection and Super-Speciation Assay	62
2.1 Introduction	63
2.2 Materials and Methods	66
2.2.1 Sample preparation	66
2.2.2 Growth of enteroviruses in cell culture	66
2.2.3 Nucleic acid extraction	67
2.2.4 Preparation of Control Material	67
2.2.5 Super Speciation Pan-enterovirus Reverse Transcription Polymerase Chain Reaction	68
2.3 Results	71
2.3.1 Novel Super-Speciation Semi-Nested Real-Time RT-PCR Assay	71
2.3.2 Novel Super-Speciation One-Step Real-Time PCR Assay	76
2.4 Discussion	80

Chapter Three

Characterisation of Enteroviruses Associated with Acute Flaccid Paralysis	82
3.1 Manuscripts Published in Association with this Chapter	83
3.2 Introduction	83
3.3 Materials and Methods	85
3.3.1 Partial VP1 Sequencing of Enterovirus Positive Materials	85
3.3.2 Determination of Enterovirus Type	87
3.3.3 Sequencing of the EV-A71 Capsid Region	88
3.3.4 Sequencing of Newly Described and Novel Virus Capsid Region	91
3.3.5 Phylogenetic Analysis of Enteroviruses	92
3.4 Results	94
3.4.1 Typing of Enteroviruses Associated with Acute Flaccid Paralysis	94
3.4.2 Characterisation of EV-A71 in Australia	97
3.4.3 Characterisation of Variants of EV-C96	101
3.4.4 Characterisation of the Novel Virus EV-A120	103
3.5 Discussion	107

Chapter Four

Reconstruction and Atomistic Molecular Dynamics Simulation of Poliovirus . . .	110
4.1 Manuscripts Published in Association with this Chapter	111
4.2 Introduction	111
4.3 Materials and Methods	115
4.3.1 <i>In-Silico</i> Reconstruction of Poliovirus Capsids	115
4.3.2 Integration of Poliovirus RNA using Steered Molecular Dynamics	116
4.3.3 Molecular Dynamics Simulation of a Complete Poliovirus Virion	118
4.3.4 Examination of trajectory data using VMD	119
4.4 Results	121
4.4.1 Optimization of Simulation Cell Morphology and Size	121
4.4.2 Examination of NAMD on Different Processor Architectures	122
4.4.3 Analysis of model stability	124
4.4.4 Analysis of Virus Structure after Molecular Dynamics Simulation	126
4.4.5 Examination of Viral RNA Interactions with the Capsid	132
4.4.6 Disruption of Virus Capsid Morphology	134
4.5 Discussion	138

Chapter Five

Comparative Protein Modelling of Enteroviruses	143
5.1 Manuscripts Published in Association with this Chapter	144
5.2 Introduction	144
5.3 Materials and Methods	146
5.4 Results	148
5.4.1 Comparative Modelling of Complete Virus Capsid Structures	148
5.4.1.1 Coxsackievirus A21	148
5.4.1.2 Enterovirus A71	154
5.4.2 Comparative Modelling of Novel and Newly Described Viruses	161
5.4.2.1 Enterovirus C96	162
5.4.2.2 Enterovirus EV-A120	165
5.4.3 Comparison of Picornavirus Capsid Topography	170
5.5 Discussion	172

Chapter Six

Concluding Remarks and Future Research	177
6.1 Manuscripts Published in Association with this Chapter	178
6.2 Concluding Remarks	178
6.3 Future Research	180

References	184
-------------------------	-----

Appendices	205
A1 Material and Reagents listing	205
A2 Equipment	206
A3 Software	207
A4 Sequencing Primers	208
A5 Computational Biology Scripts	209
A5.1 TCL scripts written for atom selection for application of constraints and virus substructure positioning	209
A5.1.1 File Name: Beta_select.tcl	209
A5.1.2 File Name: RNA_supercoil_betaselect_step1	209
A5.1.3 File Name: Origin_contraints.tcl	210
A5.2 TCL scripts written for the calculation of virus diameter, RMSD and RMSF	211
A5.2.1 File Name: diameterCA	211
A5.2.2 File Name: RMSD_segments	211
A5.2.3 File Name: VirusCapsidRMSD	212
A5.2.4 File Name: RMSF_Capsidv2	214
A5.3 TCL script written for the renaming of virus protein chains	215
A5.4 TCL script written to position all 60 virus protomers in the same space as protomer No. 1	216
A5.5 Selection parameters for the selection and comparison of multiple poliovirus capsid proteins with standard naming convention	217

A5.6. Small molecule parameterization, customised topology data file for virus construction in VMD	218
A5.7 Small molecule parameterization, customised simulation structure data file	223
A5.7.1 File Name: Picornavirus_small_molecules.par	223
A5.7.2 File Name: Disoxaril.par	224
A5.7.3 File Name: 1EAH_SCH4.par	225
A5.8 NAMD NPT simulation parameters for BlueGene/P and /Q supercomputers with file compression activated and custom load balancing intervals	226
A5.9 Simulation Construction Flowchart	228
A5.10 Simulation Analysis Flowchart	229
A6 Supporting Grants	230
A7 Awards	231
A8 Dissemination Details	232
A8.1 Conference Proceedings, Seminars and Public Media	232
A8.1.1 Poster: Refinement of the wild poliovirus capsid structure by atomistic molecular dynamics simulation of a complete virion	234
A8.1.2 Poster: Atomistic molecular dynamics simulation of complete human pathogenic virions	235
A8.2 Publications	236

List of Figures

Chapter One

1.1	Egyptian funerary stele 1403-1365BC	4
1.2	The complete poliovirus genome	7
1.3	Enterovirus taxonomy	9
1.4	Electron photomicrograph of enterovirus aggregates	10
1.5	Schematic diagram of enterovirus genome	11
1.6	Jelly-roll β -barrel similarity between major protomer subunits	12
1.7	Neutralising antibody (Fc region) bound to BC loop antigenic determinant	13
1.8	Picornavirus structural Features	14
1.9	Picornavirus replication cycle (cellular)	16
1.10	Microscopic morphology of enterovirus infected cells	19
1.11	Enterovirus infected cell immunofluorescence	21
1.12	Principle of real-time polymerase chain reaction	23
1.13	Schematic diagram of enterovirus genome with primer binding sites	24
1.14	Criteria for the determination of enterovirus serotype	25
1.15	Schematic Diagram of the COnsensus DEgenerate Hybrid Oligonucleotide Primer methodology (CODEHOP)	26
1.16	Phylogenetic analysis of members of the genus <i>Enterovirus</i> using the VP1 encoding region	29
1.19	Global map of polio cases 1988 and 2013	34
1.20	Map indicating global vaccine derived poliovirus distribution 2001 to 2011	36
1.21	Antiviral targets and the picornavirus replication cycle	38
1.22	Schematic diagram of an electron microscope	43
1.23	Thin section transmission electron micrograph of an L20B cell	44
1.24	Principles of single particle cryo-EM methods	46
1.25	Bonded particle interactions and the calculation of associated energies	53
1.26	Non-bonded particle interactions and the calculation of associated energies	53
1.27	Graph of common molecular activities and the associated timeframes	56
1.28	Evolution of supercomputer computational power over time	57
1.29	Illustration of the Satellite Tobacco Mosaic Virus simulation	59

Chapter Two

2.1	Phylogenetic analysis of members of the genus <i>Enterovirus</i> based on 5' NTR sequence	71
2.2	Sequence alignment indicating primer and probe binding locations	72
2.3	EV Super-speciation snrRT-PCR results for serial dilutions of control material	75
2.4	One-Step EV Super-Speciation qRT-PCR results for serial dilutions of control material	76

Chapter Three

3.1	Schematic diagram of aligned enterovirus genomes of four prototype species <i>A Enteroviruses</i>	92
3.2	Cumulative graph of the 10 dominant circulating enterovirus types	97
3.3	Phylogeographic analysis of enterovirus 71 using Bayesian methods 1997-2003	98
3.4	Phylogeographic analysis of enterovirus 71 using Bayesian methods overlaid on a global map	98
3.5	EV-A71 genogroup shift during the period of 2009 to 2013	99
3.6	Phylogenetic analysis of enterovirus A71 identified in Australia during the period 2009 to 2013	100
3.7	Phylogenetic analysis of the newly described EV C96	102
3.8	EV-A120 virus titre from original faecal sample	103
3.9	EV-A120 sequence identity to species <i>A Enterovirus</i> prototypes	104
3.10	EV-A120 Maximum likelihood phylogenetic tree (HKY+ G+ I) with 1000 psuedoreplicates	105
3.11	Simplot and BootScan outputs for EV-A120 and species <i>A Enterovirus</i> prototypes	106

Chapter Four

4.1	Helical arrangement of complete genome prior to Steered MD	117
4.2	Rhombic-dodecahedral periodic boundary conditions	121
4.3	Comparison of Computer Architectures for Simulation Environment (RMSD)	123
4.4	Comparison of Computer Architectures for Simulation Environment (RMSF)	124
4.5	RMSD of poliovirus after 10ns atomistic MD simulation	125
4.6	Average capsid radius (Å) over 0.1µs simulation time	126
4.7	RMSF values calculated over 7.5 to 10 ns time period	127
4.8	Representation of ion movement in a 10 ns equilibration of poliovirus virion using molecular dynamics	128
4.9	Examination of protomer structural motifs using protomer fit-averaging	129
4.10	Thermal comparison of VP1 structural proteins	130
4.11	View of fit-averaged VP1 structural protein	130
4.12	N-terminus of VP1 after 10 ns simulation	131
4.13	Bar graph depicting the frequency of amino acids occurring within a distance of 5 Å relative to the viral RNA	132
4.14	Examination of capsid chain RMSD by residue	133
4.15	10 ns simulations of poliovirus virion at different temperatures	135
4.16	Poliovirus type 1 (Mahoney strain) after 0.1 µs of atomistic MD simulation ..	136
4.17	RMSD and virus capsid diameter variation for simulations incorporating different pocket factors	137

Chapter Five

5.1	Predicted models of coxsackievirus A21 VP1 protein	149
5.2	Final structure output after 20 ns of atomistic molecular dynamics simulation for CV-A21	150
5.3	RMSD values for capsid simulation of CV-A21 homology model	151
5.4	Virus diameter measurements for CV-A21 over the x , y and Z axes in Å	151

5.5	Protomer fit-averaged structure output after 20 ns of molecular dynamics simulation	152
5.6	Final structure output after 10 ns of atomistic molecular dynamics simulation for EV-A71	155
5.7	RMSD values for capsid simulation of EV-A71 homology model	155
5.8	Virus diameter measurements for EV-A71 over the x , y and z axes in Å	156
5.9	Structural comparisons of enterovirus 71 homology models with X-ray crystal data	158
5.10	Comparison of hydrophobic pocket x-ray diffraction data and the predicted structure for EV-A71	160
5.11	RMSD values for capsid simulation of EV-C96 homology model	163
5.12	Virus diameter measurements for EV-C96 over the x , y and z axes in Å	163
5.13	Final structure output after 10 ns of atomistic molecular dynamics simulation for EV-C96	164
5.14	Ancestral relationships from the available crystal structures for EV-A120 ...	166
5.15	RMSD values for capsid simulation of EV-A120 homology model	167
5.16	Virus diameter measurements for EV-A120 over the x , y and z axes in Å	167
5.17	Final structure output after 10 ns of atomistic molecular dynamics simulation for EV-A120	168
5.18	Multiple amino acid sequence alignment for CV-A7, CV-A14, CV-A16, EV-A71 and EV-A120	169
5.19	3-Dimensional reconstruction of EV-A120 after 10ns MD simulation indicating residues considered significant to SCARB2 binding	169
5.20	Picornavirus phylogeny compared with capsid topography	171
5.21	Structural variation of poliovirus and EV-A71 VP4 crystal structures	174

Chapter Six

6.1	Vaccine-Derived Poliovirus type 2 with Myristate as the pocket factor	181
-----	---	-----

List of Tables

Chapter Two

2.1	Cell culture medium ingredients	66
2.2	Primer sequences for EV Semi-nested Super-speciation PCR	69
2.3	Inhibition control primer sequences for EV Super-speciation assay	69
2.4	Enterovirus types tested using EV Super-speciation snrRT-PCR	74
2.5	Enterovirus types tested using EV Super-speciation qRT-PCR	77
2.6	Enterovirus types tested using PanEV qRT-PCR	78
2.7	Limit of Detection Results for EV RT-PCR Assays	79

Chapter Three

3.1	Primers used in the “CODEHOP” assay	85
3.2	EV-A71 amplification and sequencing primers	88
3.3	Primers used in sequencing of the novel virus EV-A120	91
3.4	Breakdown of enteroviruses identified by country	95
3.5	Top 10 enteroviruses identified from cases of acute flaccid paralysis	96
3.6	<i>Enterovirus</i> species identified in Australia, Pacific Island countries and Papua New Guinea	96
3.7	Cell culture results of EV-C96 isolates	101

Chapter Four

4.1	Table outlining specific details of each simulation	116
4.2	Performance comparison of different periodic boundary conditions	122

Chapter Five

5.1	CV-A21 protomer amino acid percentage identity to poliovirus type 1 template structure	148
5.2	Backbone RMSD values for predicted models of VP1 for CV-A21 using comparative modelling methods	149
5.3	RMSD values for individual proteins and protomer for CVA 21	153
5.4	Quality assessment of CV-A21 predicted model using the ProQ Protein quality prediction website	153
5.5	EV-A71 protomer amino acid percentage identity to poliovirus type 1 template structure	154
5.6	Backbone RMSD values calculated for predicted models compared to available EV-A71 structures	156
5.7	Pairwise comparison of “Backbone” RMSD values calculated for predicted models compared to available EV-A71 structures	157
5.8	Quality assessment of EV-A71 predicted model using the ProQ Protein quality prediction website	159
5.9	RMSD values for predicted EV-A71 models compared to the reference structure 4AED	159
5.10	RMSD values calculated between original x-ray diffraction data derived for EV-A71	161
5.11	EV-C96 protomer amino acid percentage identity to poliovirus type 1 template structure	162
5.12	Quality assessment of EV-C96 predicted model using the ProQ Protein quality prediction website	164
5.13	EV-A120 protomer amino acid percentage identity to poliovirus type 1 template structure	165
5.14	Quality assessment of EV-A120 predicted model using the ProQ Protein quality prediction website	168

Abbreviations

3'NTR	3- prime Non-translated Region (synonymous with 3'UTR)
3'UTR	3- prime Untranslated Region
5'NTR	5- prime Non-translated Region (synonymous with 5'UTR)
5'UTR	5- prime Untranslated Region
AFP	Acute Flaccid Paralysis
CCID ₅₀	Cell Culture Infectious Dose
CDC	Centers for Disease Control
CHARMM	Chemistry at HARvard Molecular Mechanics
CPE	Cytopathic Effect
CPU	Central Processing Unit
cryo-EM	Cryogenic Electron Microscopy
CUDA	Compute-Unified Device Architecture
CV	Coxsackievirus
EFLOPs	Exa (10^{18}) Floating Operations per Second
ELISA	Enzyme-Linked Immunosorbent Assay
EM	Electron Microscopy
ERLNA	Enterovirus Reference Laboratory Network of Australia
EV	Enterovirus
FEP	Free Energy Perturbation
FLOPs	Floating Operations per Second
GFLOPs	Giga (10^9) Floating Operations per Second
GPEI	Global Poliomyelitis Eradication Initiative
GPU	Graphics Processing Unit
HEV	Human Enterovirus
HRV	Human Rhinovirus
HFMD	Hand Foot and Mouth Disease
IF	ImmunoFluorescence
LBM	Lim Benyesh-Melnick
M	Molarity (Moles/Litre)
MD	Molecular Dynamics
NAMD	Nanoscale Molecular Dynamics

NERL	National Enterovirus Reference Laboratory
NMR	Nuclear Magnetic Resonance
NPRL	National Poliovirus Reference Laboratory
NPT	Constant Number (atoms), Pressure and Temperature
PBC	Periodic Boundary Conditions
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PFLOPs	Peta (10^{15}) Floating Operations per Second
PV	Poliovirus
qRT-PCR	Quantitative Real-Time Reverse Transcription Polymerase Chain Reaction
RCSB	Research Collaboratory for Structural Bioinformatics
RMSD	Root-Mean Square Deviation
RMSF	Root-Mean Square Fluctuation
RNA	Ribonucleic Acid
RRL	Regional Poliomyelitis Reference Laboratory
RT-PCR	Reverse-transcription Polymerase Chain Reaction
SDHC	Secure-Digital High Capacity
snRT-PCR	Semi-nested Reverse-transcription Polymerase Chain Reaction
snrRT-PCR	Semi-nested Real-Time Reverse-transcription Polymerase Chain Reaction
SMD	Steered Molecular Dynamics
STMV	Satellite Tobacco Mosaic Virus
Tcl	Tool Command Language
TEM	Transmission Electron Microscopy
TFLOPs	Terra (10^{12}) Floating Operations per Second
TMV	Tobacco Mosaic Virus
VDPV	Vaccine-Derived Poliovirus
VDW	Van der Waals
VIDRL	Victorian Infectious Diseases Reference Laboratory
VLSCI	Victorian Life Sciences Computation Initiative
VMD	Visual Molecular Dynamics
VP	Viral Protein
WHO	World Health Organization
WPRO	Western Pacific Region

Thesis Abstract

This PhD thesis examines the integration of complex computational methodologies with the surveillance and research of a genus of viruses implicated in a wide variety of clinical conditions, ranging from asymptomatic infection to death. These viruses, known as the enteroviruses, are some of the most studied viruses in history and as a result are represented by a vast body of literature. The fact that enterovirus research and surveillance rests upon such an extensive foundation of published material, makes enteroviruses a perfect candidate for the experimental application of modern computational methods, or *in-silico* experimentation.

Fundamental to the understanding of virus behaviour is the determination of molecular structure and function, a fact which applies not only to viruses, but also to biological entities in general. As computational power has increased, the potential for the application of advanced analytic and predictive *in-silico* methods to further our understanding of virus behaviour, cannot be ignored. A practical example of this advance in computational power can be found in the one million atom *in-silico* reconstruction and simulation of Satellite Tobacco Mosaic Virus (STMV). This particular experiment, performed in 2006, required a supercomputer to calculate 50 nanoseconds of data at a rate of 1 nanosecond per day¹, a feat that can be now performed on a modern desktop computer in the same time-frame.

In-silico experimentation, similar to that applied to STMV, has resulted in the formulation of this thesis and of the hypothesis that computational power currently available can be utilised for multiple stages of virus study incorporating identification, epidemiology and atomic structure prediction. Extensive work was performed during the course of this thesis in adapting classical molecular dynamics techniques to the large scale simulation of a prototype poliovirus, using millions of simulated atoms. The successful application of these techniques has resulted in microsecond-timescale, atomistic simulations of complete virus particles. These simulations represent the first published instance of the simulation of a biologically complete pathogenic microorganism, incorporating the encoding genetic information².

This thesis also examines the use of bioinformatics methods in the development and application of an advanced quantitative multiplex real-time reverse-transcription polymerase chain reaction (qRT-PCR) methodology, for the primary screening of samples from patients suffering acute flaccid paralysis (AFP), which is one of the most debilitating presentations of enterovirus infection. The application of this novel qRT-PCR method reduces the initial screening time of samples derived from a symptomatic patient from 4-5 days using virus culture, to four hours using the novel qRT-PCR. This novel qRT-PCR method can be rapidly scaled-up in response to an outbreak situation. The ability to screen large numbers of samples during an outbreak situation is important and is hampered when using virus culture methods exclusively.

In Australia and the Western Pacific region over the last decade, the rate at which non-polio enteroviruses in cases of AFP have been identified, is on average 18%. With the introduction of PCR screening methods, a number of non-cultivable enteroviruses were identified, along with newly described and a previously undescribed enterovirus. Little is known about these newly described and novel enteroviruses. This thesis aimed to investigate the identification of viruses that may represent a significant public health threat and to then use their genetic sequence information to recreate major virus structural components *in-silico*. This reconstruction process was achieved by exploiting advances in comparative protein modelling and molecular dynamics simulation methods. In order to apply these methods to the reconstruction of previously undescribed viruses for which no structural data exist, validation of different comparative protein modelling techniques was required.

The validated methods examined in this thesis were applied to the newly described species-C *Enterovirus*, EV-C96 and a novel species-A *Enterovirus* EV-A120, which was discovered during the PhD candidature period. The predictive *in-silico* methods generated reliable atomic coordinates, representing structures suitable for the reconstruction of virus capsid models for further study. Examples of potential applications for these models are provided in the form of the observation of anti-viral binding activity in poliovirus and in the examination of spatial data for putative receptor binding domains using proposed receptor binding motifs from homologous enteroviruses.

Chapter One

Literature Review



*“It was believed that without parental genomes,
no daughter cells or progeny viruses would arise.
We have broken this fundamental law of biology
by reducing poliovirus to a chemical entity,
which can be synthesized on the basis of
information stored in the public domain” Ekard Wimmer, 2006.*

* Chemical “formula” for Poliovirus³

1.1 Introduction

Diseases caused by viruses have likely plagued man since before recorded history. Viruses belonging to the genus *Enterovirus* (including poliovirus) offer possibly the oldest evidence to support the observation of viral infection. Evidence of enterovirus (EV) infection has been observed in Egyptian skeletal remains dating back to 3700BC⁴, and evidence of possible polio infection dating back to 1400BC has been identified (Figure 1.1)⁵.



Figure 1.1 Egyptian Funerary stele 1403-1365BC. Funerary stele (stone slab) depicting the Doorkeeper Ruma, exhibiting a typical presentation of poliomyelitis. Note the atrophy of the right leg, also known as equinus foot^{4,5}. Figure derived from global polio eradication initiative website. Accessed, 15th December 2013.

The term “virus” first appears in 1599 and is the Latin term for slimy poison or venom⁶. Experiments with viruses that caused disease in plants, specifically tobacco mosaic virus (TMV), by Dimitri Ivanovsky in 1892, showed that the agents responsible for tobacco mosaic disease could be passed through very fine filters which would remove the smallest known bacteria, resulting in a solution containing a “filterable agent” that was found to induce disease when applied to healthy plants⁷. Whether or not this “filterable agent” was a living entity or a toxin was initially a scientific question, but is today a philosophical one.

Upon the widespread availability of antibiotics (a term first coined by Selman Waksman in 1942)⁸, cells cultured on the outside of an animal host could be used for the propagation of viruses. Antibiotics permitted the propagation of cells without the problems associated with extraneous growth of contaminating fungal or bacterial agents. These cell culture techniques allowed researchers to investigate virus growth and

pathogenesis as an adjunct to, and subsequently in replacement of, methods such as live animal inoculation⁹. During this time it was discovered that members of the genus *Enterovirus* exhibited diversity in cell tropism, leading to the initial characterisation of these viruses being determined on this basis¹⁰⁻¹².

After the successful crystallisation of TMV in 1935 by Wendall Stanley¹³, x-ray diffraction data describing the atomic structure of the virus was obtained by Bernal and Fankuchen in 1941¹⁴. The advances in plant virus structure determination ultimately lead to the elucidation of the first structures for viruses affecting mammalian hosts, specifically the enteroviruses - rhinovirus and poliovirus - in 1985^{15,16}.

Among the pivotal events that drove the laboratory-based investigation of virus infection in the western world, were the massive outbreaks of poliomyelitis that occurred in the United States in the mid-20th century, peaking in the 1950s, when tens of thousands of paralysis cases occurred with thousands of deaths^{17,18}. Initially it was thought that the disease was the result of a single virus type and this idea complicated the development of vaccines to combat the disease. It was later determined that there are in fact three separate and distinct serotypes of poliovirus¹⁹, an observation that was first indicated by Burnet and MacNamara in 1931²⁰. This discovery rapidly resulted in the production of two effective vaccines against all three serotypes, which remain in current use^{21,22}.

The intensity of research related to enterovirus identification, pathology, epidemiology, structure and function means there is an extensive body of literature derived from over 100 years of observation and experimentation, making enteroviruses the perfect candidate for modern *in-vitro* experiments. Recent advances in bioinformatics and computational biology now permit the researcher to combine 2-dimensional results, in the form of sequence data, with 3-dimensional results from structural analysis and to perform experiments computationally, termed *in-silico* experimentation.

The following literature review examines historic and biologic aspects of infectious agents belonging to the genus *Enterovirus*, specifically members of the species A, B, C and D enteroviruses, and examines common methods used to derive structural data. The literature review also examines an increasing role being played by modern bioinformatics and computational biology, specifically in the field of computational biophysics, and provides specific examples of their utility with respect to the examination of virus structure.

1.2 The Viral Pathogen: Enterovirus

Members of the genus *Enterovirus* incorporate established human pathogens responsible for a diverse range of clinical conditions, ranging from subclinical infection to severe illness and occasionally death²³⁻²⁵. The epidemic potential of members of the genus is well established²⁶⁻³¹. Rhinovirus and poliovirus are well known examples of enteroviruses with epidemic potential, the latter being responsible for millions of paralysis cases throughout history³².

Enteroviruses, in particular the polioviruses, are among the most studied viruses in history. Indeed, with respect to animal disease, one of the first virus structures to be elucidated using x-ray crystallography was derived from poliovirus type 1¹⁶. Poliovirus has a long history of firsts; it was the first mammalian RNA virus to be cloned and sequenced^{33,34}, the first RNA virus infecting animals to be reconstructed in the form of an infectious clone³⁵ and the first virus to be constructed entirely from synthetic oligonucleotides³⁶.

The study of poliovirus continues to be of benefit to the infectious disease research community to this day. A significant point to note is the relatively tiny size of the enterovirus genome. At approximately 7,500 nucleotides, the entire genome can be easily displayed on a single A4 page (Figure 1.2). It is humbling to realise that such a relatively simple and efficient piece of “biological code”, can be responsible for such complex disease presentation and be the source of wide-spread global morbidity and mortality.

1.2.1 Enterovirus Taxonomy

Enteroviruses are a genus of viruses classified in the order Picornvirales, family Picornaviridae. There are 17 genera classified within the picornavirus family- *Aphthovirus*, *Aquamavirus*, *Avihepatovirus*, *Cardiovirus*, *Cosavirus*, *Dicipivirus*, *Enterovirus*, *Erbovirus*, *Hepatovirus*, *Kobuvirus*, *Megrivirus*, *Parechovirus*, *Salivirus*, *Sapelovirus*, *Senecavirus*, *Teschovirus* and *Tremovirus*³⁷.

With the introduction of new classifications by the International Committee for the Taxonomy of Viruses (ICTV) proposed in 2011 and published in 2013³⁷, the structure and nomenclature of the family *Picornaviridae*, now based on genomic sequence data, have changed significantly. With the inclusion of the rhinoviruses in the genus *Enterovirus*, more than 200 enterovirus types exist and are classified within twelve species, *Rhinovirus A, B, C, Enterovirus A, B, C, D, E, F, G, H* and *J* (Figure 1.3)³⁷. The family *Picornaviridae* include pathogens with vertebrate and invertebrate hosts. This literature review will discuss only members of the genus *Enterovirus* that are known to be associated with disease in humans, specifically EV- A, B, C and D.

Prior to the reclassification of the genus *Enterovirus*, human rhinoviruses were classified in their own genus, *Rhinovirus*³⁸. This was based on physiochemical properties of the two viruses and to a certain extent, the presentation of symptomatic disease indicating a variation in the sites of replication. Rhinoviruses were for the most part acid labile³⁹⁻⁴¹ and thought to be unable to survive passage through the alimentary canal in the same manner as the human enteroviruses, with which replication has been associated with Peyers' patches in the small intestine⁴². In the National Enterovirus Reference Laboratory of Australia (NERL) a number of rhinoviruses have been identified in stool samples using polymerase chain reaction (PCR) techniques (refer to chapter 2), an observation also made in other laboratories⁴³. The significance of this finding is poorly understood. It may indicate a resistance of these particular rhinoviruses to acid destabilisation, but may certainly warrant further investigation by a rhinovirus laboratory.

Many of the initial discoveries of new members of the genus *Enterovirus* have been in response to actively searching for poliovirus in samples from outbreaks of paralytic disease or via surveillance programs aimed at polio eradication⁴⁴⁻⁴⁶. In 1948, the

coxsackieviruses were isolated from samples derived from children during an outbreak in Coxsackie New York, USA¹¹. The samples were inoculated into the cerebrum of suckling mice and the viruses isolated were subsequently named Coxsackievirus after the location from which they were derived. In 1951, echoviruses were identified after the introduction of cell culture methods. The derivation of the name echovirus is derived from the term enteric cytopathic human orphan viruses, since no disease was attributed to the virus at the time of their discovery¹². However, the echoviruses are now known to be an important pathogen associated with neurological illness and cardiomyopathy^{47,48}. Subsequent enteroviruses are named numerically and are characterised based on sequence analysis⁴⁹.

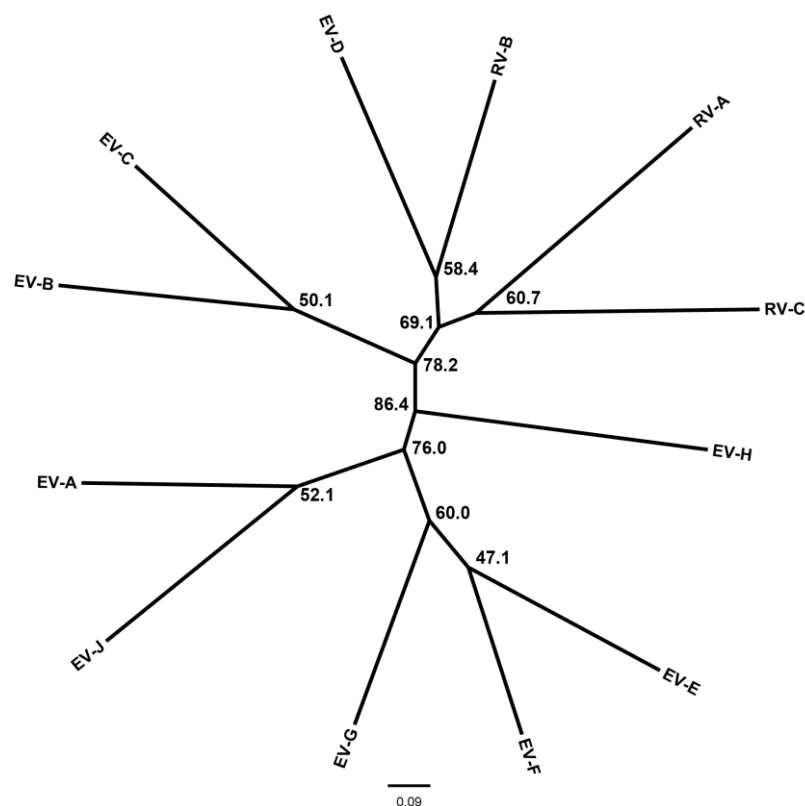


Figure 1.3 Enterovirus taxonomy. Nucleic acid maximum-likelihood radial phylogram, indicating genetic relationships of the VP1 encoding region between representative members of the genus *Enterovirus*. Node labels are described as follows: **Species EV-A** (Serotype CV-A24, *Accession No. AY421760*), **EV-B** (CV-B2, *AF081485*), **EV-C** (PV-1, *NC_002058*), **EV-D** (EV-68 *AY426531*), **EV-E** (BEV-1, *KC667561*), **EV-F** (BEV-2, *NC_021220*), **EV-G** (PEV-9, *HM131607*), **EV-H** (SEV-4, *AF326759*), **EV-J** (SpVN203, *NC_013695*), **RV-A** (RV-A1, *FJ445111*), **RV-B** (RV-B3, *DQ473485*), **RV-C** (RV-C1, *EF077279*). Scale bar indicates nucleotide substitutions per site. Alignment performed using the “Muscle” algorithm within the MEGA software package version 5.2⁵⁰. Branch lengths calculated using the Hasegawa, Kishino and Yano (HKY)⁵¹ model with discrete variables for gamma distribution and invariant sites (HKY+G+I). Bootstrap analysis performed with 1000 psuedoreplicates

1.2.2 Enterovirus structure

Enteroviruses are approximately 30 nm in diameter, roughly spherical and follow an icosahedral symmetry⁵², a configuration that is likely a result of a favourable energy minima and is observed in a number of biological systems including bacterial organelles⁵³. Observation of virus particles in infected cell culture material, indicate a honeycomb-like aggregation, arranged in a manner that is indicative of the virus's icosahedral symmetry (Figure 1.4). The capsids, which are non-enveloped, contain a positive-sense single-stranded RNA molecule roughly 7,500 base pairs in length. The RNA molecule contains a 22 amino acid viral protein (VP) which is covalently bound to the 5' non-translated region (5'NTR) of the genome, abbreviated as VPg (Viral Protein genome-linked)^{54,55}.

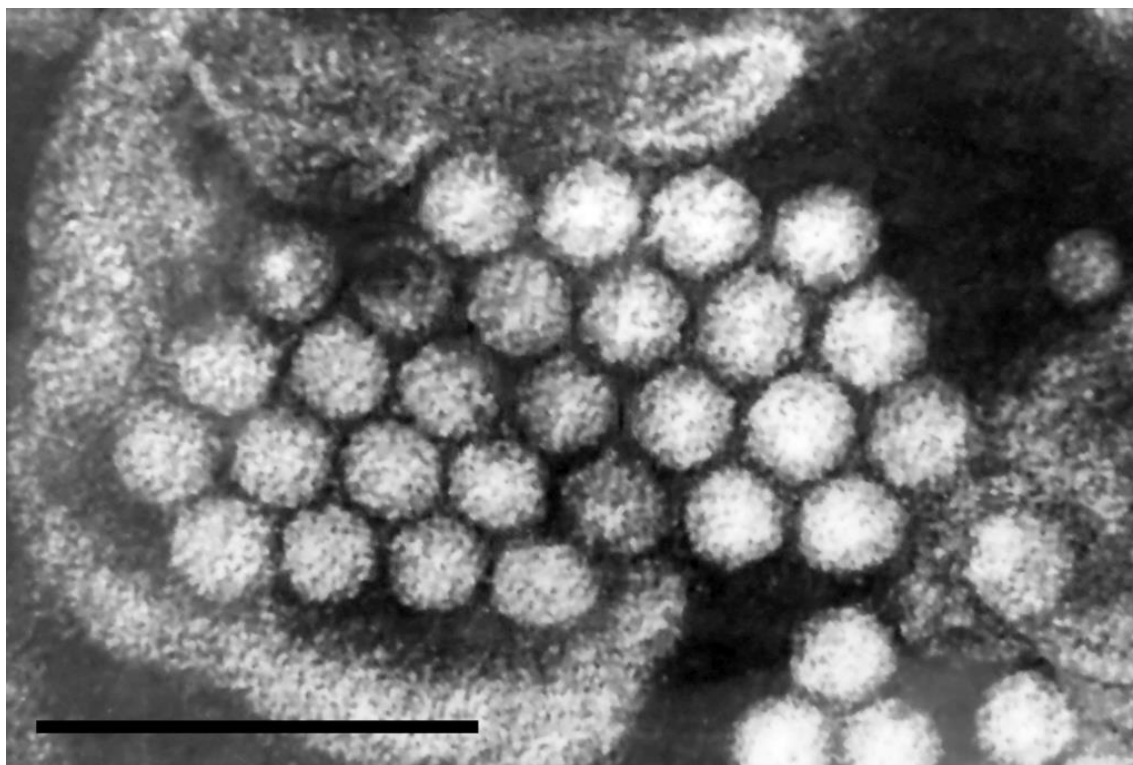


Figure 1.4 Electron photomicrograph of enterovirus aggregates. Negative stained transmission electron micrograph of aggregated enterovirus particles. Image courtesy of the Electron Microscopy Unit VIDRL (reproduced with permission 2012), scale bar = 100nm.

Also covalently bound within the capsid structure are 60 myristate molecules attached to the amino termini of the VP0 structural protein⁵⁶, which in the majority of enteroviruses, is cleaved to form the VP4 and VP2 structural proteins^{57,58}. The virus genome is

translated in the cytosol of the infected cell and acts in a similar manner to that of mRNA without the requirement for capping^{59,60}. The genome contains a complex structure within the 5' NTR (approximately the first 740 bp in poliovirus) that acts as an internal ribosome entry site (IRES)⁶¹⁻⁶³. This site initiates the formation of translation complexes which synthesize a polyprotein containing structural and non-structural proteins, which undergo post-translational proteolytic self-cleavage, (refer to section 1.2.3)⁶⁴.

The genome is sub divided into five main sections; the 5' NTR, three polyprotein regions P1, P2, P3 and a 3' non-translated region (3' NTR), (Figure 1.5). The 5' NTR region contains an internal ribosome entry site (IRES) encoding sequence, essential for initiation of translation by host ribosomes. The P1 region encodes structural proteins VP0, VP1 and VP3, with VP0 undergoing further cleavage to VP2 and VP4 in the case of members of the genus *Enterovirus*. Region P2 encodes non-structural proteins in the form of proteinases 2A, 2B and 2C, whilst region P3 encodes 3A, 3B (VPg), the proteinase 3C and the RNA dependant RNA polymerase 3D⁵², which occurs in a 3CD^{pro} precursor form.

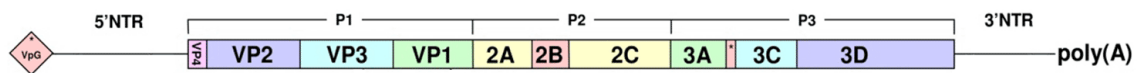


Figure 1.5 Schematic diagram of enterovirus genome. Simplified representation of the enterovirus genome showing the virus protein VPg on the distal aspect of the 5' NTR (pink diamond), 5' NTR, capsid encoding region P1 (VP1, VP2, VP3, VP4), P2 and P3 regions encoding non-structural proteins (proteinases and the RNA-dependent RNA polymerase), 3' NTR incorporating a poly-A tail. Adapted from Roberts *et al* (2010)⁶⁵.

During processing of the polyprotein phase of replication, the P1 region begins folding the structural proteins and is cleaved by protein 3CD^{pro} to form VP0, VP1 and VP3. It is not until RNA encapsidation that VP0 is cleaved to form VP2 and VP4, in a process not completely understood. It is hypothesised that proteinase activity or RNA mediated autolytic cleavage mediates this cleavage event^{66,67}. Once cleavage of VP0 has occurred, the capsid structure is “set” and is irreversible, as opposed to the pro-capsid phase, which describes the reversible formation of virus pro-capsid particles which do not contain

genomic material^{66,68}. Other picornaviruses such as parechoviruses differ from enteroviruses, in that VP0 is not cleaved to form VP2 and VP4⁶⁹.

The most important advance in the understanding of enterovirus atomic structure occurred in 1985, when the simultaneous determination of structure using x-ray diffraction techniques for poliovirus type 1 and rhinovirus type 14 were published^{15,16}. The atomic structures revealed that the structural proteins follow a common theme of β -barrels configured in a “jelly-roll” formation, as seen in many virus coat proteins and pore-forming toxins such as bacteriolysins⁷⁰⁻⁷². Combined, the three major virus capsid proteins, VP0 (VP2+VP4), VP1 and VP3, comprise a single protomer. Each of these capsid proteins exhibits a similar gross structural morphology (Figure 1.6)¹⁶. Such phenotypic presentations are not discernible at the genomic level when performing sequence alignment and comparative phylogenetic analysis, but the conservation of structure may infer a common ancestor or provide evidence of convergent evolution^{15,73}.

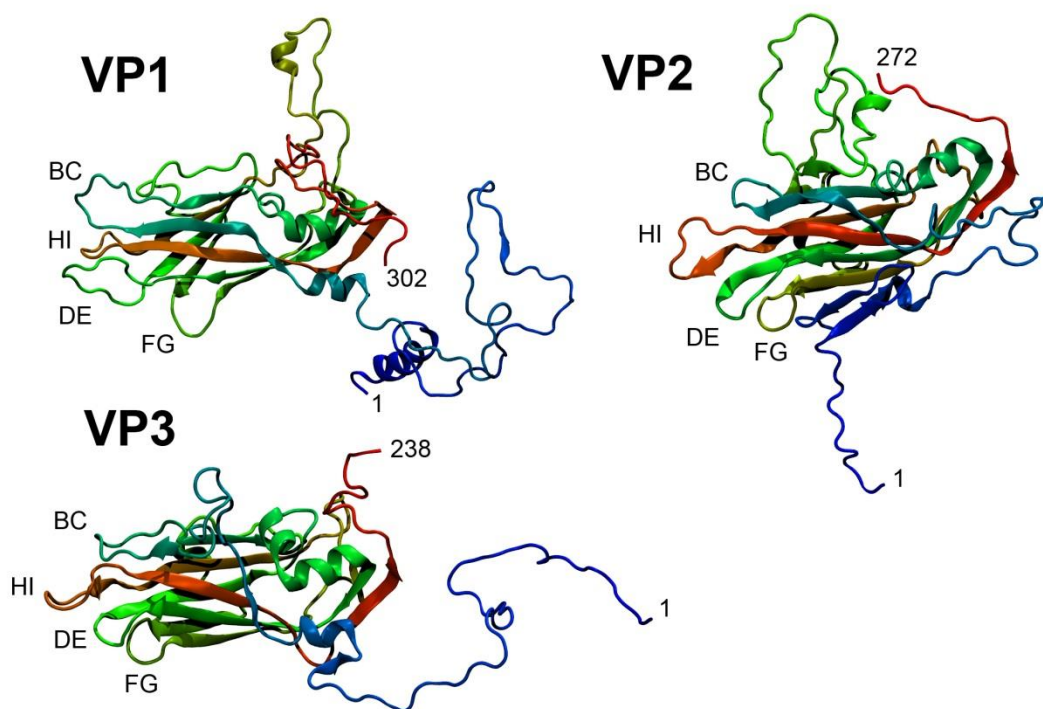


Figure 1.6 Jelly-roll β -barrel similarity between major protomer subunits. Illustration of the primary protein chains associated with a single structural subunits or protomer, Colouring, Blue= N-terminus, Red= C-terminus (derived from the poliovirus type 1 capsid structure file 1HXS and reconstructed as described in chapter 4), highlighting the homogeneity of the β -barrel structural core of VP1, VP2 and VP3, interspersed with flexible loop regions.

The loop like structures that are interspersed between alternating beta strands form important surface structures, such as the BC loop of VP1 and the EF loop of VP2, which are involved in neutralising antibody binding (Figure 1.7)^{74,75}.

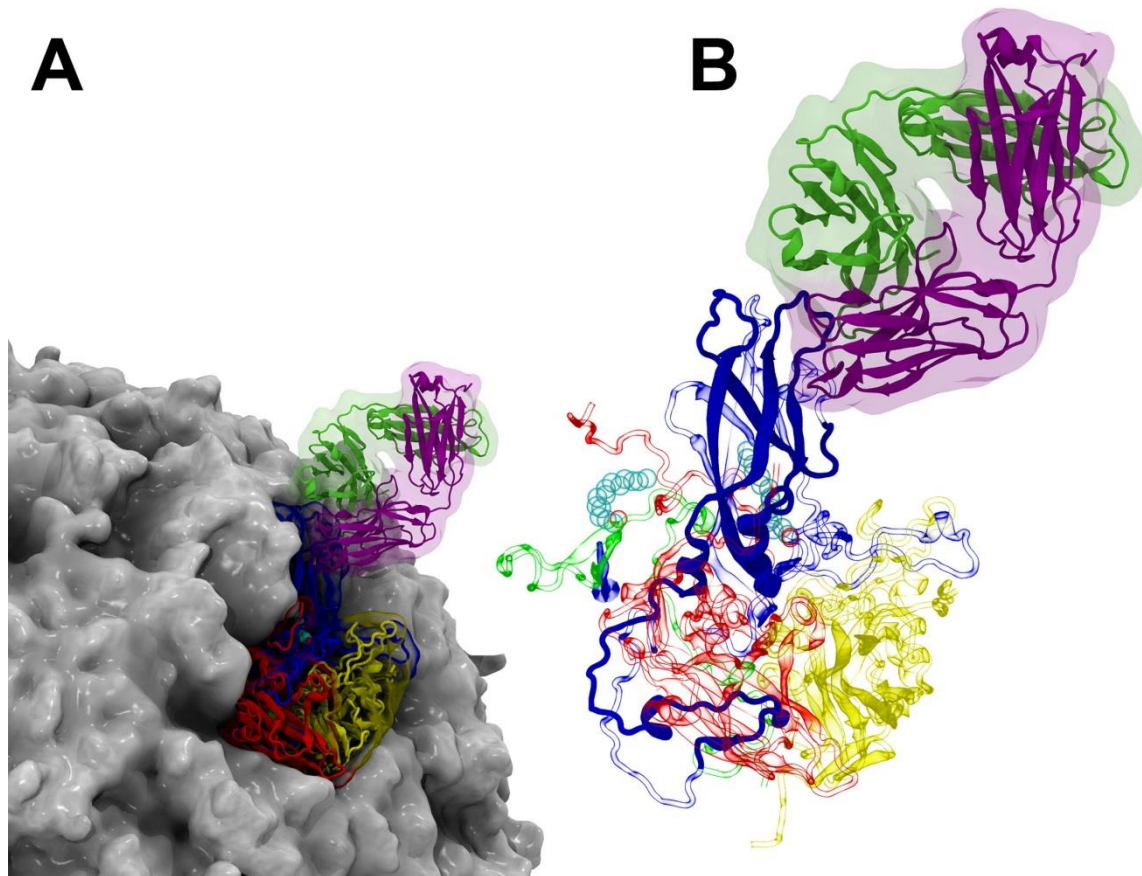


Figure 1.7 Neutralising antibody (Fc region) bound to BC loop antigenic determinant. A) 3-dimensional representation of a complete poliovirus with antibody fragment bound to a major antigenic determinant; a single virus protomer is represented. Blue – VP1, yellow – VP2, red – VP3, purple and green – variable region of neutralising antibody. B) Breakout Figure showing the protomeric subunit with a section of the VP1 protein highlighted (solid blue colour) which is targeted using PCR amplification techniques and sequencing to identify virus type. Note the variable region of neutralising antibody shown interacting with the BC loop of VP1. Reconstructed from Protein Data Bank x-ray crystallography files 1FPT⁷⁶ and 1HXS (poliovirus type 1)¹⁶.

The enteroviruses are roughly spherical and conform to an icosahedral symmetry, consisting of 60 protomers arranged in a pseudo T=3 architecture. The resulting structure consists of 12 vertices and 20 faces (Figure 1.8). VP1 proteins are assembled around the five-fold axis of symmetry, with alternating VP2 and VP3 protein chains¹⁶. This

configuration results in distinct surface features that are consistent across all members of the genus *Enterovirus*. A prominent feature of the solvent accessible virus landscape, is a canyon-like structure that exists between the VP1 five-fold axis of symmetry and the raised surfaces of the loop of VP2 and VP3 (Figure 1.8).

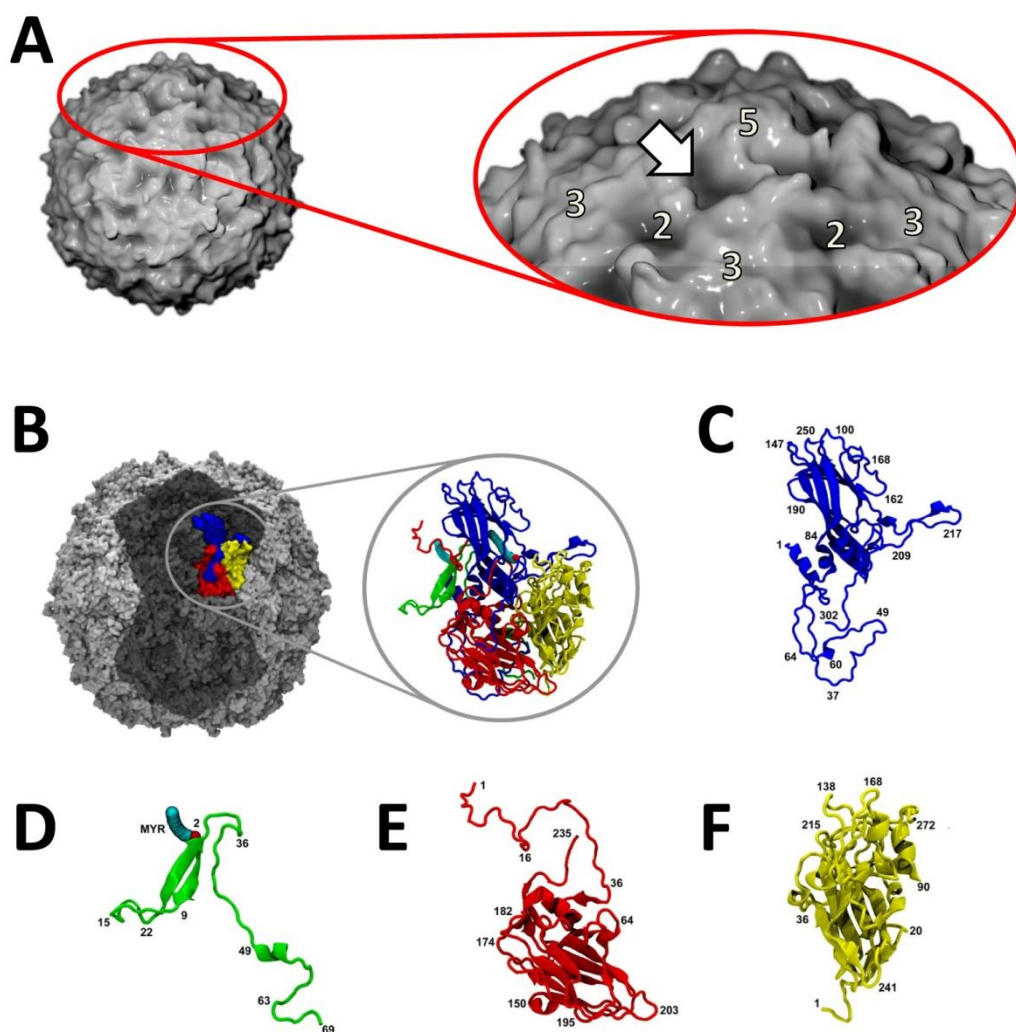


Figure 1.8. Picornavirus structural features. A) 3-dimensional representation of a complete poliovirus particle, breakout on the right indicates structural highlights, arrow indicates the canyon, 5- indicates the five-fold axis of symmetry, 2- indicates the two-fold axis of symmetry and 3- indicates the pseudo three-fold axis of symmetry, B) Coloured section shows the relative positioning of a single protomer (1 of 60) within the icosahedral virus capsid. Dark grey shading indicates the relative position of pentamers intersecting at the 2-fold axis of symmetry. The breakout image on the right of “B” shows a cartoon representation of a protomer coloured by chain. C-F) Individual protomer subunits with selective amino acid residues numbered C) VP1=blue, D) VP4=green, E) VP3=red, F) VP2=yellow. N-terminal myristate is shown in cyan and red, attached to VP4 (D) An adaptation of this image appears in *Fields Virology*⁵² and the *Journal of Molecular Graphics and Modelling*².

An important structural feature of enteroviruses is the formation of a hydrophobic pocket which lies at the base of the canyon, with a solvent accessible opening that typically contains a fatty acid such as myristate, palmitate or sphingosine, although in some viruses this pocket is thought to be empty⁷⁷. The hydrophobic pocket has been the subject of study for anti-picornavirus agents such as the WIN compounds, an example of which (known as Pleconaril) was trialled as a potential treatment for the common cold⁷⁸, with the possibility of usage in other members of the genus *Enterovirus*. These compounds are hydrophobic structures with charged regions and are thought to bind tightly within the hydrophobic pocket, inhibiting the uncoating process and thus preventing the transfer of viral RNA to the host cell for replication⁷⁹.

1.2.3 The Enterovirus Replication Cycle.

Enterovirus infection in the host is initiated in the upper respiratory tract and is spread via the faecal-oral route⁸⁰, with the exception of some enteroviruses such as enterovirus 70, which can result in haemorrhagic conjunctivitis after coming into contact with the eyes⁸¹. Studies of poliovirus have yielded significant information regarding the specific processes of infection and pathogenesis in humans, in simians and in transgenic mice which express the poliovirus receptor CD155^{82,83}. Initial infection occurs within the tonsils, with virus that is acid stable passing through the digestive tract and infecting mucosal tissue along that route, specifically the Peyer's patches^{80,84}. The virus can be detected in these sites typically within 1 to 3 days after infection. Approximately a quarter of all cases results in a viraemia after the virus migrates via the lymphatics⁸⁰. In cases of central nervous system infection, invasion occurs typically during the viraemia phase⁴². Damage to the central nervous system is evident as either aseptic meningitis, which accounts for approximately 1% to 10% of cases, or paralysis, which accounts for approximately 0.1% to 1% of cases⁸⁰. Invasion of the central nervous system by poliovirus has been demonstrated as a process of retrograde axonal flow⁸³, with evidence of a possible receptor-mediated permeation of virus across the blood-brain barrier⁸⁵.

The replication of enteroviruses in cells is a relatively rapid process ranging from 5 to 10 hours³⁶ and occurs in cytoplasm of the infected cell after initial attachment and entry. The virus undergoes a conformation shift releasing the RNA into the cytoplasm, which is translated almost immediately due to the viral RNA mimicking that of the host's messenger RNA⁸⁶.

The different stages of enterovirus replication in the host cell can be broken down into the steps described in Figure 1.9 for poliovirus. The importance of these stages of replication is emphasized in the discussion of picornavirus antivirals and their targets (refer to section 1.2.9).

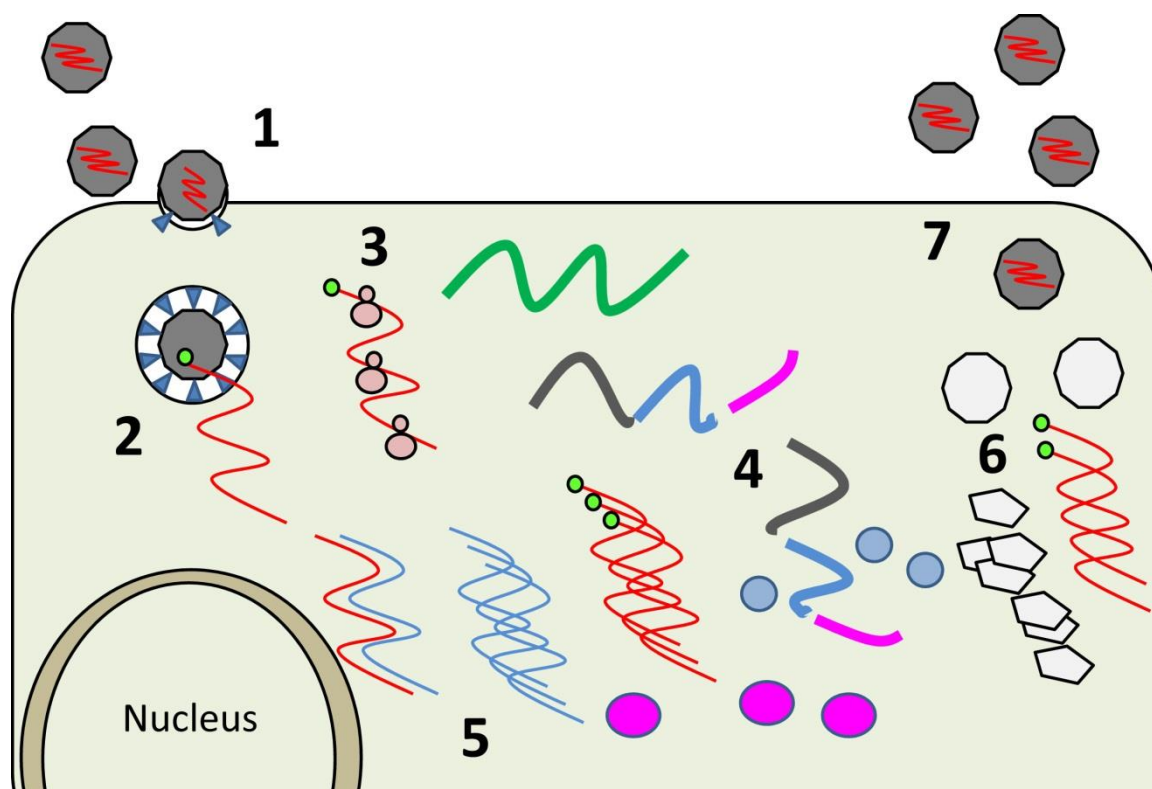


Figure 1.9 Picornavirus replication cycle (cellular). Simplified schematic representation of the stages of cellular replication of enterovirus. Numbered stages are described in further detail in the text below.

- 1) Attachment:** The virus is introduced into an environment where it can interact with a specific receptor or receptors and attachment occurs on the surface of the host cell. In the case of poliovirus, the primary receptor is CD155, which interestingly in a purified state when introduced to viruses in solution, is capable of inducing a conformational shift that results in the loss of genomic RNA. CD155 is an adhesion molecule which has a role in recognition for natural killer cells and interacts with the number of cell surface receptors on cells involved in activation of cytotoxic activity⁸⁷.

- 2) **Entry and uncoating:** Once bound to a host cell, the virus almost immediately undergoes a conformational shift that results in extrusion of the RNA into the host cell cytoplasm. For this process to occur in some enteroviruses, a pH change is required. This change has been shown not to be the case for poliovirus and more recently coxsackievirus A9⁸⁸. In the case of foot and mouth disease virus and rhinoviruses, entry to the cells is via receptor-mediated endocytosis⁸⁹. In this instance, uncoating occurs as a result of pH mediated processes associated with the host's endocytic pathway. It was previously thought that the uncoating process involved extrusion of viral RNA at the five-fold axis of symmetry⁹⁰, however subsequent studies have shown that this may not be the case and that the RNA is released on the two-fold axis of symmetry⁹¹. As the single-stranded, positive sense viral RNA resembles that of eukaryotic messenger RNA, translation is able to proceed almost immediately⁹².
- 3) **Translation:** In the case of poliovirus, up to 1500 base pairs of additional RNA can be added to the virus genome, implying that there is some available space within the capsid⁵². Such space is limited and no enzymes are carried over from the virion into the host cell. Protein translation therefore occurs primarily due to the secondary structure of the IRES and its ability to initiate translation by ribosomes⁶¹. This initiation is thought to occur due to specific binding of the IRES structure to the 40S ribosomal subunits⁹³. The enterovirus genome is translated as a polyprotein, which undergoes post-translational cleavage into a number of structural and non-structural proteins once VPg has been removed^{64,94}.
- 4) **Post-translational cleavage:** Picornaviruses encode a number of structural and non-structural proteins, specifically proteases and an RNA-dependent RNA polymerase. The three domains of the polyprotein can be broken down into P1, P2 and P3 regions. Cleavage of the P3 region is facilitated by the 3C proteinase whilst cleavage of P1 from P2 is facilitated by the 2A proteinase. The P2 region contains a proteinase, whilst the P3 region contains a proteinase and the RNA-dependent RNA polymerase 3D⁵².
- 5) **Genome replication:** Strand synthesis of RNA occurs primarily on membrane vesicles, where the viral RNA polymerase 3D⁹⁵ copies the positive stranded RNA into a negative stranded format, forming a template for the subsequent production of positive stranded RNA molecules, which are packaged into newly formed virions.

- 6) **Virion assembly:** Immediately upon translation of the viral RNA, formation of the beta barrels that comprise the structural proteins VP1, VP3 and VP0, occurs in the P1 region of the polyprotein. The structural proteins are subsequently cleaved by the proteinase 3CD¹⁶. The first structure to be assembled is termed the 5S protomer and consists of VP1, VP3 and VP0. Five of these protomers are assembled to form a 14S pentamer. Once pentamers have been formed, they possess the ability to self-assemble, through a combination of electrostatic distribution and hydrogen bonding. The first stage at which the newly formed structural proteins resemble the final virus structure, occurs once 12 of the pentamers have aggregated to form an 80S empty capsid structure. At this stage, the empty capsid is still capable of disassembly and reassembly and it is not until the empty capsid contains the genomic RNA that the pro-virion undergoes hydrolytic cleavage of VP0 to VP2 and VP4, so that the structure is essentially locked into place and is termed a virion or 160 S particle⁶⁸.
- 7) **Virus exit:** Virion exit is achieved in less than 16 hours after the initial infection and occurs as a result of apoptosis induced by viral replication. Virus proteins involved in apoptosis include proteinases 2A, 3C, and the 2C NTPase, whereas inhibitors of apoptosis include 2B and 3A. Stages leading to apoptosis in the host cell can be uniquely identified microscopically and can be attributed to the cytopathic effect (CPE) characteristic of an enterovirus infection, following infection of susceptible mammalian cell lines⁵².

In an immuno-competent infected host, virus shedding via the faeces can continue for weeks⁹⁶. In patients suffering from primary immune disorders such as B-cell mediated immunodeficiency, virus replication and subsequent shedding of virus can occur for decades⁹⁷.

1.2.4 Cultivation and Serotype Identification of Enteroviruses

The desire to detect and to characterise a pathogen, is ultimately determined by a need to understand the aetiology of a disease. Often, advancement in this understanding is proportional to the severity of the condition. Accordingly, the understanding of the “filterable agent” responsible for diseases such as Hand Foot and Mouth disease, meningitis and more importantly poliomyelitis among others, gained traction in the discovery of the viruses ability to transmit the between live hosts *in-vitro*⁹⁸. Indeed, with the agent of poliomyelitis, great strides in the isolation and characterisation of the virus

were made in 1908 after the successful propagation of the virus in primate hosts by Landsteiner and Popper⁹⁹.

The history of cell culture methods in the propagation of human enteroviruses can best be outlined in the context of polio. The discovery that poliovirus could be propagated in non-neuronal tissue by Enders *et al.*, (1949)¹⁰⁰, led to the use of a variety of cell lines which displayed a CPE (Figure 1.10). The use of multiple cell lines for the propagation of poliovirus, combined with the use of antisera neutralisation techniques, was seminal to the ability to initially differentiate between virus types. The ability to further characterise enteroviruses by serotype, proved useful to the association of a virus serotype with disease presentation⁸⁰.

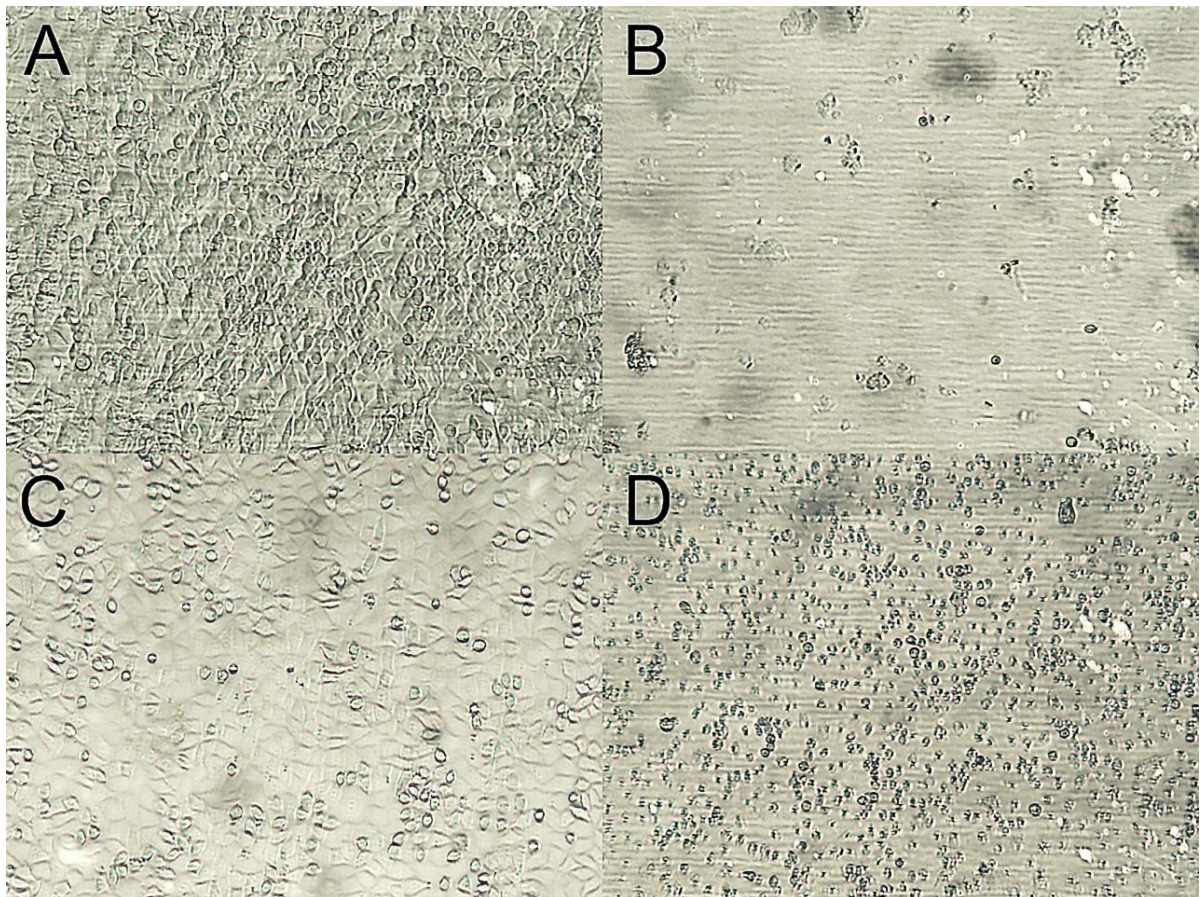


Figure 1.10 Microscopic morphology of enterovirus infected cells. Light microscope photomicrograph highlighting cytopathic effect of cells infected with poliovirus. (A) Uninfected RD-A (rhabdomyosarcoma) cells and (B) poliovirus infected RD-A cells. (C) Uninfected L20B (CD155 transgenic murine fibroblast) cells (D) poliovirus infected L20B cells at 100x magnification.

In the 1940s, the widespread introduction of antibiotics in primary cell culture methods permitted the efficient use of a variety of cell lines, both primary and continuous, for the propagation of viral agents. The use of cell culture methods was also pushed forward in a rapid manner in response to the large scale manufacture of the Salk inactivated polio vaccine²². Cell culture methods in this respect were useful not only in the detection and characterisation of viruses, but also provided a suitable platform for the mass production and purification of viral agents.

Although enteroviruses have the potential to induce a CPE in cell culture when inoculated with extracts of clinical specimens, the results are not always consistent and include limitations. Primarily, the methods involved in the continued cultivation of cell lines are expensive and time consuming. Not all enteroviruses are cultivable by cell culture and it is necessary to use a variety of cell lines to identify the wide range that can be grown. Once a CPE is observed, a subjective association with virus infection is made and this association must be confirmed with a secondary method such as antibody-based typing or molecular characterisation, as toxicity derived from the inoculum can sometimes be misidentified as CPE⁸⁰. Before the widespread adoption of molecular detection and characterisation techniques, cell lines exhibiting lytic CPE characteristic of enterovirus infection would be identified to the serotype level, by the use of serotype-specific antibodies. The most common of these methods are antisera neutralisation, enzyme-linked immunosorbent assay (ELISA) or fluorescent antibody staining¹⁰¹.

Antisera neutralisation employs the principle of opsonisation of the virus particle with a virus specific antibody, resulting in the inhibition of cell infection, thus preventing the development of CPE. Antisera neutralisation methods for enterovirus serotype identification involved the use of intersecting specific antisera pools; the Lim Benyesh-Melnick (LBM) was the most widely implemented method with eight pools¹⁰². LBM antisera pools have become scarce not only in availability but also usage and are often reserved for use in reference laboratories performing enteroviral typing, with the use of molecular sequencing methods supplanting this method.

Prior to the widespread use of PCR detection methods, rapid methods involving the use of cell-coated coverslips, inoculated with sample, incubated for a short period of time and combined with an immuno-fluorescence (IF) detection method, known as a “shell-vial”,

were commonly used in diagnostic laboratories for the detection of enteroviruses¹⁰³. Although direct detection of rhinovirus species in respiratory samples using immunofluorescence staining methods was relatively common, this was not the case for the enterovirus species affecting humans. Immunofluorescence methods for the detection of enteroviruses were typically attempted, subsequent to amplification of virus particles using culture methods (Figure 1.11).

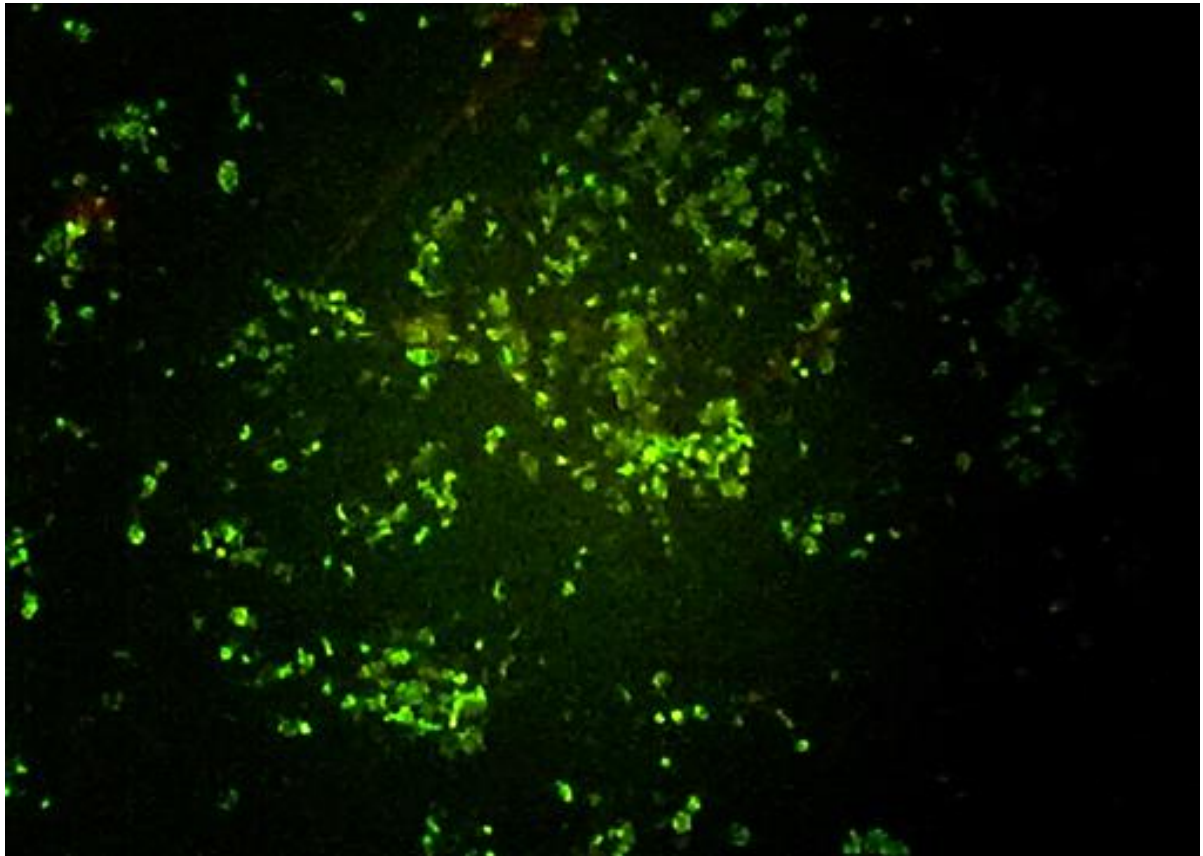


Figure 1.11 Enterovirus infected cell immunofluorescence. Indirect fluorescent immunoassay staining of enterovirus infected rhabdomyosarcoma cells showing cell fragments and associated green fluorescence due to mouse-anti-enterovirus antibody conjugated with an anti-mouse antibody labelled with fluorescein. Photomicrograph at 40x magnification.

Immunoassays which utilise ELISA or IF techniques, can use reporter molecules in either a direct or indirect manner. Direct assays are useful in that the time taken to perform the assay is minimal and the sensitivity can be considered appropriate for certain clinical settings, although the specificity of the antibody can be problematic with signal-to-noise ratios becoming unacceptable. This problem can be overcome with the use of indirect methods, employing an antibody directed typically to the Fc region of the original

antibody. The reporter molecule is placed on the second antibody as opposed to the original antibody.

Overall, cell culture-based techniques can take days to weeks from the receipt of a sample for testing and the final reporting of an enterovirus type. This timeframe is not generally useful in the clinical management of a patient and as a result, cell culture based methods have fallen out of favour, being replaced by more rapid methods based on the PCR amplification of specific viral genome sequences in clinical samples.

1.2.5 Enterovirus Detection and Characterisation using Molecular Methods.

Contemporary methodology for the detection and characterisation of enteroviruses has moved from cell culture-based antisera neutralisation of virus to the more rapid PCR methods, the use of which became more widespread in the 1990s^{104,105}. The combination of fluorescent moieties and genetic sequence-specific primers and probes, have allowed for the introduction of rapid detection methods^{106,107}. These are becoming the mainstay of diagnostic virology laboratories globally, due to the high throughput nature of their implementation and the rapid turnaround times, which can be useful in situations where rapid clinical intervention decisions need to be made, such as the differentiation of herpetic or enteroviral aseptic meningitis¹⁰⁸.

The PCR method was originally developed by Kary Mullis when working for the Cetus Corporation as a chemist¹⁰⁹. PCR methods rely on a combination of specific oligonucleotide primers that flank the regions of an area of interest on the complementary strands of DNA or RNA. When combined with a thermostable polymerase, typically variants of the *Thermophilus aquaticus* polymerase, transcription of the complementary strand in DNA form is achieved from the site of binding of the initiating oligonucleotide primer. The repetitive cycling of thermal conditions from high temperature, (typically 95°C, which results in the denaturation of the DNA strands), followed by a lower annealing temperature (allowing the binding of the oligonucleotide primers), is then followed by an extension stage usually performed between 60°C to 72°C. As this process continues, the forward and reverse primers that flank the region of interest are bound in tandem on each amplified DNA product, resulting in an exponential increase in that particular fragment of interest (Figure 1.12).

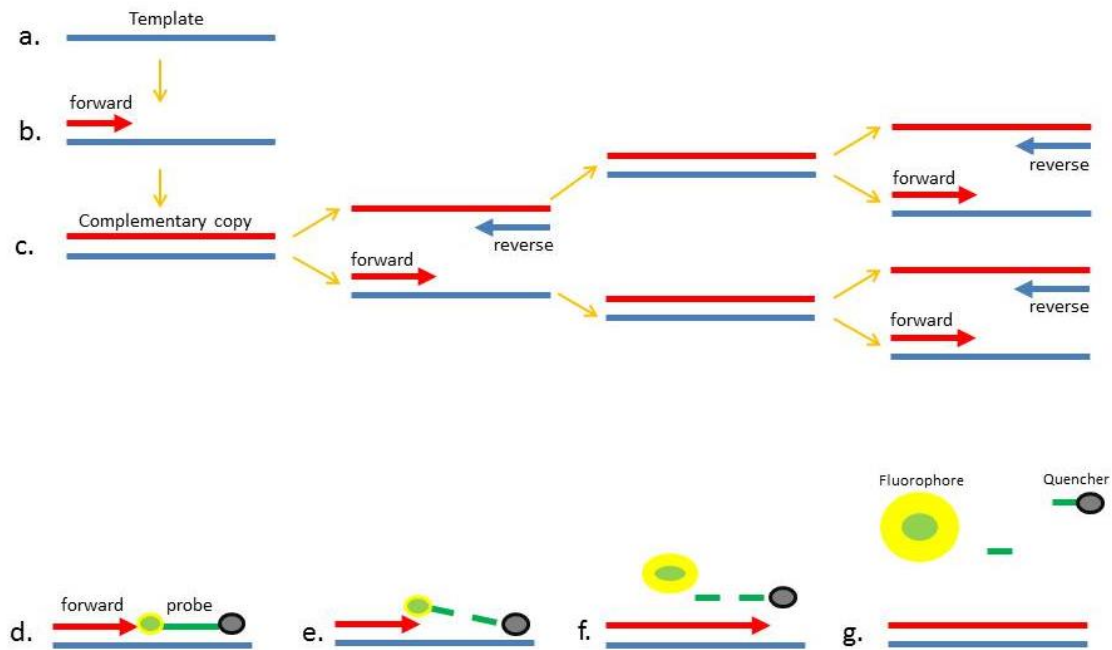


Figure 1.12 Principle of real-time polymerase chain reaction. Figure showing a) template strand of DNA ready for amplification, b) Forward primer binds to the DNA template and is extended in a complementary DNA sequence by DNA polymerase, c) this process is repeated using forward and reverse primers resulting in an exponential replication of DNA copies, d) binding of a TaqMan probe to the target of interest, followed by, e) strand synthesis DNA polymerase, f) the probe undergoes hydrolytic cleavage due to the exonuclease activity of the DNA polymerase releasing the fluorophore, g) separation of the fluorophore from the quencher results in an increase in signal intensity.

An adaptation of this method involves the incorporation of a probe sequence that is complementary to sequence located between primers of the target fragment. In the real-time implementation of this technique, alterations in fluorescent signals are exploited to indicate that amplification has occurred. A commonly used variant of this technique is termed the TaqMan probe system, and involves the 5' labelling of the probe with a fluorophore and the 3' labelling of the probe with a molecule that is specifically tailored to absorb or “quench” the fluorescence emitted by the 5' labelled fluorophore. During DNA strand synthesis, the bound oligonucleotide probe undergoes hydrolysis by the polymerase enzyme releasing the fluorophore, resulting in an increase in fluorescent signal (Figure 1.12). Each cycle of the thermal profile results in the scanning of the reaction by a detector and the amount of fluorescence is calculated. This method allows for a quantitative assessment of the amplification process in “real-time” as opposed to standard “end-point” PCR assays.

In instances where the genetic target is not DNA-based, an initial process involving the reverse transcription of RNA to DNA using retrovirus derived RNA-dependent RNA polymerase (such as those obtained from the murine leukaemia viruses), is performed prior to PCR amplification. This process is termed reverse-transcription PCR (RT-PCR). When combined with real-time PCR amplification systems, they are termed rRT-PCR for qualitative assays, or qRT-PCR for quantitative assays.

The most common enterovirus PCR detection methods employed in laboratories are based on the 5'NTR region (Figure 1.13), first targeted for detection by Olive *et al.*, (1990)¹⁰⁴. This region is highly conserved amongst all enteroviruses and although the method enables sensitive detection of enteroviruses from clinical specimens, further characterisation to the serotype level cannot be reliably determined from the PCR amplicon. In the context of public health surveillance, it is desirable to identify enterovirus type, particularly when dealing with enteroviruses associated with serious neurological syndromes such as those induced by poliovirus and EV-A71.

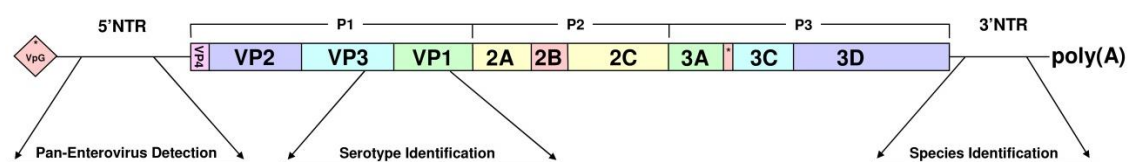


Figure 1.13 Schematic diagram of enterovirus genome with primer binding sites. Depiction of the PCR primer annealing sites within the ~7,500 nucleotide enterovirus genome, based on poliovirus type 1 Mahoney, Genbank accession number V01150 (not to scale). A viral encoded protein, VPg, is covalently attached to the 5' terminus. Examples of the areas targeted for different levels of enterovirus characterisation are indicated. Adapted from Roberts *et al.*, (2010)⁶⁵.

The application of PCR methods to the characterisation of enteroviruses, as opposed to the traditional methodologies based on antibody-based detection systems, has been a significant advance for laboratories providing detailed information for use by clinicians and public health authorities.

The availability of rapid sequencing methods based on the Sanger dideoxy nucleotide sequencing principles¹¹⁰, has allowed the amplification and elucidation of sequence data

that are specific to phenotypic characteristics or antigenic determinants that determine virus type. In the case of the enteroviruses, one of the most useful areas for phenotypic examination and characterisation is the BC loop, which appears on the external surface of the virus on the five-fold axis of symmetry (refer Figure 1.7). This region undergoes selective pressure in response to the host immune system and therefore is highly variable, making it a useful target for use in the assessment of molecular epidemiology, using sequence analysis in phylogenetic inference¹¹¹.

Guidelines have been established by Oberste *et al.*, (1999), relating to the determination of homologous serotypes when compared to prototypic enterovirus sequences of the VP1 region¹¹². Sequence matches greater than 75% from nucleic acids, or 85% from amino acids, are indicative of a homologous serotype (Figure 1.14). Percentage identities below these figures can be used to infer heterologous serotype, heterologous clusters or heterologous genera. A proposal has to been made to apply the same identification criteria to rhinoviruses¹¹³.

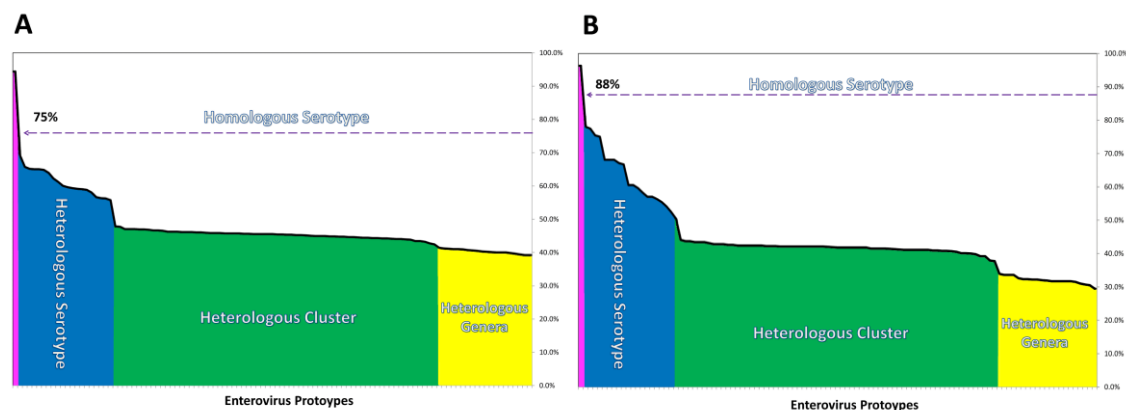


Figure 1.14 Criteria for the determination of enterovirus serotype. The Figure indicates the guidelines used for the determination of homologous serotypes for unknown sequences. Chart A) indicates the percentage identity (y-axis) calculated for nucleic acids for all available prototype enteroviruses (x-axis) when compared to an unknown sequence. Sequence matches above 75% indicate a homologous serotype. Chart B) indicates the percentage identity calculated for amino acids from available prototype enteroviruses compared to unknown sequence. Matches above 88% indicate a homologous serotype^{112,114}.

Given that there are in excess of 100 enterovirus types, there is significant difficulty in determining the etiological basis of a particular clinical manifestation in a rapid manner. Standard laboratory protocols based on 5'NTR detection methods are useful only in determining the presence of enterovirus. More specific methods based on sequencing of

the VP1 genomic region are most commonly applied and specifically the CODEHOP protocol, as described in Figure 1.15. Serotype identification methods based on VP2 and VP4 genomic regions also exist^{115,116}. Sequencing of the VP1 region using the CODEHOP method is used primarily in the research related to this thesis (chapter 3)

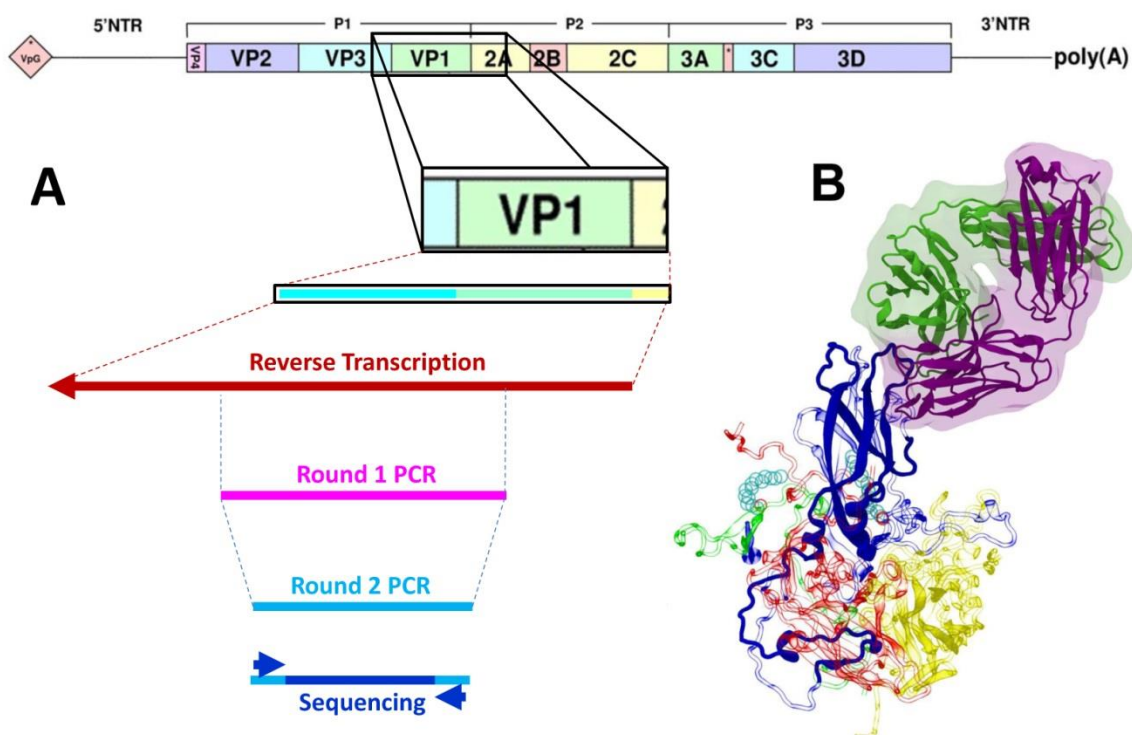


Figure 1.15 Schematic Diagram of the Consensus Degenerate Hybrid Oligonucleotide Primer Methodology (CODEHOP). A) Schematic diagram of the region amplified by the CODEHOP PCR primers, as relates to the type identification of enteroviruses. Reverse transcription of the capsid encoding region of the virus genome using a series of six base pair reverse transcription primers, followed by two round PCR. Sequencing is performed using non-degenerate primers targeted at the conserved ends of the second round product defined by the second round hybrid primers. B) 3-dimensional representation of an enterovirus protomer with all four protein chains identified VP1 (blue), VP2 (yellow), VP3 (red), VP4 (green), target sequence for the CODEHOP assay¹¹⁷ is indicated by solid blue colouring in the VP1 region. The BC loop antigenic determinant can be identified by its interaction with antibody (purple, green).

Once sequence data are available, comparative analyses can be performed using multiple sequence alignment methods, combining data for comparison from the public domain with local data. Sequence alignment methods based on similarity are available, with the initial steps involving a progressive pair-wise construction process, involving the use of a simplistic tree used to guide the subsequent alignment (guide tree)^{118,119}.

An important component of multiple sequence alignment concerns the introduction of gaps in nucleic acid or amino acid sequence data during the alignment process, particularly in areas of heterogeneity. These gaps are required to represent insertion or deletion events which may have occurred during the evolutionary processes. The correct alignment of multiple sequences that contain gaps is crucial and the most common approach to the problem of where a sequence should be “broken” in order to accommodate a gap, is to assign numeric penalties which accumulate and then mandate insertion of a gap after a predetermined threshold has been reached¹¹⁸.

Typically, sequences are aligned based on a scoring method with the least homologous sequences scoring a higher value¹²⁰. Various scoring mechanisms exist, such as the PAM¹²¹ and BLOSUM¹²² substitution matrices for amino acid substitutions, or the HKY⁵¹ model for nucleic acid substitutions, which provide a value that represents the likelihood of a substitution event occurring. The accumulation of these values is used as a scoring mechanism for later pairwise alignment. One of the most widely used programs for sequence alignment is the Clustal family of alignment algorithms, which uses a scoring mechanism combined with a progressive pair-wise construction process¹²³.

In order to convert multiple sequence alignments into a visual representation that is illustrative of sequence similarity, the use of phylogenetic trees is the most common methodology, with the first recorded instance sketched by Charles Darwin appearing in his first notebook on the “Transmutation of Species” in 1837. Generally there are two methods for the calculation of phylogenetic trees, the first based on algorithmic methods (distance matrix) and those based on discrete data methods (tree searching methods)^{124,125}.

The inference of phylogeny using the distance matrices essentially calculates the difference between sequence data, which is calculated for all pairwise combinations of

the sequence being analysed. The value for each pairwise alignment is then used to construct a tree with branch lengths based on values represented as distances.

Although discrete data methods can be more computationally expensive, they can also be more informative and allow for the incorporation of information related to the evolutionary history of the data being analysed. These methods incorporate techniques such as parsimony, maximum likelihood and Bayesian inference¹²⁶. Essentially these methods examine each column of an alignment and determine a phylogenetic tree that best represents this grouping of information. An example of such a tree used to determine enterovirus type and ancestral lineage, is shown in Figure 1.16.

Scoring methods for variation in each column vary widely, with a number of methods incorporating empirical evidence of the rates of change in nature related to the probability of a particular variation occurring^{121,122}. For example, an adenine to guanine mutation is more likely to occur than an adenine to cytosine change or in the case of amino acid alignments, the probability of an alanine to glycine change will have a greater impact than an alanine to histidine change, given the physical difference in size and charge of the side-chains. In the case of virus sequence analysis, the rate of change or introduction of mutations is vastly greater than that of the eukaryotes; this is particularly evident for the RNA viruses as opposed to DNA viruses. Therefore the incorporation of dates of collection (historic time point) and the geographic location, as discrete data, enhances the investigation of virus ancestral lineages from a temporal and geographic perspective. Such investigations are termed phylogeography¹²⁷.

Phylogeographic analyses of viral outbreaks have been used to determine ancestry of the rabies and influenza viruses^{127,128}. These methods can be useful in a public health sense to assist in determining the source of epidemics, or chains of transmission, and can help in planning to restrict or prevent disease spread. The representation of phylogeny using trees can be further enhanced when combined with global mapping programs such as Google Earth¹²⁹. This technique can give the observer an efficient and simplified overview of virus transmission with respect to a particular global region of interest. (refer to section 3.3.5 and 3.4.2 for information and examples of Bayesian phylogeography). The use of phylogenetic inference methods was critical to the investigations of 5'NTR heterogeneity

in the design of novel PCR methods (chapter 2) and to the examination of the molecular epidemiology of enteroviruses associated with severe disease (chapter 3).

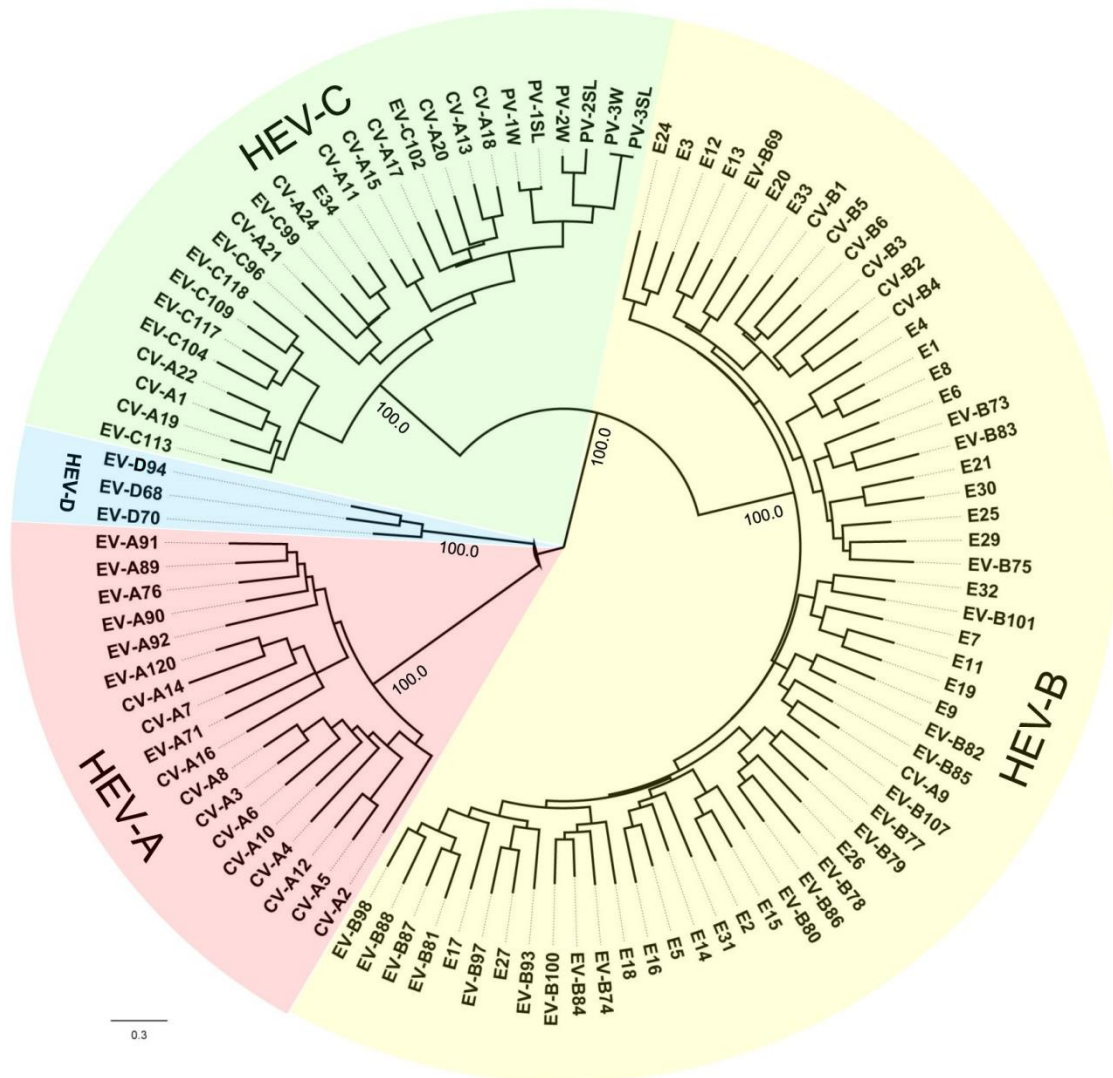


Figure 1.16 Phylogenetic analysis of members of the genus *Enterovirus* using the VP1 encoding region. Analysis performed using the Maximum Likelihood method with 1000 bootstrap pseudoreplicates, and the HKY+G+I nucleotide substitution model. Sequences aligned using ClustalW. Polar representation with scale indicating nucleotide substitutions per site.

1.2.6 The Association of Enteroviruses with Human Disease

The majority of enterovirus infections are asymptomatic. In cases where symptomatic infection does occur, disease states range from mild febrile illness to more serious sequelae such as neurological illnesses, paralysis and death. The enterovirus most synonymous with serious neurological illness involving devastating outcomes such as permanent paralysis or death is without doubt, poliovirus. Poliovirus infection involving the spinal cord is termed “poliomyelitis” and is derived from the Greek words grey (polio) and marrow (myelon).

Poliomyelitis, or anterior horn cell disease, is not necessarily restricted to infection by poliovirus, but is a condition that can be attributed to infection with other members of the genus *Enterovirus* such as the echoviruses, coxsackieviruses and enterovirus A71 (EV-A71)^{130,131}, the clinical syndrome exhibited by poliomyelitis, acute flaccid paralysis, may also be attributed to other causes including, but not limited to, auto-immune processes such as Guillain-Barré syndrome or acute disseminated encephalomyelitis¹³². An historic example of a polio-like disease outbreak due to a non-polio enterovirus can be attributed to the circulation of coxsackievirus A7. In the 1950’s this virus was responsible for large outbreaks and was, for a time, referred to as Russian type 4 poliovirus¹³³. Recently, serious neurological outcomes associated with infection by enterovirus A71, particularly in China and Southeast Asia, have become more prevalent^{134,135}. In Australia in 2012 to 2013, infections resulting in AFP and death have occurred. Such a high number of reported AFP cases have not been seen in Australia since the endogenous circulation of poliovirus in Australia in the 1960’s.

Approximately 90% of all poliovirus infections are asymptomatic with less than 10% involving abortive infection, 1% suffering aseptic meningitis and 0.1% suffering acute flaccid paralysis and sometimes death⁸⁰. Neurological sequelae are usually evident after a short period of a flulike illness and thus are referred to as a bimodal disease presentation¹³⁶.

Paralysis is typically a result of widespread damage to neural cells in the anterior horn region of the spinal cord. The damage is visible in magnetic resonance imaging as an increase in signal intensity with a characteristic “owls eyes” appearance¹³⁷. Paralysis

typically presents as initial weakness involving single or multiple muscle groups, gradually progressing to paralysis in one or more limbs. Some patients do not recover full use of the affected limb(s) over time and in the case of cranial nerve involvement (also referred to as bulbar paralysis), may suffer resulting disturbances of speech, swallowing or breathing. Paralysis of the muscles of the diaphragm results in the patient being unable to breathe without the assistance of a ventilator. In the case of medullary centre involvement, the outcome can be fatal¹³⁸. This has been seen in cases of EV-A71 meningo-encephalitis^{139,140}.

The most common neurological manifestation of enterovirus infection is meningitis¹⁴¹. Typically, meningitis is a self-limiting infection, with the exception of infection within the first two weeks of birth, with such cases having a poor prognosis^{142,143}.

The majority of enterovirus-related respiratory illnesses are caused by rhinoviruses, making them the most commonly isolated member of the genus *Enterovirus* when investigating cases of the common cold. The identification of rhinoviruses in clinical samples has been observed more frequently since the introduction of PCR detection methods. This has also resulted in the identification of species C *Rhinoviruses*, which were previously uncultivable^{144,145}. Recently, enterovirus 68 has been associated with severe cases of pneumonia¹⁴⁶ and has also been isolated from cases of acute flaccid paralysis (refer to chapter 3), although its role in this clinical condition is poorly understood.

Enterovirus 68 is a member of the species D *Enteroviruses* and is thus related to enterovirus 70. In the late 1960s and early 1970s, acute haemorrhagic conjunctivitis was pandemic in Asia and Africa¹⁴⁷. During this pandemic, it was observed that cases of haemorrhagic conjunctivitis due to enterovirus 70 infection, were associated with significant neurologic sequelae¹⁴⁸. Coxsackievirus A24 has also been involved in widespread outbreaks of haemorrhagic conjunctivitis⁸⁰.

Perhaps one of the most visually striking clinical presentations of enterovirus infection in early childhood would be that of hand, foot and mouth disease (HFMD). Patients can also present with herpangina which involves the appearance of greyish white vesicles in the mouth and throat. HFMD is usually mild, beginning with a sore throat and low-grade

fever and is typically associated with members of the *Enterovirus* species A¹⁴⁹⁻¹⁵². The majority of patients develop vesicles 3 to 5 mm in diameter and erythematous in appearance, typically occurring on the palms of the hands and soles of the feet consistent with the naming of the disease¹³⁴. Recent outbreaks of HFMD in Australia have been associated with coxsackievirus A6, coxsackievirus A16 and EV-A71 (Chapter 3.4.1). HFMD is typically associated with members of the *Enterovirus* species A¹⁴⁹⁻¹⁵².

Members of the species B *Enteroviruses*, specifically the echoviruses, are often associated with neurological illness¹⁴¹. However other members of the species B *Enteroviruses* including coxsackievirus B1 to B5, are not only associated with neurological illness such as meningitis, but also with cardiomyopathies such as acute myocarditis. Neonates and younger infants have a higher susceptibility to coxsackievirus B associated myocarditis and associated systemic infection, which can be fatal^{141,142}. Approximately 0.01% of acute myocarditis cases may progress to chronic dilated cardiomyopathy⁸⁰. In a similar manner to acute myocarditis, chest pain arrhythmias and potentially heart failure are symptomatic of chronic dilated cardiomyopathy¹⁴¹.

The coxsackieviruses B1 to B5 have been implicated in the onset of juvenile diabetes mellitus (type 1 diabetes), with evidence of enterovirus RNA in pancreatic tissue samples coincident with patients developing anti-islet-cell antibodies^{153,154}. The apparent seasonality of type 1 diabetes presentation supports a post-infectious disease mechanism, preceding an autoimmune response¹⁵⁵⁻¹⁵⁷.

1.2.7 Enterovirus as a Public Health Consideration

The primary means of transmission for enteroviruses is via the faecal-oral route⁸⁰. Transmission from person-to-person is efficient in environments where personal hygiene and environmental sanitation are poor.

In the context of public health considerations for enterovirus infection, poliovirus has served as the prototype virus for over 100 years. The first clinical descriptions of poliomyelitis occurred in Europe in 1784 by Dr Michael Underwood in which he described children suffering weakness and paralysis of the lower extremities^{158,159}. Subsequent outbreaks were described early in the 19th century in Europe and the United States in the late 19th century⁸⁰. By 1938, significant funding had been established in an

attempt to curb cases of paralytic poliomyelitis, with the announcement of the Salk inactivated vaccine, licenced in 1955²², followed by the Sabin attenuated vaccine, licenced in 1961-1962^{21,80,160}. As a result of the introduction of these vaccines, the incidence of paralytic poliomyelitis dropped dramatically in the countries where it was used. The last case of poliomyelitis due to wild poliovirus infection in the United States was reported in 1979 and the last indigenous case of wild poliovirus infection reported in Australia was in the 1960s¹⁶¹.

Whilst enterovirus infection occurs on a global scale, the majority of studies into the public health considerations of enterovirus infection have been performed in the United States and typically in response to large outbreaks of poliomyelitis. Of particular interest are the “virus-watch” programs performed during the 1960s and 1970s, in which the prevalence of enteroviruses, predominantly the coxsackieviruses and echoviruses, were observed to be the highest in children¹⁶². As these studies were performed using virus culture methods, non-cultivable enteroviruses would not have been identified. Only those viruses of a high enough titre and tropism for the cell lines used would have been observed.

Typically, the public health response to non-polio enterovirus infection is considered to be a low priority, given that it is usually a self-limiting illness with a low mortality rate. Since the late 1990s, emergence of EV-A71 causing HFMD has garnered international attention, due to the associated risk of developing meningo-encephalitis with bulbar involvement and death^{134,140,163-165}. Although mostly restricted to China and Southeast Asia, Australia has in recent years reported significant neurological illness and fatalities associated with EV-A71 infection and as a result, has established enhanced surveillance for the virus (refer to section 3.4.2). It is now a primary focus of the Enterovirus Reference Laboratory Network of Australia (ERLNA), established in 2009 by the NERL and funded by the Commonwealth Government of Australia.

There is concern that EV-A71 may continue to evolve and that widespread periodic outbreaks of HFMD associated with severe neurological sequelae may become commonplace. Recently, vaccines for enterovirus EVA-71 subgenogroup C4 have become available¹⁶⁶⁻¹⁶⁸; some of these vaccines have undergone clinical trials in China, but remain unavailable for use at this time¹⁶⁹. Subgenogroup C4 has been responsible for

a larger number of fatalities in China and Southeast Asia, and there is evidence in animal models that these new vaccines may provide a degree of cross-protection from other genogroups of EV-A71 such as genogroup B. Cross protective-immunity in humans is unknown and investigations are ongoing¹⁷⁰.

The ERLNA, in conjunction with the respective State and Federal departments of health, continue to monitor the situation with EV-A71, but such surveillance is not just restricted to this virus. All enteroviruses identified by serotype are collated and significant events are investigated and reported to the relevant parties (Chapter 3).

1.2.8 The Polio Eradication Initiative and the Evolution of VDPVs

In 1988, the World Health Assembly passed a resolution to eradicate poliomyelitis worldwide. Since then, the annual global disease burden of poliomyelitis has been reduced from 350,000 cases to 406 in 2013 representing a 99.9% reduction (Figure 1.19). The World Health Organization (WHO) Global Polio Eradication Initiative¹⁷¹ is based on maintaining high levels of polio vaccine coverage, clinical surveillance for cases of AFP in children less than 15 years of age, and laboratory confirmation of wild poliovirus infection. Each country has a WHO accredited polio reference laboratory, to which specimens from AFP cases can be referred.

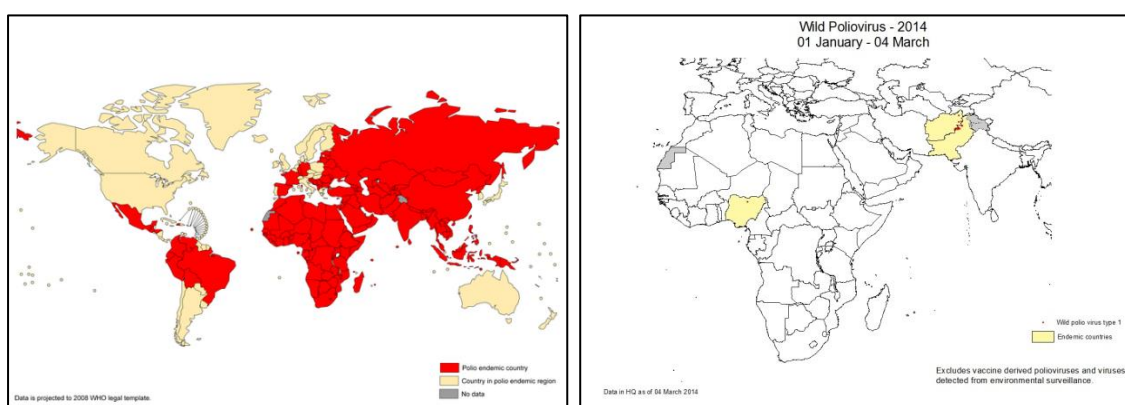


Figure 1.19 Global map of polio cases 1988 and 2013. A) Polio affected countries 1988 highlighted in red, B) polio affected countries September 2013. Images obtained from the global polio eradication initiative website March 2014. Source: www.polioeradication.org¹⁷¹.

So far, the polio eradication initiative has been successful in the eradication of wild poliovirus types 2 (last identified in India in 1999) and apparently type 3 (most recent

identification in Nigeria in November 2012), despite significant difficulties and challenges maintaining funding¹⁷¹. The most recent success was the declaration that India was certified free of wild poliovirus transmission in 2012¹⁷², an event which is considered to be possibly one of the most difficult public health challenges¹⁷³. With India's success in eradicating the virus comes the next challenge, that of dealing with the three remaining countries with circulating wild type poliovirus, where interruption of transmission has never occurred. These three countries, Afghanistan, Pakistan and Nigeria, are currently beset by military conflict or geopolitical instability. One outcome of this circumstance has been the murder of polio vaccinators and doctors associated with the eradication program, making continued efforts to maintain high vaccination levels extraordinarily difficult¹⁷⁴.

Indigenous viruses from these three countries have been identified in surrounding countries and have initiated large outbreaks from reintroduced viruses, the most recent being Somalia where over 194 cases were reported in 2013^{171,175}. With public and private financial support wavering, such introductions are potentially devastating to the eradication of poliovirus. Calls have been made to change the program from an eradication program to one of control, but costings for a control-based approach to poliovirus have shown that over the long term, the resultant suffering and financial burden to the global community far outweighs the financial costs of eradication, even in instances where virus introduction into previously virus free countries occurs¹⁷⁶.

A side effect of the use of live attenuated vaccine strains in the vaccination of children, typically in poorer nations where multiple factors conspire against making the safer inactivated form of the vaccine available, is the emergence of vaccine derived polioviruses (VDPV)¹⁷⁷. There are three categories of VDPV; circulating VDPVs (cVDPV), immune-deficient associated VDPVs (iVDPV) and ambiguous VDPVs (aVDPV).

Circulating vaccine derived polioviruses (cVDPVs) are typically found in areas where poliovirus vaccination uptake is poor or significant gaps in maintaining vaccination campaigns are evident (Figure 1.20). Such an instance occurred in Nigeria in 2002, when community leaders in northern Nigeria advocated that the vaccine was being used for nefarious reasons, specifically the transmission of human immunodeficiency virus (HIV), and was being touted as a potential cause of sterility in young people, none of which were

true¹⁷⁸⁻¹⁸⁰. During this time of non-compliance with recommended dosage schedules, the live attenuated virus may have transmitted between naïve hosts and accumulated point mutation to such a degree (approximately 1% per annum) that there was a concomitant loss of attenuation. cVDPVs are considered to be as threatening to human health as that of a wild virus. In northern Nigeria cVDPV type 2 accounted for 392 cases of polio, between 2005 and 2013¹⁷¹.

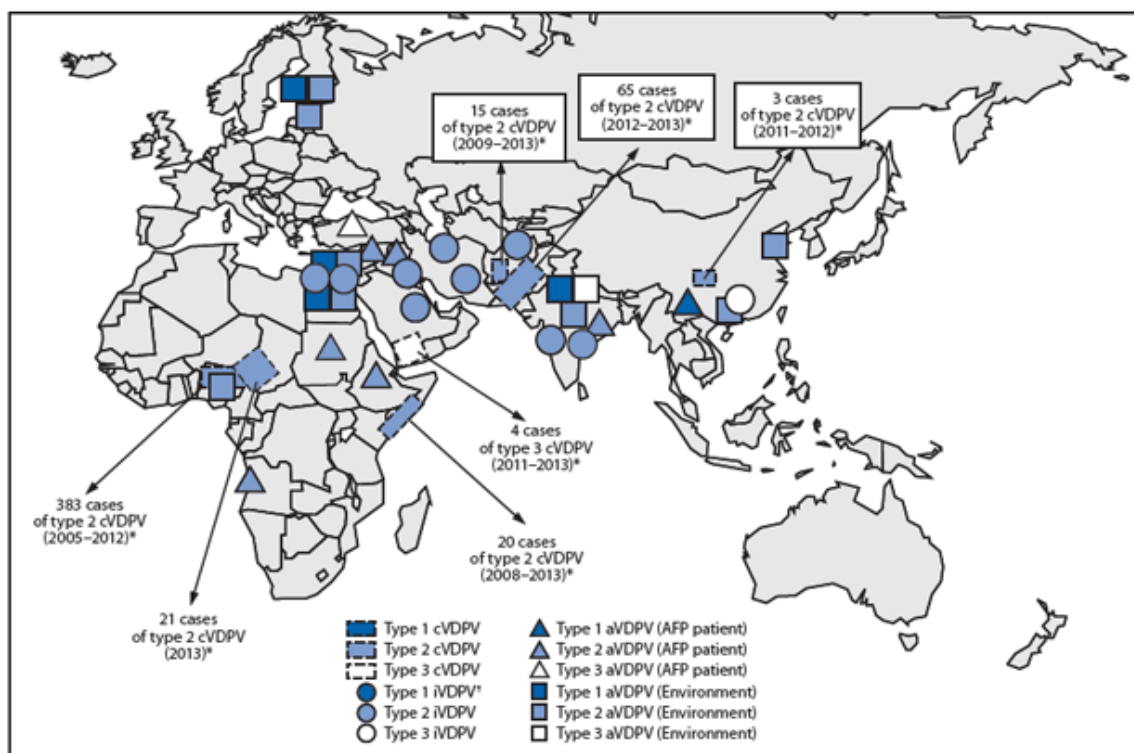


Figure 1.20 Map indicating global vaccine derived poliovirus distribution 2001 to 2013. Source: March 21, 2014 / 63(11);242-248¹⁸¹.

Revertant viruses need not only be the result of community-wide transmission amongst non-vaccinated individuals. In the case of an iVDPV, the patient is typically suffering a primary immunodeficiency involving B-cells. This immune deficiency results in failure to completely clear the poliovirus infection and patients have been known to excrete virus for periods greater than 20 years^{97,182}. In 2009, a patient with an underlying immunodeficiency was found to be suffering from poliomyelitis; the virus was examined and thought to have originated from the vaccination of a close contact many years prior. Due to the lack of specific antiviral therapies for such an instance, the patient subsequently died¹⁸³.

As the polio eradication initiative has progressed, enhanced surveillance measures are being implemented. One method used to improve surveillance for poliovirus, is environmental testing, involving the testing of sewage samples for the presence of poliovirus. If a vaccine-derived poliovirus is detected in such an environmental sample and its origin cannot be traced, it is deemed to be a VDPV of ambiguous origin or an aVDPV. aVDPVs have been identified in sewage samples in Israel as part of routine environmental surveillance¹⁸¹. However, during this course of this surveillance, wild poliovirus type 1 has also been identified¹⁸⁴, highlighting the importance of supplementary surveillance systems as an adjunct to AFP surveillance.

Once thought to be an advantage of using live attenuated poliovirus strains in a vaccine, the transmission of vaccine strains from recipients to non-vaccinated hosts is of serious concern and is an element of the OPV paradox. The OPV paradox refers to the live vaccine that prevents poliomyelitis, but can also cause poliomyelitis. Furthermore, once the wild strains of poliovirus are eradicated, Vaccine Associated Paralytic Poliomyelitis (VAPP) cases and VDPV cases resulting from live attenuated vaccine strains will become the only sources of poliomyelitis. Challenges exist in the removal of live attenuated strains, because the persistence of live virus in the community or environment represents the potential for reintroduction of the viruses¹⁸⁵. This is a potent reminder, considering that the only paralysis cases caused by a type 2 poliovirus are those of vaccine origin.

As described above, the lack of a standard approach to the treatment of a patient with an acute presentation of poliomyelitis in the background of chronic infection is of significant concern. As a result an initiative has been formed to focus on the creation of strategies for the treatment of infection with poliovirus¹⁸⁶.

Upon the interruption of transmission of wild-type poliovirus, a period of time will pass before the WHO announces the cessation of the use of oral poliovirus as a vaccination strategy. This will be done in order to minimise the risk of the emergence of VDPVs and the associated reduction in risk of paralytic disease associated with these viruses.

1.2.9 Poliovirus Endgame Strategy and the Poliovirus Antiviral Initiative

In 2006, the Poliovirus Antiviral Initiative (PAI) was established in response to the approaching endgame strategy for the eradication of poliovirus¹⁸⁶. The emergence of

Sabin vaccine poliovirus strains in an unvaccinated population is a serious concern to the eradication initiative. Such events are possible, in particular in the case of individuals with a primary immunodeficiency, who are unable to clear a chronic poliovirus infection. Such an infection can be sub-clinical and the patient may be excreting virus into the environment or to close contacts. Over time, the chronic replication of the virus in the gut can lead to point mutations that confer an increase in neurovirulence^{177,182,185}. At present, no therapeutic options exist for the treatment of individuals with a chronic enteroviral infection.

In response to the formation of the PAI, a number of small molecules developed for the treatment of picornavirus infection have been examined for their efficacy in treating poliovirus infection. Figure 1.21 indicates some examples of these antivirals and their activity on multiple targets within the replication cycle. Specific information relating to the defined stages of the picornavirus replication cycle can be found in section 1.2.3.

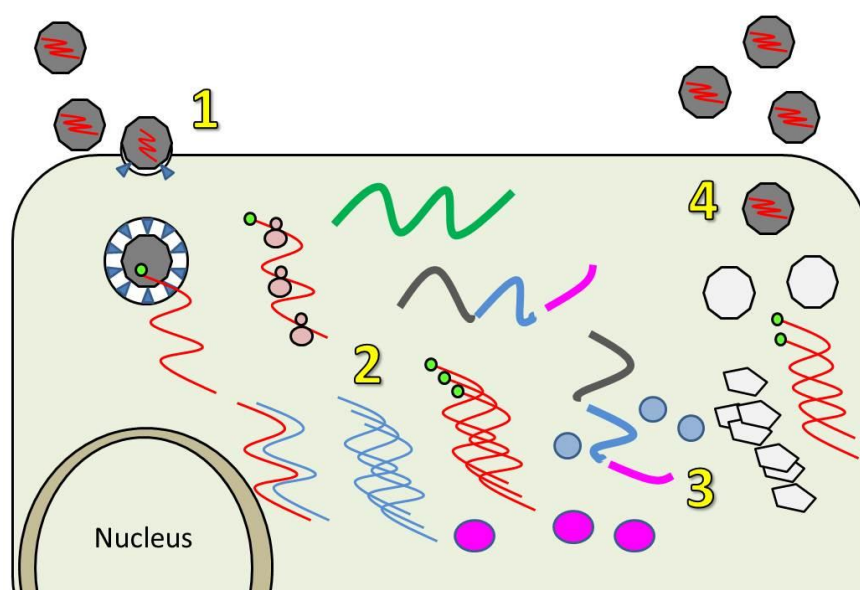


Figure 1.21 Antiviral targets and the picornavirus replication cycle. Simplified schematic representation of the stages of cellular replication of enteroviruses, highlighting stages of virus replication targeted for inhibition using antiviral agents. Figure adapted from De Palma *et al.*, (2008)¹⁸⁷.

- 1) Attachment or uncoating inhibitors such as the capsid binders Pleconaril, BTA-798 or V-073^{188,189}.
- 2) Inhibitors of replication targeting proteins 2C (eg. Guanidine Hydrochloride)¹⁹⁰, 3A (eg. Enviroxime)¹⁹¹ or the 3D polymerase (eg. Gliotoxin, nucleoside analogues)¹⁹².
- 3) Inhibitors of post-translational proteolytic cleavage (eg. Rupintrivir)¹⁹³.
- 4) Assembly and release inhibitors such as Hydantoin¹⁹⁴.

It was the opinion of an expert panel that a multipronged approach would be taken to use antivirals in this manner. Specifically, the choice of antivirals with different targets such as capsid binders and enzyme inhibitors¹⁹⁵ (see Figure 1.21), which is similar to the approach taken for HIV, utilising multiple antivirals targeting different virus products or regions, in order to minimise the risk of drug resistant mutants being created.

The development of potential agents for the inhibition of enterovirus replication and the understanding of their mode of action is aided significantly by the observation of the interactions of drug and target at the atomic level. A significant amount of work has been performed in the determination of enterovirus structure, typically via the use of x-ray diffraction methods, in order to gain a “snap-shot” of virus-drug interactions at the atomic level. The earliest example of the examination of picornavirus-antiviral interaction using x-ray diffraction methods was in 1988, showing the atomic details of WIN compounds binding to rhinovirus capsids^{196,197}.

The following section 1.3 examines common *in-vitro* experimental methods used to determine the atomic structure of molecular structures. A brief historical background for each method is provided, with specific information relating to the principles of the methodology. More importantly, the advantages and limitations of each method are examined. Specific examples are given, relating to the application of these methods with respect to the investigation of virus structure.

1.3 The Determination and Visualisation of Virus Structure.

The elucidation of structural virus components at the atomic or near-atomic level is pivotal to our understanding of virus macromolecular structure. The first indications of the macromolecular configurations of a virus was attributed to TMV and was determined using X-ray diffraction patterns derived from the methods of Stanley and Wendall in 1935¹³ and electron microscopy by Kausche *et al.*, (1939)¹⁹⁸.

The ability to provide detailed structural data has progressed hand-in-hand with an increase in computer processing power. The ability to automate data recording and analysis has resulted in the accumulation of data many orders of magnitude greater than is possible by manual methods. The subsequent analysis of these data using highly efficient software algorithms has also yielded massive increases in resolving power for structural components of viruses^{199,200}.

Combining the power of modern computer technology with traditional methods for structure resolution, has resulted in methods such as transmission electron microscopy (TEM) being capable of resolving complete adenovirus capsids at resolutions matching that of X-ray crystallographic methods²⁰¹. However, even with the latest in computational prediction methods, it is still not possible to predict the structural elements in all proteins^{202,203}. The deduction of final protein conformation must still be performed experimentally and the resultant outputs can be refined when combined with amino acid sequence data.

1.3.1 X-ray Crystallography

X-ray diffraction is the physical outcome of the interaction of X-rays with a crystalline lattice. The principle of x-ray diffraction is based on the passing of very narrow parallel beams of x-ray radiation through a crystalline substrate of protein (or material to be examined)²⁰⁴. If no atoms are encountered, the x-rays pass through unaffected. Where x-rays encounter an atom, the x-rays will be scattered. If the crystalline lattice is well ordered, then the resultant backscatter will occur in an ordered manner and can be captured as a diffraction pattern²⁰⁵.

X-rays form part of the electromagnetic radiation spectrum with a wavelength of approximately 1.0 \AA (0.1 nanometres), but unlike electromagnetic radiation in the visible light range, x-rays cannot be refracted with any known lens. This becomes a significant problem when employing x-rays to examine the atomic structure of a molecule using diffraction methods. In order to reverse calculate the diffraction point of an x-ray and identify the electron cloud responsible for this diffraction in 3-dimensional space using the Fourier transform algorithm, three key points of information are necessary. 1) The angle of diffraction as it relates to the source of the x-rays and the collection point, 2) the amplitude of the x-ray which is derived from the intensity of the data point collected in 2-dimensions and 3) the phase of the x-ray as it makes its way through the crystalline lattice. Data relating to phase, resulting from the absence of any available lens apparatus, remains absent from the collected diffraction data in x-ray diffraction methods. This is referred to as the “phase problem” and present a significant issue in the application of the Fourier transform algorithm required to reverse calculate the derived 2-dimensional data to a 3-dimensional space. Different approaches have been used historically to assist in the elucidation of large macromolecular structures. The multiple isomorphous replacement (MIR) “Patterson” method, was used in the first determination of protein structures in 1958, specifically haemoglobin (Perutz)²⁰⁶ and myoglobin (Kendrew)²⁰⁷, earning them both the Nobel Prize in 1962. Subsequent methods to solve the “phase problem” are the multi-wavelength anomalous diffraction (MAD)²⁰⁸ and molecular replacement (MR) methods²⁰⁹. The methods are computationally very intensive and the resultant electron density maps are then used to fit the amino acid sequence data for the specific protein being investigated.

The theoretical experimental resolution of a density map is a function of the quanta of data points obtained, versus the size of the unit cell examined. For example, a unit cell size $100 \times 100 \times 100 \text{ \AA}$ would require 1000 points in each unit cell to achieve 0.1 \AA resolution for each direction, so $1000 \times 1000 \times 1000$ points (one billion points) would be required to resolve a 100 \AA cubic cell. This theoretical limit is restricted by the intensity of the x-ray source and the rate of destruction of the material being examined over time. One of the most important advances afforded by powerful computer systems and automation, is the ability to approach the theoretical limit of resolution, by rapidly taking thousands of diffraction patterns from a single crystal moved along multiple axes and, using the “Fourier transform” method, back-calculate the positions of individual atoms

with a potential resolution of less than 1 Å to form an electron density²¹⁰. Such resolving capacity obviously requires the assistance of powerful computers and much progress has been made in the automated collection and interpretation of X-ray diffraction data²¹¹⁻²¹³. The final resolution is very important, as it enables the recognition of structural motifs which can facilitate the fitting of amino acid sequence data.

Although the theoretical limit for the resolution of x-ray diffraction methods is less than 1 Å, this figure varies considerably depending on the type of structure being examined and the quality of the crystals available. Spatial variation within the molecule being examined is significant, as disordered or mobile elements of the molecule will result in unreliable density data, for which in the case of proteins, affects the ability to fit amino acid sequence data. The observation of fluctuations in spatial coordinates is referred to as Debye-Waller factors²¹⁴ or “B-factor”, with low values indicating lower vibrational (thermal) states, correlating with better quality data.

In the case of viruses, which are relatively large macromolecular structures with complex interaction often involving repeating subunits, the actual resolution derived using x-ray diffraction methods can be significantly higher than 1 Å. An examination of the RCSB Protein Data Bank at the time of writing, indicates that the resolution of x-ray diffraction derived capsid structures for members of the genus *Enterovirus* the range from 2.2 Å for the wild-type poliovirus structure (PDB ID 1HXS) originally examined in 1985¹⁶, to 4.0 Å for the empty capsid structure of EV-A71 (PDB ID 3VBU) released in 2012²¹⁵.

X-ray crystallographic methods, whilst being able to provide very high resolution data relating to protein structure, are unable to provide a large scale biological juxtaposition of the protein (or virus) being examined. Often the relative positioning of a molecule within a biological system is important. For this purpose, electron microscopy methods are useful, as described in the following section 1.3.2.

1.3.2 Transmission Electron Microscopy

In 1933, Ruska and Knoll created the first prototype transmission electron microscope²¹⁶. In much the same manner as a light microscope passes light through a sample and then a series of lenses which magnify the specimen, transmission electron microscopy (TEM) channels a fine stream of electrons in a vacuum through a thin sample preparation. The

occlusion of these electrons by electron dense material results in an image represented on a fluorescent screen (Figure 1.22).

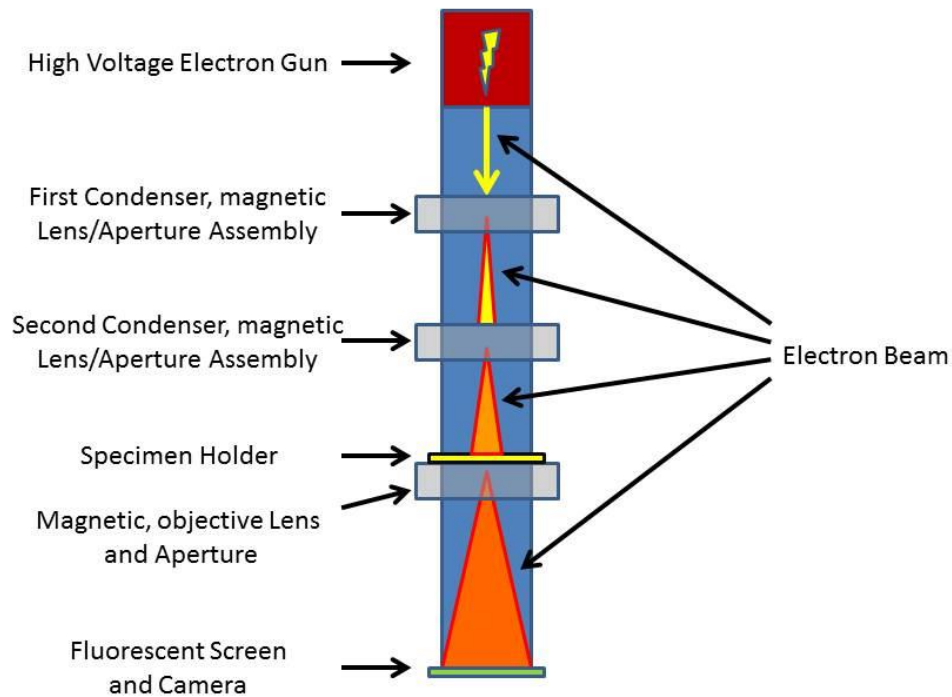


Figure 1.22 Schematic diagram of an electron microscope. A simplified diagram of an electron microscope, indicating the major components and their relative positioning. Electrons for use are generated by a high voltage electron gun (Red). The electrons are accelerated along the tube length and focused using magnetic lens assemblies. Once the electron beam has passed through the specimen, the scattered electrons are detected using a fluorescent screen creating an image which is then recorded using a high resolution camera.

The introduction of the transmission method of electron microscopy enabled the visualisation of TMV virus particles in 1939 by Kausche *et al.*,¹⁹⁸, thereby settling a long standing debate as to whether the TMV agent was a liquid or a particle.

The resolving power of TEM is measured in the millions, with resolutions of 0.5 Å theoretically achievable using high energy systems and methodologies involving crystallography and diffraction, although these methods are not in routine use and in reality, result in structures between 1-10 Å^{217,218}. Typically, lower power TEM systems in the range of 80-300 kVA are used and represent a useful adjunct in virus structural determination, in that they provide a gross morphological description of a virus at resolutions of 20-40 Å and when combined with thin section methods, can also provide a

biological context. These thin section methods can be performed on samples such as cell culture positive material and the virus can be visualised in various stages of its life cycle at an effective resolution of between 50 to 75 Å (Figure 1.23) with ultra-low power versions (5kVA) capable of 100 Å resolutions²¹⁹.

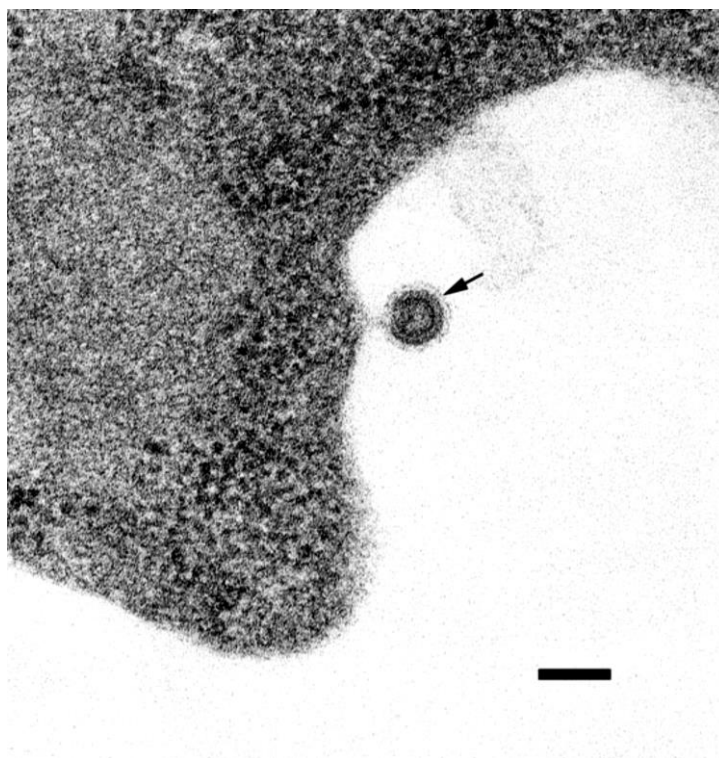


Figure 1.23 Thin section transmission electron micrograph of an L20B cell. An endogenous retrovirus-like particle (arrow) is depicted budding from the cytoplasm of an L20B mouse fibroblast, a cell line expressing the CD155 receptor that is commonly used in the isolation of poliovirus. Scale bar: 100nm. Image courtesy of John Marshall, Electron Microscopy Division, VIDRL²²⁰ (appendix A8.2.4).

Disadvantages of this method include the costs associated with purchasing, maintaining and operating a TEM system. For a period of time, electron microscopy methods were becoming less prevalent, particularly after the introduction of antibody-based virus detection systems coupled with light microscopy and the advent of the PCR era. However, combining TEM with cryogenic freezing methods, computer based imaging and 3-dimensional reconstruction, there has been resurgence in the use of TEM²²¹.

1.3.3 Cryogenic Electron Microscopy

Electron microscopy is a destructive process for the subject matter being examined. Damage to biological specimens begins during the preparation process and may be caused by dehydration, or while attaching the specimen to be examined onto the substrate

before observation. As a result, the specimen observed using TEM techniques does not necessarily reflect the actual biological state of the original sample. Given a long enough time, the sample examined under a stream of electrons will be irreparably damaged^{222,223}.

Initial experiments by Taylor *et al.*, (1975), showed that the process of cryo-electron microscopy (cryo-EM) was capable of preserving frozen hydrated catalase crystals, which were destroyed when prepared in an air dried manner²²². Critical to the process of cryo-EM microscopy was the discovery that rapid cooling of pure water would prevent the formation of ice crystals. Techniques for rapid freezing using liquid nitrogen and/or helium, such as those described by Dubochet *et al.*, (1982)²²³ and McDowell *et al.*, (1984)²²⁴, resulted in a resurgence of electron microscopy techniques, producing results which rivalled those of x-ray crystallography.

Given the appropriate resources, it is possible for cryo-EM to be used under optimal conditions to resolve structures at a resolution of less than 3 Å in vitreous (glass-like) ice at liquid nitrogen or helium temperatures (4K to 77K)^{225,226}. To examine virus structures at high resolutions, the techniques most commonly encountered in laboratories are cryo-electron tomography and single particle cryo-EM²²⁵.

Tomography is the process of imaging via sections. The word tomography is a derivative of the Greek words tomē or tomos (cut, part or section) and graphein (to write). In the context of cryo-electron tomography, a series of images are obtained from the specimen typically frozen in vitreous ice and placed on an electron microscopy stage which is able to tilt. Using computational methods, these images can be reconstructed into a 3-dimensional density, the resolution of which is dependent on the thickness of the original specimen. Typically this requires specimens to be 0.5 μm in order for the electron beam to penetrate the specimen²²⁷. Such work is also typically suited to electron microscopes with accelerating voltages of 200 to 300 kV. For microscopes of low voltages, samples prepared using traditional methods such as negative staining can also be useful for tomography.

Cryo-electron tomography methods can easily produce images with resolutions of 50 to 100 Å which can be useful for cellular structures and bacteria, but in terms of resolving virus structures, these resolutions are not suitable²²⁸. One method that can be applied to

increase the resolution is sub volume averaging, which can provide structural information at 20 Å resolutions. These methods have been applied to viral surface proteins such as HIV trimers and other envelope glycoproteins²²⁹.

One of the most commonly used forms of cryo-EM techniques is the single particle method. This method involves the aggregation of numerous 2-dimensional images showing the particles of interest in different orientations, which are then reconstructed to generate a 3-dimensional representation of the particle (Figure 1.24). This method also involves the plunge freezing of high titre preparations of the particle of interest in order to create a thin layer of vitreous ice. With respect to the examination of viruses using single particle methods, the exploitation of high symmetry such as the icosahedral 60-subunit symmetry found in members of the family *Picornaviridae*, allows averaging of subunits to refine the final structure. This can result in structures of 4 Å or better, approaching the resolution conferred by x-ray diffraction methods^{226,230}.

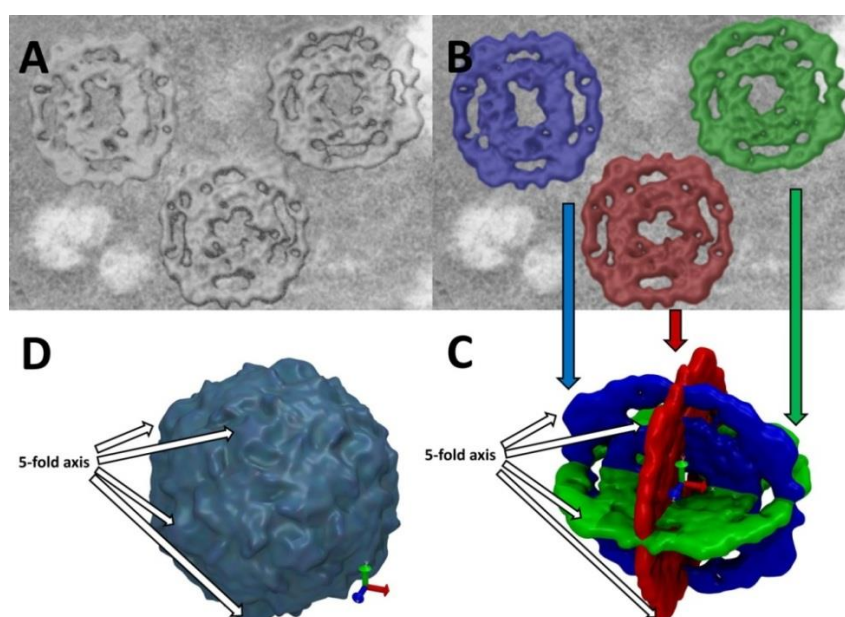


Figure 1.24 Principles of single particle cryo-EM methods. Highly simplified, schematic representation of the principles used in single-particle cryo-EM methods, for the reconstruction of a virus particle. A) Simple representation of the initial electron micrograph containing the molecule of interest in differing axial rotations and planes. B) highlights the differing axial rotations as they relate to C), D) represents the final density constructed from a large number of sample sets similar to those represented in A) and B), subunit averaging of protomers around the five-fold axis of symmetry results in an increase in resolution^{226,230}.

The acquisition of 2-dimensional data and the 3-dimensional reconstruction process is demanding for the researcher and is computationally intensive, but advances in computational software and hardware technology allow more efficient gathering of a larger number of quality 2-dimensional images, which can be used to increase the signal-to-noise ratio, critical to the reconstruction process^{228,231}. Once knowledge of the relative orientation of the particle is established, the Fourier transform of each 2-dimensional projection can be extrapolated to a 3-dimensional object^{225,232}. This method can also be combined with tilt-imaging methods, allowing the inclusion of more 2-dimensional images with known variations in orientation²³³.

Hybrid methods, involving the derivation of cryo-EM densities with atomic coordinate data obtained using nuclear magnetic resonance (NMR) techniques and/or x-ray crystallography, are becoming more frequently used and can result in very high resolution structures²²⁸. The resolution of publicly available enterovirus capsid structures obtained using cryo-EM methods range from 6.09 Å (CV-A7, PDB ID 4AGY)²³⁴ to 30 Å (Poliovirus with receptor, EMDB ID EMD-5144)²³⁵.

1.3.4 Nuclear Magnetic Resonance

NMR or Nuclear Magnetic Resonance spectroscopy is frequently used in the elucidation of small molecules such as proteins, advances in technology and methodologies, such as those incorporating the use of isotopic labelling, have resulted in the determination of structures approaching 100,000 daltons in mass²³⁶. Unlike x-ray crystallography, a crystallised version of the structure (such as a small protein) is not required, nor is the sample required to be fixed in or on a solid substrate as in electron microscopy. In fact all that is required is an aqueous version of the target in a pure and concentrated form.

NMR has been used to examine the structure of the uridylated form of the VPg molecule, which is critical to the reproductive cycle of the virus family *Picornaviridae*. In poliovirus, a member of the family *Picornaviridae*, the VPg is a 22 amino acid protein that is covalently bound to the 5' non-translated region of the virus genome^{237,238}.

NMR relies on the principle of atomic “spin” and the potential of atomic nuclei to be aligned to a strong magnetic field. In sufficiently high quantities, “magnetic moments” of atomic nuclei in molecules can be deliberately misaligned and “tipped of axis” by the use of radio frequency (RF) radiation. The resulting alterations in magnetic field in

response to a specific RF frequency can be observed and recorded in the form of spectral data for analysis. The apparatus that maintains the strong magnetic field, RF control pulses and the detection of magnetic induction is referred to as an NMR spectrometer²³⁹. The emission of RF data by atomic nuclei is directly influenced by neighbouring nuclei and thus information relating to the positioning of hydrogen atoms can be extrapolated. The size of shift in RF signals indicates the distance between atomic nuclei and when combined with sequence data, the 3-dimensional structure can be derived along with information pertaining to variations in conformational state of the protein²⁴⁰. Structures solved using NMR are typically presented as “sets” of data.

Although NMR allows the inference of differing conformational states of a protein, none of the methods described using x-ray diffraction or cryo-EM can enable the observation of large-scale molecular movement and interactions over a given time-frame. End-point observations of molecular structures are obtained and often the methods described above are combined to provide a detailed and complete picture of macro-molecular structures²⁴¹. A significant amount of effort and resources are required to observe virus structures at the atomic level, using the techniques just reviewed. Thus, it may not be feasible to elucidate the capsid structure of an enterovirus identified to have the potential to be a public health threat, in a timely manner. It is for these reasons, that the prediction of virus structure using methods such as comparative protein modelling is useful.

1.4 Virus Structure Prediction Using Comparative Protein

Modelling Methods

The elucidation of structures using conventional methods such as x-ray crystallography or cryo-electron methodologies, rely upon the availability of large amounts of highly purified protein to be investigated. In the case of cultivable viruses this is not necessarily an issue, as large numbers of virus particles can be produced in the laboratory using standard cell culture methodology, combined with high speed centrifugation and purification techniques. The amount of time to perform these procedures is still a significant consideration, and this is particularly evident in the case of x-ray diffraction methods, which are reliant upon the availability of a crystal for examination. The production of crystals for x-ray diffraction has evolved significantly over time and a number of viable commercial applications are available. With respect to enteroviruses from high titre cultures, a number of x-ray diffraction-derived structures representing virus particles in different states are available and a lot has been learned from the early days of virus structure prediction, but such work remains nonetheless time-consuming and expensive.

In the case of non-cultivable viruses, such as the recently described *Rhinovirus* species C and other members of the genus *Enterovirus* affecting humans, high virus titre methods are not available for purification and for use in x-ray diffraction methods or cryo-EM methods²⁴². The use of comparative protein modelling methods, which will now be reviewed, is a viable alternative in the absence of methods using infectious DNA clones for the production of complete virus particles, such as those described by Racaniello *et al.*, (1981)³⁴.

The application of 3-dimensional structural prediction methods for protein sequences derived from nucleic acid sequence data, has allowed the reconstruction of novel proteins by comparison with homologous data²⁴³. The Protein Data Bank (PDB) is an ever-growing resource of 3-dimensional structural data that can be used to predict homologous structural motifs²⁴⁴. These methods can be used to reconstruct novel proteins with homologous regions such as enzymatic active sites, or structural motifs such as beta barrels, commonly found in the capsids of viruses. Whilst the reconstruction of common motifs such as β -barrels or α -helices results in models of an acceptable quality, a problem arises when loop regions of a protein are modelled using comparative methods.

The problem of modelling these loops can be attributed to three main causes:

Firstly, missing data in the template structure due to thermal fluctuation or spatial disorder, that cannot be resolved using methods such as x-ray diffraction, complicates the prediction of appropriate loop positioning²⁴⁵. Secondly, if an analogous loop structure is available for comparative modelling, insertions or deletions may exist in the query sequence when compared to the template structure (as a consequence of evolution), complicating the threading of query sequence within the 3-dimensional coordinates afforded by the template structure²⁴⁶. Thirdly, it may be that some “loop” regions contain sequences that are highly flexible and exhibit numerous conformations at finite temperature, or have no intrinsic structure²⁴⁷, in which case x-ray diffraction may not be useful. Conditions required to force a loop into a particular conformation, so it that can be resolved, might not be biologically relevant and in these cases, molecular dynamics (MD) simulation is necessary.

These problems are particularly evident in the loop-like structures that are interspersed between the extended-beta formations that comprise the “jelly-roll” structures, central to the subunits of a protomer in the capsid of enteroviruses, as described in section 1.2.2. Loops such as the BC loop of VP1 and the EF loop of VP2 are known antigenic determinants and as such, undergo selective pressure by the host, resulting in hyper-variability of the amino acid sequence²⁴⁸.

Previous study of virus proteins in the early 1990s, utilised comparative protein modelling techniques that were initially limited to the investigation of non-structural proteins, such as viral proteases or polymerases, based on homology with previously elucidated structural information^{249,250}. It is not surprising that the complexity of comparative protein modelling efforts has had a strong correlation with the increased availability of computational power, which in the 1990s was beginning to make significant inroads due to reduced cost.

By the latter half of the 1990's, comparative protein modelling techniques had extended to the examination of interactions of multiple molecules such as receptor molecule interactions and virus glycoproteins²⁵¹⁻²⁵³. Computer limitations still existed in at this stage, which made the reconstruction of structures measuring hundreds of Å in diameter not only difficult, but essentially impossible.

In 1999, Yoneda *et al.*, described the reconstruction of an icosahedral viral capsid using symmetry methods in the reconstruction of poliovirus, using rhinovirus X-ray crystal data as a template for comparative protein modelling²⁵⁴. Although not a complete capsid reconstruction, this work gave valuable insight into the limitations of modelling protomer subunits as opposed to that of a complete virus capsid, even when performed using symmetry methods to reconstruct a capsid.

Enterovirus structural proteins were targeted for comparative protein modelling in 2001²⁵⁵. This timing coincided with the emergence of an increased incidence of neurovirulence caused by EV-A71, which previously, was mainly associated with HFMD. The use of homology methods allowed the reconstruction of virus surface structures and the hydrophobic pocket associated primarily with the VP1 chain of the virus capsid. These EV-A71 models were reconstructed using homologous regions found within the structural regions of poliovirus and bovine enterovirus.

Given the high degree of sequence identity between the human enteroviruses and the availability of verified structures in the form of x-ray crystallography, cryo-EM and nuclear magnetic resonance data, enteroviruses make ideal candidates for comparative protein modelling. The generation and validation of homology models derived from existing protein data has the potential to facilitate the application of MD simulation to elucidate the molecular mechanics underlying viral pathogenesis, including receptor-binding, cell entry, RNA release and temperature dependant stability of critical enzymes of viral origin. Such data would ultimately result in the recognition of new targets for the development of antivirals.

An examination of the historical aspects of MD simulation methods and the application of these methods to virus structural research is presented in the following section 1.5. Comparative protein modelling methods combined with atomistic MD simulation are used extensively in the research work presented in this thesis, the results of which are presented in chapter 5.

1.5 Molecular Dynamics Simulation

Molecular modelling and mechanics are based primarily on the concept that molecules can be represented as mechanical systems subject to Newtonian (or classical) physics, enabling various molecular properties to be calculated, including energy and particle juxtaposition in geometric terms. In essence, this involves solving Newton's equations²⁵⁶ (1.1, 1.2 below) for each atom (i) as defined by:

$$F_i = m_i a_i \quad (1.1)$$

$$m_i \frac{d^2 r_i}{dt^2} = - \frac{dV}{dr_i} \quad (1.2)$$

Where:

F_i = force on the particle i

m_i = mass of the particle i

a_i = acceleration of the particle i

r_i = Cartesian position vectors of the particle i

t = time

V = potential energy

Atomic interaction potentials (V) in classical molecular mechanics models can be described using a number of types of force fields. The most commonly applied force fields for biological systems involve only pairwise interactions of particles, where these force-fields are derived for the most part from empirical observations^{257,258}.

Force fields involve bonded and non-bonded terms. Bonded terms involve forces which act between bonded atoms and are approximated using harmonic "spring" potentials (balls and springs). These interactions can be expressed in terms of bond lengths, bond angles and torsional rotations (Figure 1.25)^{258,259}.

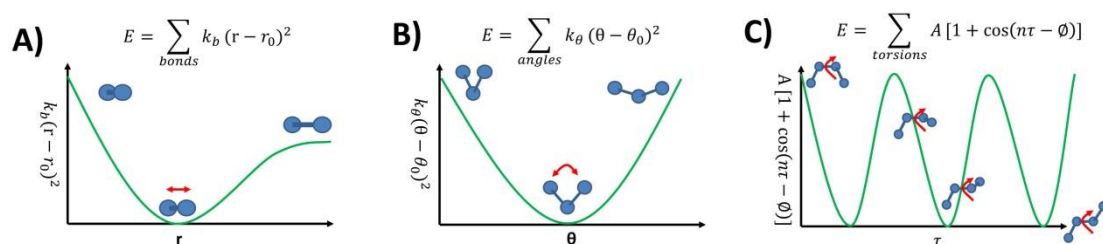


Figure 1.25 Bonded particle interactions and the calculation of associated energies. A) bond stretching energies, B) bond angle energies, C) bond torsion energies. Figure adapted from Fox²⁵⁹ and Lindahl²⁵⁸.

Empirical parameters k_b = bond spring constant, k_θ = angle spring constant, A = torsion rotational energy barrier height.

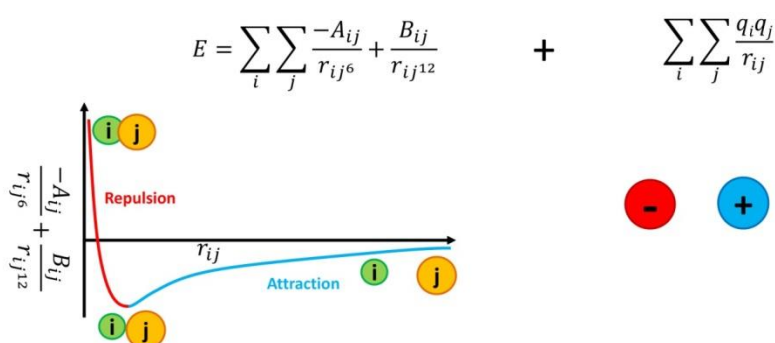


Figure 1.26 Non-bonded particle interactions and the calculation of associated energies. Van Der Waals energy calculations combined with electrostatic calculations. Figure adapted from Fox²⁵⁹ and Lindahl²⁵⁸. r_{ij} = interatomic distance between atoms i and j , $A_{ij} B_{ij}$ Lennard-Jones coefficients for an atom pair (attraction-repulsion), $q_i q_j$ Columb partial charges.

In any molecular system there also exist non-bonded interactions for which the primary terms are van der Waals forces and electrostatic potentials (Figure 1.26). Indeed, molecular mechanics has its roots in the concepts of molecular bonding and van der Waals forces, with the fundamental ideas dating back to the early 1930s. In the molecular mechanics model, the potential energy V is written as a sum of bonded (Figure 1.25) and non-bonded terms (Figure 1.26) The solving of Newton's equations in the context of molecular mechanics force-fields within a system of particles, enables the determination of atomic trajectories when the atoms "interact" over time, thus permitting the inference of motion to a static system^{256,259}. Whilst this approach appears relatively straightforward, in reality, the calculations for large molecular systems are not feasible to solve in any meaningful time-frame. The solution to this problem is to achieve integration of these

equations at a number of discrete time steps, usually calculated at 1 to 2 femtoseconds, thereby generating step-wise atomic motions²⁶⁰.

At present, only a relatively small number of atoms can be simulated, in the order of several million atoms. An example used in this research is a poliovirus and associated water and ions (~3.5 million atoms). This presents another problem, since any simulation would in essence be of a virus surrounded by a thin sheet of water, where the water would behave like it does at a surface, not in a submerged aqueous state. This is not biologically realistic, since surface effects would cause artefacts on the virus structure^{261,262}. In order to partly solve this problem, periodic boundary conditions (PBC) have been created, which represent a simple rule that defines a finite box around the virus, solvated with water^{249,250}. Any atom that crosses the edge of the box boundary is mathematically removed and transferred to the opposite boundary, with the same velocity (speed and direction) as before, essentially creating a hall of mirrors where each atom "sees" an infinite number of atoms as a reflection of themselves, thus approximating viruses solvated in an infinite "bulk" solution²⁵⁶.

Periodic boundary conditions are an elegant solution, but they are still problematic. In an MD simulation, electrostatic effects on atoms must be taken into account. Since electrostatics are long-range forces, and each atom "sees" an infinite number of other atoms, then it becomes intractable to calculate electrostatic interactions in a PBC system. To counteract this problem, certain strategies have been devised. (i) Cut-off methods, where electrostatic interactions beyond a specified radius are reduced to zero, either abruptly, or using a smoothing function. (ii) Ewald summation such as the Particle Mesh Ewald (PME) summation²⁶³, which involves deconstructing electrostatic energy calculations into long-range and short-range components. Short-range components are summed within a defined radius in a manner similar to cut-off methods, whilst the long-range electrostatic components undergo Fourier-transformation and are thus summed in Fourier-space, enabling rapid convergence. The PME method, which is used in all simulations presented in this thesis, have been demonstrated to produce structures more accurately than simulations using cut-off methods²⁶⁴. An important aid to the progression of molecular dynamics simulation has been the evolution of computational power. The evolution of methods to utilise this power to generate simulation data, is discussed below.

In 1946, Frank Westheimer's calculations of the stereoisomers of biphenyl derivatives enabled the definitive determination of molecular mechanics²⁶⁵. One of the primary reasons for the relatively slow progression of the field during this time was the lack of appropriate computational solutions. Quite simply, adequate computer systems were not available. This observation is confirmed by the fact that at the time of Frank Westheimer's work in 1946, what is considered to be the first general purpose electronic computer was born. It was housed at the United States Army's ballistic research laboratory and was known as the "Electronic Numerical Integrator And Computer" (ENIAC)²⁶⁶. Although it used decimal arithmetic, ENIAC still required manual rewiring via plug-board, to change its programming.

In the 1950s, transistor-based systems began to replace vacuum tube counterparts, resulting in the first transistorised computer being demonstrated at the University of Manchester in 1953^{267,268}. Increases in manufacturing efficiency resulted in wider availability of computational power in the 1960s to 1970s. This timing coincided with the initial development of systematic force fields, based on observational methods in the form of spectroscopic information, heats of formation and the sharing of small compound structures with similar chemical groups, combined with quantum mechanical information and other experimental data. These works were pioneered by Lifeson at the Weizmann Institute of science, Scheraga at Cornell University and Allinger at Wayne State University²⁶⁹.

In 1969, Lifeson and Levitt reported the first energy calculation of myoglobin and lysozyme, which is the first instance of such calculations of entire proteins²⁷⁰. These experiments paved the way for the modern day implementation of energy minimisation methods. Subsequent to this, Rahman and Stillinger described the first method for the dynamics simulation of a polar molecule, water^{271,272}. By the 1970s and the availability of supercomputers, Karplus et al at Harvard University had created the first generation of biomolecular force fields, on which modern force fields are based, including the force fields used for the work in this thesis - Chemistry at HARvard Macromolecular Mechanics (CHARMM)²⁷³.

Perhaps the most famous and widely referenced MD simulation was performed in 1977, on the relatively tiny protein bovine pancreatic trypsin inhibitor (BPTI)²⁷⁴. This simulation was within the picosecond timescale and indicated the atomic fluctuations that occur during such a timeframe.

With respect to MD simulations, timeframe is extremely important. The amount of time required for a simulation, is dependent on the questions being asked. For example, to examine ion transport mechanisms, timescales in the order of nanoseconds would be required²⁷⁵. To examine tertiary structure formation of proteins, these timescales approach milliseconds in length and require very substantial computational resources as a result (Figure 1.27)²⁵⁸.

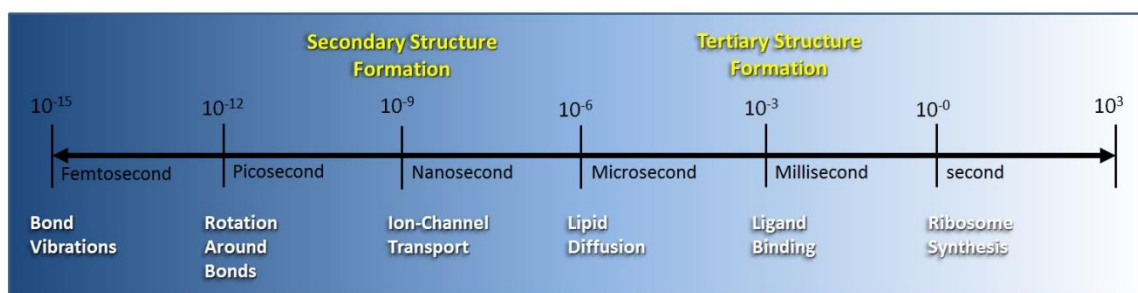


Figure 1.27 Graph of common molecular activities and the associated timeframes. Timeline shown in seconds starting from simple atomic bond vibrations calculated in femtoseconds extending beyond protein tertiary structure formation in milliseconds. Figure adapted from Lindahl²⁵⁸.

For an example of how computational resource power limits the ability to simulate large structures for long timescales, consider the following comparison. A five base pair DNA simulation performed in 1985, consisting of 2800 atoms with a trajectory length of 0.5 ns, required 20 hours of simulation time on a Cray X-MP supercomputer²⁷⁶, which was capable of an estimated 0.8 giga-floating operations per second (GFLOPs) of double precision calculation capacity. Some 20 years later in 2006¹, a one- million atom simulation of a complete virus particle was performed using a system with approximately 0.8 terra-floating operations per second (TFLOPs) for a total trajectory time of 50 ns, over a period of 55 days, translating to approximately 1 nanosecond per day. This represented a simulation size increase of approximately 350 times with double the trajectory data calculated per day, compared to the work performed in 1985. Current

supercomputers are now capable of 10^{15} floating operations per second (PFLOPs) Figure 1.28, with 10^{18} (EFLOPs) supercomputer systems on the horizon.

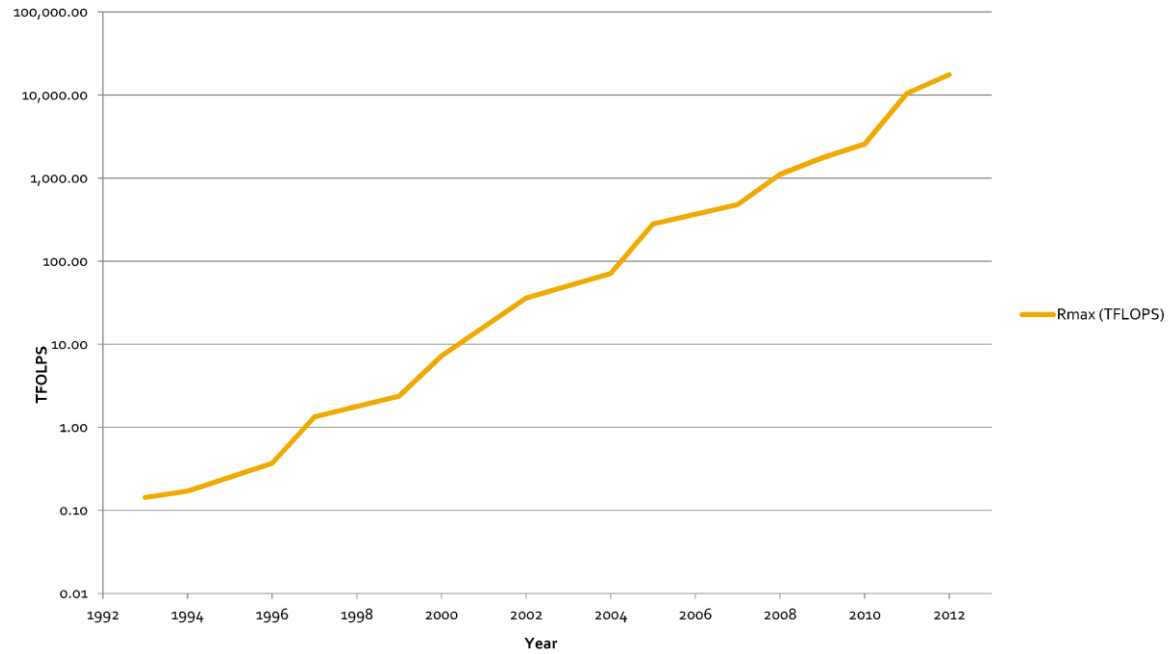


Figure 1.28 Evolution of supercomputer computational power over time. Line graph depicting the increase in supercomputer calculation capacity expressed in TFLOPs (10^{12}) using double precision data processing methods calculated using the LINPACK benchmark²⁷⁷. y-axis is logarithmic in scale.

1.5.1 Simulation of Virus Components

The increase in computational capacity has served not only to facilitate the derivation and examination of static structural data such as those derived using x-ray diffraction methods, but has also allowed the incorporation of temporal variables via the use of MD simulation.

Initial work in the simulation of virus structures was limited to relatively short timescale (picosecond) simulations, examining protein-drug interactions for rhinoviruses and 3-dimensional predictions for amphipathic segments of the HIVgp41 envelope protein in 1989^{278,279}. Extension of these methods became more commonplace as a complementary feature of structure derivation for influenza virus hemagglutinin, and the examination of receptor binding in the 1990s²⁸⁰. Ultimately by the turn of the 21st century, the use of computational methods had assisted in the refinement of potential antiviral agents termed “rational drug design”, which resulted in the development of the influenza neuraminidase inhibitors amantadine (Symmetrel), rimantadine (Relenza) and seven years later, GS4104 (Oseltamivir)^{281,282}. During this same period, the simulation of rhinovirus capsid proteins had progressed to a more extensive examination of the vibrational states of the virus, in relation to the capsid inhibitor WIN52084s^{283,284}.

By 2008, the use of computational methods for the rational design of antiviral agents had become more widespread, resulting in the application of the technique to the design of a Hepatitis C drug (Boceprevir), targeting a shallow pocket within the virus protease. This was a challenging proposition, due to the shallow nature of the pocket with respect to the surface features of the protease²⁸⁵. Given the increases in computational power available to researchers in the 21st century, the evolution of simulation methods for virus components into the simulation of complete virus structures, had become inevitable.

1.5.2 Simulation of Complete Virions

In 2006, the first MD simulation of a structure representing a complete virion at the atomic level was published¹. The virus simulated was that of STMV and consisted of approximately 1,000,000 atoms including water and associated ions required to suspend the virtual virus in solution (Fig 1.29). Although the protein's composition was representative of the virus and was derived from x-ray crystallography experimental data, the encoding RNA in this simulation was not the actual genomic sequence for the virus. Instead, the RNA sequence consisted of multiple tracts of adenine and uracil, in order to facilitate ordered folding of the RNA.

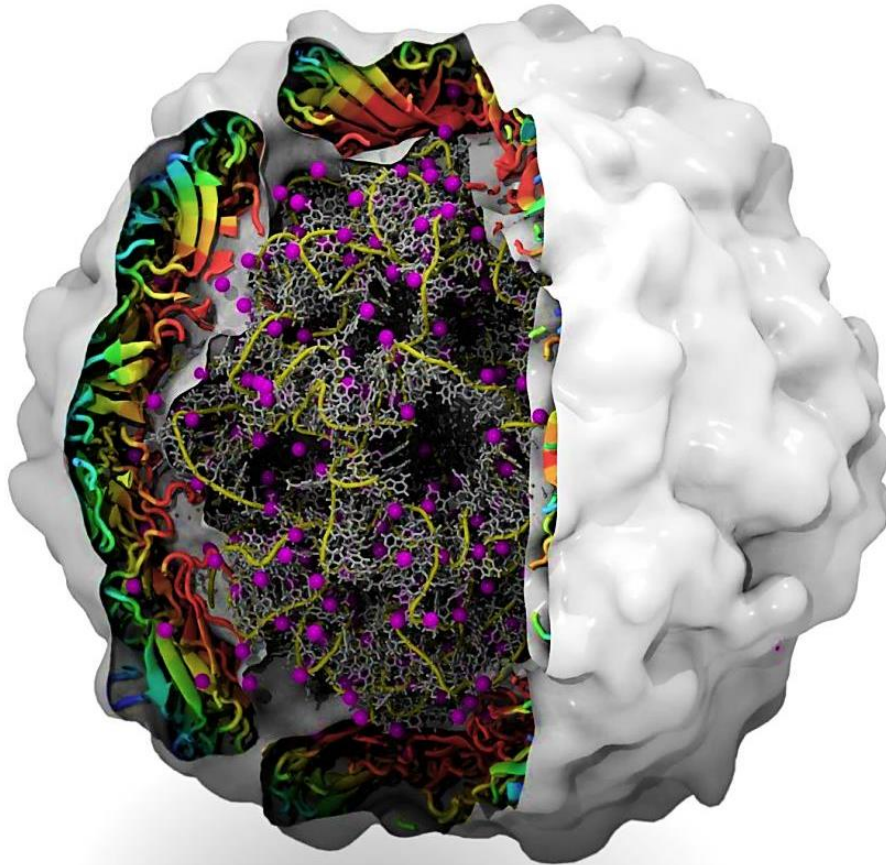


Figure 1.29 Illustration of the Satellite Tobacco Mosaic Virus simulation. Capsid represented as a density at eight Å in grey structural proteins represented in new cartoon (red – 70 Å, blue – 90 Å from centre. RNA displayed in a “tube and liquorice” representation, with associated magnesium ions in purple. Total simulation size was approximately one-million atoms. Representation derived from original data and rendered using tachyon renderer in VMD.

No other complete virus structures have been described in literature since, with the exception of an attempt to reconstruct influenza virus from partial x-ray diffraction data for picosecond timescales, from which limited information was published²⁸⁶. Other simulations of complete virus capsid structures have been performed, but this has been done using coarse-grained modelling (Monte Carlo methods), which is likely due to the very large simulation size (millions of atoms) and the associated amount of computational resources required to run the simulations for any relevant timescale²⁸⁷. For example, multi-million atom simulations require weeks of simulation time to achieve nanoscale times on high-end institution-based supercomputers, which is expensive and time consuming^{2,266} (refer to chapter 4).

The trend of requiring institutionalised computational resources is beginning to change, with the increased incorporation of graphics processing unit (GPU) based parallel processing methods in high-end workstations, such as those used in graphics manipulation industries. Such systems lag by only five years behind the required computational resources for running multi-million atom simulations. For example, to run a 4 million atom simulation for a 100ns timescale within one month would require a 50 TFLOPs system. Currently, it is possible to purchase a commodity system with a 5 TFLOPs capacity, due to increases in computational capacity afforded by GPUs²⁸⁸. If this trend is to continue logarithmically, in a similar manner to that shown in figure 1.28, then it is reasonable to expect that commodity systems capable of exceeding 50 TFLOPs double precision computing capacity, will be available before the year 2020.

1.6 Project Aims

This project aimed to exploit the advancements made in computational biology and bioinformatics and to apply them to the examination of viruses that are of significant public health concern, namely the genus *Enterovirus*.

The areas of interest targeted by this research were development and enhancement of enterovirus detection and characterisation methods via the use of sequence data-mining and phylogenetic examination, in order to exploit similarities and differences in enterovirus genomes. It was anticipated that more specific detection assays could be developed by examination of hyper-variable regions related to major antigenic determinants for the rapid identification of serotype (chapter 2). Phylogenetic inference of these regions was performed and it was anticipated that these data would be amenable to the investigation of virus transmission patterns (chapter 3).

The use of molecular epidemiological methods to determine enterovirus transmission patterns was of use in recognising enteroviruses that pose a specific threat to human health (chapter 3). Viruses of public health significance were subjected to further examination using *in-silico* methods incorporating predicted protein models of the capsid structure and MD simulation, to determine their integrity (chapters 4 and 5). Once appropriately capable computational models are developed, it is postulated that they will assist in the development of methods to predict the effectiveness of antiviral agents targeting the capsid region, specifically the hydrophobic pocket, for which a number of drugs already exist (chapters 4 and 5).

Chapter Two

Development of a Novel Enterovirus

Detection and Super-Speciation

Assay

2.1 Introduction

The continuous increase in computational power, combined with remote database accessibility via the internet, has had a direct and measurable effect on frontline diagnostics and public health surveillance²⁸⁹⁻²⁹². The ability to store, access and process large amounts of data, specifically in the form of nucleic acid and amino acid sequences, gives diagnostic and public health laboratories the ability to examine pathogen diversity at the genomic level and to design specifically targeted genetic assays, typically in the form of PCR methods. This power, combined with the use of modern bioinformatics approaches, allows sequence data derived from PCR-based assays to be analysed in the context of a broader range of sequences available in the public domain, via databases such as GenBank (refer to section 1.2.5).

The introduction of PCR techniques into the public health laboratory repertoire has enabled the rapid identification of pathogens of all descriptions from all sites of the body and is able to provide results in timely manner, with a beneficial impact on treatment strategy for clinicians (refer to sections 1.2.4 and 1.2.5).

The use of PCR methods is not exclusively in the domain of human diagnostic pathology. PCR detection methods have been successfully applied to studies of the environment, for example in the examination of environmental samples from sewage, as an adjunct to public health surveillance programs²⁹³. Environmental surveillance, when integrated with PCR methods and sequence analysis, can be a powerful tool for public surveillance and epidemiology through the detection of specific pathogens. Such strategies are currently being implemented in the Global Polio Eradication Initiative for the surveillance of wild poliovirus and VDPVs^{184,294}.

In the context of public health surveillance, the high throughput nature of PCR methods, particularly in the form of real-time technology, makes the technique especially amenable to outbreak investigation and management.

A suitable example involved the importation of wild poliovirus in 2007, by a traveller returning to Melbourne, Australia^{137,161}. The patient returned to Australia from a polio endemic region in Pakistan, exhibiting mild symptoms of AFP. At the time, the standard

procedure at the National Polio Reference Laboratory in Melbourne was to perform virus culture of stool samples as per the WHO recommended procedures and wild poliovirus was isolated after four days. Relying upon cell culture as the principal test proved to be problematic, in that the requirement to rapidly expand testing of samples to the traveller's close and public contacts, increased the work-load in a manner that stretched the capability of standard cell culture methodology. In such a situation, the turnaround time of sample results, including negative results, is critical to the coordinated public health response at the state level and potential at a national level²⁹⁵.

A more suitable solution would have been to employ real-time RT-PCR methodology, which could accommodate up to 96 samples including controls per test, with a throughput of up to 384 samples a day. Such a method, for the specific detection of poliovirus direct from clinical specimens, did not exist.

Due to the conservation of RNA sequence in the 5'NTR, a method based on existing pan-enterovirus 5'NTR detection would be unlikely to adequately differentiate poliovirus from other enteroviruses. Detection of poliovirus using PCR-based methods directed at the hyper-variable regions of the capsid encoding sequence lack sensitivity when compared to standard cell culture methods, as was the case of the 2007 importation event with a low titre wild poliovirus¹³⁷.

As a result of this experience, the development of PCR-based methods that could be easily scaled up, to differentiate potential poliovirus positive samples from other enteroviruses whilst maintaining sensitivity in a standard laboratory workflow, were investigated. It was also envisaged that such a method would be of value for surveillance in the immediate period after wild poliovirus eradication, to screen samples for the presence of poliovirus or species C *Enteroviruses*. A PCR-based method to screen for poliovirus directly in clinical specimens could drastically reduce the amount of time required when compared with sequencing, to identify the enterovirus type.

In order to increase sensitivity, and specificity, it is possible to use PCR methods using nested approaches, that is, the amplification of nucleic acid sequence using two rounds. The first round of PCR primers, flank the second round PCR inner primers. When combined with probes selected to target variations of the amplicon sequence, multiple

targets can be identified in the same sample. Such methods, termed multiplex PCR, have previously been used in diagnostics laboratories¹⁰⁸.

The application of multiplex, nested or semi-nested PCR methods with a real-time second round PCR has the advantage of maintaining specificity and providing a high level of sensitivity. Real-time PCR also permits the differentiation of amplicons that are of equal length but contain variations in sequence, as the detection of PCR product relies on the detection of specific sequence motifs rather than the amplification of templates of a specified length as in end-point PCR methods, refer to section 1.2.5.

The introduction of real-time PCR machines with multiple channel fluorescence detection, such as the ABI 7500 system, permitted up to five fluorescent probes emitting light of different wavelengths to be used in a single test. The use of multiple fluorescent channels allowed the incorporation of reagents containing a known quantity of a PCR target such as a virus, protected nucleic acid sequence or genomic sequence that is ubiquitous in the samples being tested. Examples of these internal controls include equine rhinitis virus, glyceraldehyde 3-phosphate dehydrogenase or β 2-microglobulin^{296,297}. This addressed the reporting of false negatives due to PCR inhibition. If a clinical sample contains PCR inhibitors, the internal control will not be amplified, flagging the sample as invalid. This is important in the context of enterovirus detection, given that stool samples and CSF samples, which are the most commonly received samples when investigating neurological illness, are inherently prone to inhibition.

This chapter describes the development of a novel real-time PCR assay that can be used to differentiate between *Enteroviruses* belonging to species A or B from those belonging to species C or D, termed enterovirus super-speciation. Initial development, in the form of primer and probe design, was achieved using bioinformatics methods, incorporating the acquisition of sequence data from the public domain and analysis using multiple sequence alignment and phylogenetic inference, to determine areas of the enterovirus genome that could be exploited to produce a suitable diagnostic method for routine use in a diagnostic laboratory.

2.2 Materials and Methods

2.2.1 Sample Preparation

Upon receipt, two grams of faecal specimen were treated with Minimum Essential Medium containing Earle's salts, chloroform (9.1% v/v) and foetal bovine serum (2%) in a final volume of 10 mL. The suspension was clarified using a refrigerated centrifuge at 4°C and 1500 g for 20 minutes. The supernatant was removed and 200 µL inoculated onto a series of mammalian cell lines. The remaining supernatant was stored in a -30°C non-cyclic defrost freezer for subsequent use. For reagent and equipment details, refer to appendices A1 and A2.

2.2.2 Growth of enteroviruses in cell culture

Cell lines used for the isolation of enteroviruses were:

- BGMK (buffalo green monkey kidney)
- HEL (human embryonic lung).
- Hep2C (human epidermoid carcinoma)
- L20B (a transgenic mouse epithelial cell line expressing the human poliovirus receptor, CD155)
- RD-A (human rhabdomyosarcoma)

Cell culture media was prepared using the following ingredients (table 2.1):

	Eagle MEM (Hanks) With L-Glutamine	Heat Inactivated FBS	NaHCO₃ (7.5%)	1M HEPES	Penicillin/ Streptomycin 100x
2% Growth Medium	500 mL	10.5 mL	3.6 mL	5.3 mL	5.3 mL
10% Growth Medium	500 mL	57.5 mL	5.1 mL	5.7 mL	5.7 mL

Table 2.1 Cell culture medium ingredients: Preparation of cell culture 10% growth media containing 10% foetal bovine serum (FBS) and 2% maintenance media containing 2% FBS was prepared using the concentrations listed.

Inoculated cell cultures were observed for up to 10 days for changes in monolayer appearance indicative of a viral CPE. The infected cells were harvested when CPE

involved 50 to 100% of the cell monolayer. The infected cells were then subjected to a freeze-thaw cycle at -30°C to lyse any remaining intact cells. The tubes containing the infected cells were then centrifuged at 1500 g for 10 minutes and the supernatant separated and stored at -30°C in preparation for re-inoculation onto fresh cells. Upon observation of viral CPE, the passaged cells were harvested and stored at -30°C for further examination.

2.2.3 Nucleic acid extraction

Viral RNA was extracted using either the High Pure Viral Nucleic Acid Kit (Roche cat# 11 858 874 001), QIAamp Viral RNA kit (QIAGEN cat# 28704), or MagMax (Life Technologies cat#AM1836) automated robotic extraction system using the manufacturer's recommended procedures. For each RNA extraction, 50-200 μL of cell culture supernatant was used, depending on the system used, and a final volume of 50 μL was collected. For reagent and equipment details, refer to appendices A1 and A2

2.2.4 Preparation of Control Material

Sabin poliovirus controls for each serotype were obtained from the National Institute for Biological Standards and Control in the United Kingdom. Virus stocks for EV-A71 sub-genogroup B5 were obtained from my son, William Roberts, during the course of an acute infection, and then propagated in RD cell lines.

The determination of virus titre was performed as follows: Serial dilutions of control material was performed by dispensing 900 μL of 2% maintenance medium to 6 tubes and 4.95 mL medium to an additional 4 tubes. Virus control material was thawed rapidly and 100 μL of the virus solution added to the first tube containing 900 μL of medium and mixed. This tube represented working stock solution at a 10^{-1} dilution. Subsequent dilutions were made from this stock solution to a final dilution of 10^{-10} . The titrated virus preparations were added to a microtitre plate with rows A-H and columns 1-12. 100 μL of 2% maintenance medium was added to wells A11 to H12, representing negative cell controls and 100 μL of virus dilutions from 10^{-7} to 10^{-10} were added to wells in rows A to H using 20 wells per virus dilution. For EV-A71 the dilution series was made representing 10^{-5} to 10^{-9} .

Once the titrated virus and negative control solutions were added to the microtitre plate, 100 μ L cell suspension containing 1.5×10^5 cells / mL were added to all wells. The plate was then covered with a non-toxic sealer and incubated at 36°C. The plates were examined for development of CPE, using an inverted microscope at 5 and 7 days. At the 7th day reading, virus titre was calculated using the Kärber formula and recorded.

Negative controls consisted of nuclease free water or 2% growth medium. The inclusion of an internal control to account for PCR inhibition was vital and bovine viral diarrhoea virus (BVDV), which has a single-stranded positive sense RNA genome, was included in the RT-PCR amplification steps. The BVDV RNA and initial guidance for primer design were provided by Dr Julian Druce of the Viral Identification Laboratory, VIDRL. The BVDV control material was originally obtained from the Elizabeth MacArthur Institute, NSW.

2.2.5 Super Speciation Pan-enterovirus Reverse Transcription Polymerase Chain Reaction.

The potential inhibition of PCR products due to substances found in template material is a significant issue. In addition to the detection and differentiation between members of the A/B and C/D *Enterovirus* species (super-speciation), the assay also accounted for the presence of inhibitory substances via the use of BVDV as an internal control. The PCR primers and probes for the enterovirus super-speciation assays are listed in table 2.2 and the internal control primers and probes for the detection of PCR inhibition are listed in table 2.3.

Name	Oligonucleotide Sequence 5' to 3'	Position ¹
EV1F (Sense)	CAAGCACTTCTGTTTCCCCGG	162-182
EV2F (Sense)	TCCTCCGGCCCTGAATGCG	443-462
EV3R (Anti-sense)	ATTGTCACCATAAGCAGCCA	596-577
EV probe	FAM-AACCGACTACTTTGGGTGTCCGTGTTTC-BHQ1	535-562
EV-AB probe	CR610-ACTCYGCAGCGGAACCGACTACT-BHQ2	523-545
EV-CD probe	FAM-AGTCYGTGGCGGAACCGACTACT-BHQ1	523-545

Table 2.2 Primer sequences for EV Semi-nested Super-speciation PCR. ¹ position relative to poliovirus 1 Mahoney, Genbank accession number V01150.

Name	Oligonucleotide Sequence 5' to 3'	Position ¹
BVDV85for (Sense)	CGAAGGCCGAAAAGAGGCTAGC	85-107
BVDV390rev (Anti-sense)	TCCATGTGCCATGTACAGCAGAG	369-391
BVDV107for (Sense)	CATGCCCTTAGTAGGACTAGCA	107-129
BVDV246rev (Anti-sense)	CCTCGTCCACGTGGCATCTC	231-250
BVDV probe	CO560-AGTACAGGGTAGTCGTCAGTGGTTCGA-BHQ1	181-207

Table 2.3: Inhibition control primer sequences for EV Super-speciation assay. ¹ Position relative to bovine viral diarrhoea virus 2 (BVDV) genome Genbank Accession Number KC963968

Preparation of the round 1 EV Super-speciation snrRT-PCR primer mix was as follows.

EV1 primer (50 µM)	20.0 µL
EV3 primer (50 µM)	20.0 µL
<u>Nuclease Free Water</u>	<u>60.0 µL</u>
Total (40 reactions)	100 µL

Inhibition control primer mix (BVDV)

BVDV-85for primer (50µM)	20.0 µL
BVDV-390rev primer (50 µM)	20.0 µL
<u>Nuclease Free Water</u>	<u>60.0 µL</u>
Total (66 reactions)	100 µL

10 μL of template RNA was added to a master mix containing 2X SuperScript™ III reaction mix, 0.4 μM each of primers EV1F and EV3R and 0.2 μM of BVDV-85for and BVDV-390rev, 2 μL SuperScript™ III RT/ Platinum® Taq Mix, 20 units Roche Protector RNase inhibitor and nuclease free water to a final volume of 50 μL . Amplification was performed on an Applied Biosystems Verti thermocycler under the following amplification conditions: 45°C for 30 minutes then 94°C for 2 minutes followed by 35 cycles of 94°C for 15 seconds, 57°C for 45 seconds, 68°C for 60 seconds and a final extension at 68°C for 5 minutes. For reagent and equipment details, refer to appendices A1 and A2

Preparation of the round 2 EV Super-speciation snrRT-PCR 20x TaqMan probe/primer mix was as follows.

EV2 primer (50 μM)	20.0 μL
EV3 primer (50 μM)	20.0 μL
EV-AB Probe CO560 (20 μM)	25.0 μL
EV-CD Probe FAM (20 μM)	25.0 μL
<u>Nuclease Free Water</u>	<u>3.0 μL</u>
Total (80 reactions)	100 μL

Inhibition control primer/probe mix (BVDV)

BVDV-107for primer (50 μM)	18.0 μL
BVDV-246rev primer (50 μM)	18.0 μL
BVDV Probe TAMRA (20 μM)	12.5 μL
<u>Nuclease Free Water</u>	<u>1.5 μL</u>
Total (62 reactions)	50 μL

1 μL of first round product was added to a master mix containing 2X TaqMan Fast Universal PCR master mix with no UNG, 1.25 μL of EV 20x Taqman Probe/primer mix, 0.8 μL of BVDV 20x Taqman probe/primer mix and nuclease free water to a final volume of 25 μL . Amplification was performed on an Applied Biosystems 7500FAST real time thermocycler under the following FAST amplification conditions: 95°C for 10 minutes followed by 45 cycles of 95°C for 15 seconds and 60°C for 60 seconds.

2.3 Results

2.3.1 Novel Super-Speciation Semi-Nested Real-Time RT-PCR Assay

5'NTR sequences from 276 members of the genus *Enterovirus* were obtained from the GenBank via the NCBI website²⁹⁸ and aligned to examine areas of conservation across members of the genus. Once aligned, the primer locations for EV2 and EV3 within the 5'NTR used in a previously described generic PanEV snRT-PCR^{65,108} were identified and the sequence was cropped to the outer primer boundaries. After cropping the sequence alignment, a neighbour-joining tree was created using the Kimura-2-parameter method (Figure 2.1). The resultant tree displayed a clear demarcation of three clades representing *Enterovirus* species A and B, *Enterovirus* species C and D and HRV species A, B and C.

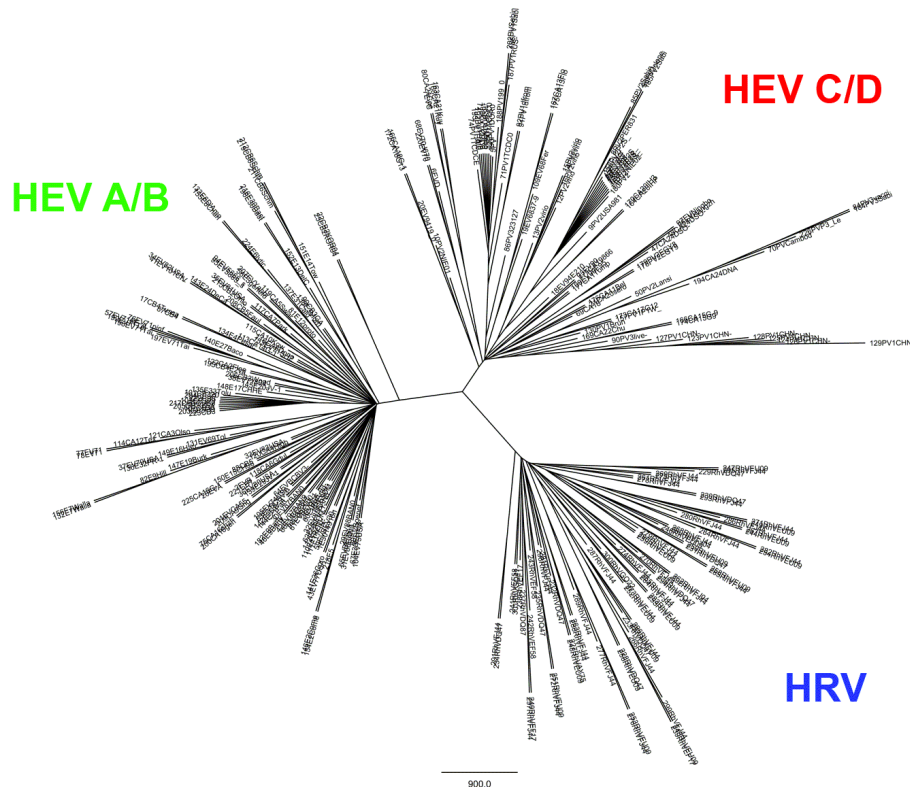


Figure 2.1 Phylogenetic analysis of members of the genus *Enterovirus* based on 5'NTR sequence.

Radial representation of 276 members of the genus *Enterovirus*, specifically EV A, B, C, D and HRV A, B, C. Partial 5'NTR fragment analysis (115bp). Clustal-W alignment, DNA distance algorithm with neighbour-joining, bootstrap analysis performed with 1000 pseudoreplicates, Phylip phylogenetic analysis package version 3.63. Scale bar indicates nucleotide substitutions per site.

The 5'NTR sequences were aligned by *Enterovirus* species, to identify sequence variation in the interior fragment of the pan-enterovirus RT-PCR amplicon. A region of interest was noted between the inner primers EV2 and EV3, upstream of the previous pan-enterovirus EV probe location (Table 2.2). This region was conserved between members of the EV A/B, C/D and HRV A/B/C species. A probe was designed with sequence complementarity to the EV A/B species and another to the EV C/D species (Figure 2.2). The probes were initially verified as specific for their respective *Enterovirus* genus species, by performing a BLAST search and examining whether the “expected value” correlated with listed members of the species for which the probe was designed. A clearly observable division in sequence similarity was present in the BLAST search results, differentiating the species A/B and C/D probes, with no apparent cross-reactivity with rhinoviruses.

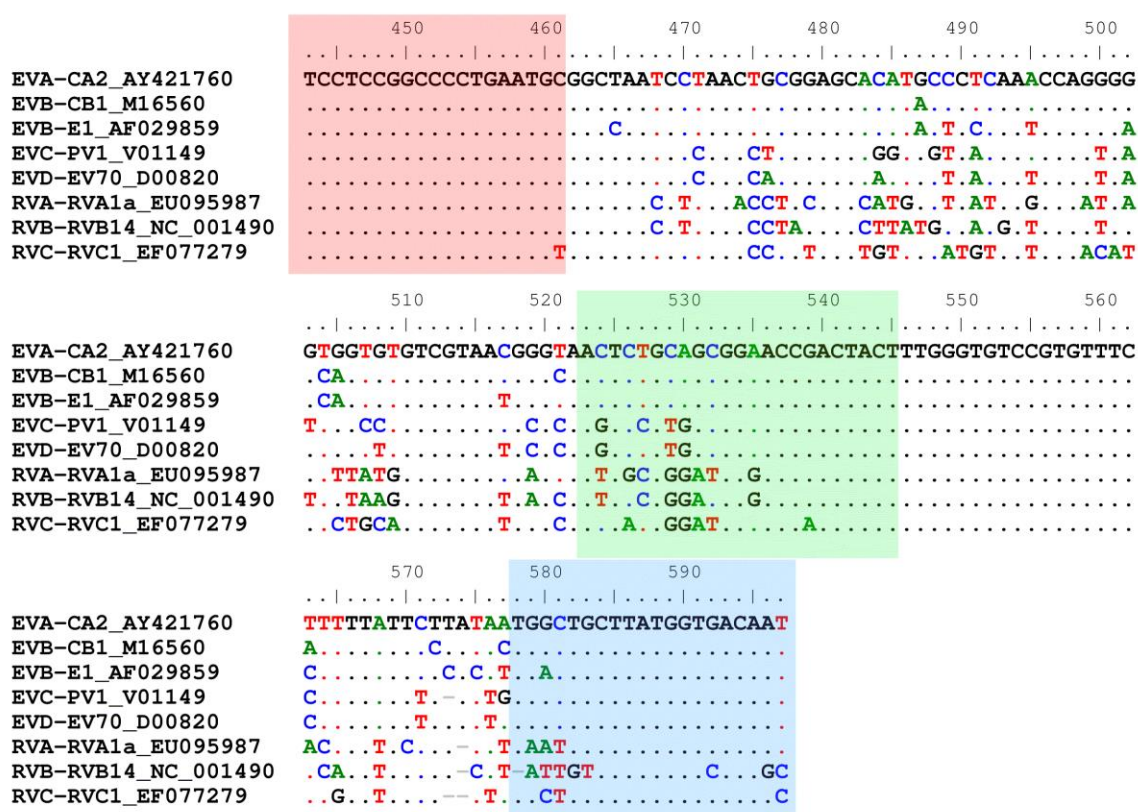


Figure 2.2 Sequence alignment indicating primer and probe binding locations. Multiple sequence alignment of representatives from each species of the *Enterovirus* genus. A fragment representing the amplicon from the newly designed EV super-speciation snRT-PCR method is shown. Pink – forward primer EV2, green – fluorescent probe location, blue – reverse primer EV3.

Two probes were synthesised with different colour fluorophores; EV C/D was paired with a 5' FAM fluorophore and a 3' BHQ1 quencher molecule whilst the EV A/B probe was synthesised with a 5' CO560 fluorophore and a BHQ1 quencher. The EV A/B probe was also synthesised using a 5' CR610 fluorophore as an alternative if problems were encountered with the channel selection and choice of quencher, such as cross-reactivity between the FAM and CO560 fluorophores.

The super speciation enterovirus A/B and C/D probes were used in replacement of the generic PanEV probe (EV probe in table 2.2). The assay was developed with the BVDV probe containing TAMRA and BHQ2 and an internal reagent control of unbound ROX dye that was provided by the manufacturer.

The EV super-speciation snRT-PCR assay was tested using known positive samples derived from a series of clinical samples with available VP1 sequence data that incorporated the antigenic BC loop region. The specificity of the assay was also tested with a series of samples that were determined as negative for enterovirus by virus culture and/or in-house PanEV snrRT-PCR used in the Enterovirus Reference Laboratory at VIDRL (Table 2.4). All developmental testing was performed in parallel with the above mentioned in-house PanEV snrRT-PCR, originally based on primers by Zoll *et al.*, (1994)¹⁰⁵. The assay had undergone rigorous examination of cross reactivity with a number of other viral pathogens, including adenovirus, HSV, VZV CMV among others^{65,105,108}. For this reason, the super-speciation assay was only tested for specificity with members of the *Enterovirus* genus.

<i>Enterovirus</i> Species	EV Type	No.	EV-A/B	EV-C/D
<i>Enterovirus A</i> (n=5)	CV-A12	1	1	0
	EV-A71	4	4	0
<i>Enterovirus B</i> (n=13)	CV-B1	1	1	0
	CV-B2	1	1	0
	CV-B4	1	1	0
	CV-B5	1	1	0
	E-6	2	2	0
	E-9	1	1	0
	E-15	1	1	0
	E-21	2	2	0
	E-25	1	1	0
	E-30	1	1	0
	E-33	1	1	0
<i>Enterovirus C</i> (n=10)	CV-A20	2	0	2
	EV-C96	5	0	5
	PV1 SL	1	0	1
	PV2 SL	1	0	1
	PV3 SL	1	0	1
EV-A/B (n=18)	5'NTR Sequencing	18	18	0
<i>Rhinovirus</i> (n=1)	Untyped	1	0	0
Negative for Enterovirus (n=24)	PCR or Culture	24	0	0
Total		71		

Table 2.4 Enterovirus types tested using EV Super-speciation snrRT-PCR. Members of the genus *Enterovirus* used in the examination of specificity of the EV super-speciation snrRT-PCR assay determined by sequence of the VP1 genomic region or 5'NTR, negative controls represented by samples with no enterovirus detected by either virus culture or in-house pan-EV snrRT-PCR.

Limits of detection for the EV super-speciation snrRT-PCR assay were determined using serial dilutions of types 1, 2 and 3 Sabin-like poliovirus and EV-A71 subgenogroup B5 control material, with dilutions from 10^{-3} to 10^{-8} depending on the virus titre (Figure 2.3). Limits of detection for the individual serotypes were: PV1SL 1.3 log CCID₅₀, PV2SL 1.0 log CCID₅₀, PV3SL 1.9 log CCID₅₀ and for EV-A71 0.9 log CCID₅₀.

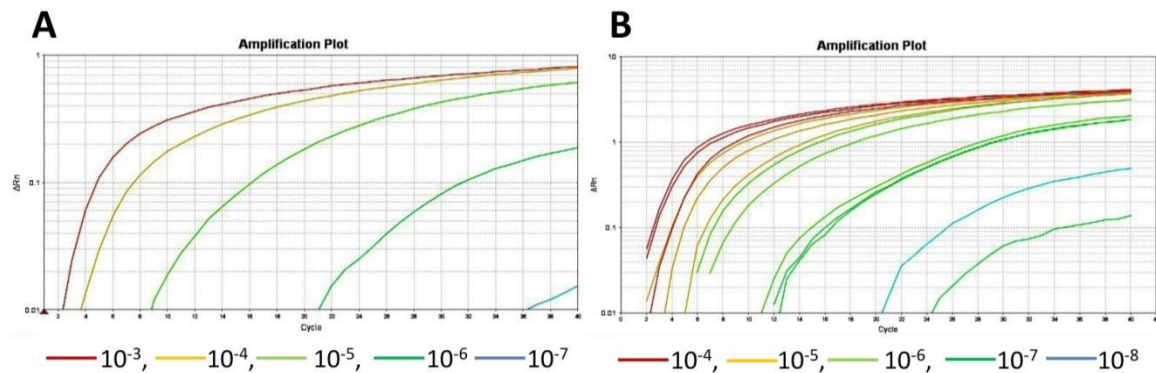


Figure 2.3 EV Super-speciation snrRT-PCR results for serial dilutions of control material. A)

Logarithmic representation of fluorescent signal increase for species A/B probe for a serial dilution of EV-A71 control material from 10^{-3} to 10^{-7} using CR610-labelled EV A/B probe B) Logarithmic representation of fluorescent signal increase for species C/D probe for a serial dilution of PV1, PV2 and PV3 Sabin-like control material from 10^{-4} to 10^{-8} FAM labelled EV-CD probe. x-axis = cycle number, y-axis = ΔRn (fluorescence)

Enterovirus specificity was examined using five members of the species A *Enteroviruses*, 13 of species B, 10 of species C and one rhinovirus. An additional 18 enteroviruses that could not be typed by amplification of the VP1 region due to low template, were examined. These viruses were determined by sequencing of the 5' NTR to be members of either the species A or B *Enteroviruses* when compared to prototype sequences and BLAST search of public sequence databases. An additional 24 clinical samples that were PCR negative using the standard in-house pan-EV snrRT-PCR method were included and none were positive in the EV super speciation snrRT-PCR assay. All known enterovirus samples tested had a direct correlation with the expected EV species-specific probe. All members of *Enterovirus* species A, B and C were correctly identified, during the initial assessment process. Only a single rhinovirus was tested for cross-reactivity. This particular sample had provided a false positive result in the standard in-house

PanEV snrRT-PCR assay, and was not detected in the new super speciation assay (table 2.4). The internal BVDV control was detected in all cases. As expected, there was evidence of BVDV signal inhibition for strongly positive enterovirus samples, indicating preferential amplification of target enterovirus sequence over that of the internal control as the amount of BVDV template added to the PCR mixture was designed to be limited.

2.3.2 Novel Super-Speciation One-Step Real-Time PCR Assay

The use of nested PCR assays is not ideal due to the extra costs in labour and reagents, but more importantly, the potential for cross contamination during the transfer of PCR product from the first to second round. A one step method was attempted using the Superscript III reagent (Life Technologies), but the limits of detection indicated the assay was unsuitable, with poliovirus type 1 at 2.3 log CCID₅₀ and EV-A71 at 1.9 log CCID₅₀ compared to 1.3 log CCID₅₀ and 0.9 log CCID₅₀, respectively, with the semi-nested version. A new one step real-time RT-PCR kit became available (AgPath-ID, Life Technologies) that proved to have highly sensitive limits of detection. To determine the limits of detection with the AgPath,-ID reagent, types 1, 2 and 3 Sabin-like poliovirus and EV-A71 subgenogroup B5 were diluted from 10⁻³ to 10⁻⁸ (Figure 2.4).

Limits of detection were as follows: PV1SL 0.3 log CCID₅₀, PV2SL 1.0 log CCID₅₀, PV3SL 0.95 log CCID₅₀ and for EV-A71 -0.1 log CCID₅₀.

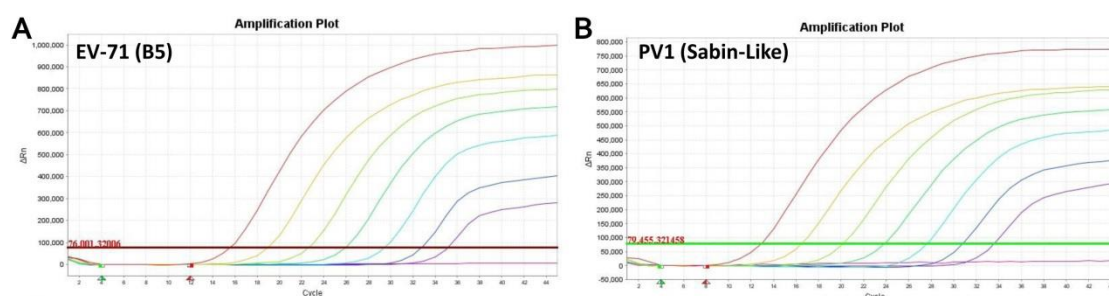


Figure 2.4 One-Step EV Super-Speciation qRT-PCR results for serial dilutions of control material.

A) CR610-labelled EV A/B and B) FAM labelled EV-CD probes. EV-A71 (B5) control material 1:10 dilutions from 6.9 to 0.9 log CCID₅₀, PV1SL control material 1:10 dilutions from 8.3 to 2.3 log CCID₅₀. x-axis = cycle number, y-axis = ΔRn (fluorescence)

Enterovirus specificity was examined using 5 members of the *Enterovirus* species A, 13 *Enterovirus* species B, 15 *Enterovirus* species C (including 6 wild type polioviruses from an Australian outbreak in the 1950s), 1 enterovirus species D and 6 members of the

species A and C rhinovirus. Among all samples tested, there was a direct correlation with *Enterovirus* species identity by sequence. All members of *Enterovirus* species A, B, C and D were correctly identified. One of six rhinoviruses that had previously tested positive with the standard in-house panEV snrRT-PCR was positive with the EV super-speciation qRT-PCR A/B probe assay, representing an 83.3% reduction in cross-reactivity (Table 2.5). An additional 15 samples that were RT-PCR negative using the standard in-house panEV snrRT-PCR method were included and none were positive with the assay.

Species	EV Super-Speciation qRT-PCR EV-A/B Positive	EV Super-Speciation qRT-PCR EV-C/D Positive
<i>Enterovirus A</i> (n=5)	5	0
<i>Enterovirus B</i> (n= 13)	12*	1 [§]
<i>Enterovirus C</i> (n=15)	0	15
<i>Enterovirus D</i> (n=1)	0	1
<i>Rhinovirus</i> (n=6)	1 [¶]	0
Negative (n=15)	0	0

Table 2.5 Enterovirus types tested using EV Super-speciation qRT-PCR. Members of the genus *Enterovirus* used in the examination of specificity of the EV super-speciation qRT-PCR assay determined by sequence of the VP1 genomic region or 5' NTR, negative controls represented by samples with no enterovirus detected by either virus culture or pan-EV snrRT-PCR. 55 samples tested: **Enterovirus Species A** - CV-A10, EV-A71 (3), EV-A120, **Enterovirus Species B** - CV-B1, CV-B5, E-6, E-7, E-11, E-18, E-19(2), E-25, E-30, E-33(2), EV-B74, **Enterovirus Species C** - EV-C96(4), EV-C99, PV1Wild(4), PV1 Sabin-like, PV2 Wild, PV2 Sabin-like, PV3 Wild(2), PV3 Sabin-like, **Enterovirus Species D** – EV-D68.

*One very low titre E-25 positive faecal sample tested negative.

[§]EV74 positive faecal sample tested positive for species A/B and C/D *Enterovirus*. Further testing of this sample using cell culture techniques indicated a mixed infection.

[¶] Rhinovirus A49 sequenced using CODEHOP primers.

An alternative version of the EV-SS qRT-PCR was examined using FAM as the fluorophore for both EV A/B and EV C/D probes, to create a generic “PanEV qRT-PCR”. Specificity of the PanEV qRT-PCR was examined using 8 members of the *Enterovirus* species A, 13 *Enterovirus* species B, 15 *Enterovirus* species C, 1 *Enterovirus* species D and 6 members of the species A and C *Rhinovirus*. An additional 28 samples that were PCR negative using the standard in-house panEV snrRT-PCR method were included and none were positive with the assay. Among all samples tested, there was a direct correlation with the *Enterovirus* species identified by VP1 sequence (Table 2.6).

Species	Pan EV qRT-PCR Positive
<i>Enterovirus A</i> (n=8)	8
<i>Enterovirus B</i> (n= 13)	12*
<i>Enterovirus C</i> (n=15)	15
<i>Enterovirus D</i> (n=1)	1
<i>Rhinovirus</i> (n=6)	1 ^φ
Negative (n=28)	0

Table 2.6 Enterovirus types tested using PanEV qRT-PCR. Members of the genus *Enterovirus* used in the examination of specificity of the EV PanEV qRT-PCR assay determined by sequence of the VP1 genomic region (or 5' NTR for rhinovirus), negative controls represented by samples with no enterovirus detected by either virus culture or pan-EV snrRT-PCR. 71 samples tested. ***Enterovirus Species A*** – CV-A5, CV-A6, CV-A10, CV-A16, EV-A71 (3), EV-A120, ***Enterovirus Species B*** - CV-A9, CV-B1, CV-B2, CV-B3, CV-B5, E-6, E-7, E-9, E-11, E-14, E-18, E-19(2), E-25, E-30, E-33(2), EV-B74, ***Enterovirus Species C*** - EV-C96(4), EV-C99, PV1Wild(4), PV1 Sabin Like, PV2 Wild, PV2 Sabin Like, PV3 Wild(2), PV3 Sabin Like, ***Enterovirus Species D*** - EV-D68.

*One very low titre E-25 positive faecal sample negative.

^φ Rhinovirus A49 sequenced using CODEHOP primers.

The limits of detection for the newly described assays were determined using the quantitated in-house quality controls, EV-A71 and Sabin-like polioviruses type 1, 2 and 3. (Table 2.7).

Assay	Log CCID₅₀ /mL	Infectious Particles/ mL
PanEV snRT-PCR (<i>Old method</i>)	0.95	8.9
EV-SS snrRT-PCR		
• EV-A71 (B5)	-0.10	0.8
• PV Sabin-like	1.00	10
EV-SS qRT-PCR		
• EV-A71 (B5)	-0.10	0.8
• PV Sabin-like	0.09	1.2
PanEV qRT-PCR		
• EV-A71 (B5)	-0.10	0.8
• PV Sabin-like	0.09	1.2

Table 2.7. Limit of Detection Results for EV RT-PCR Assays. Limits of detection expressed as log CCID₅₀ /mL and Infectious particles /mL, calculated using serial dilutions of control virus solutions. Assay abbreviations are as follows: In-house panEV snRT-PCR = PanEV snrRT-PCR, EV Super-speciation semi-nested real-time RT-PCR = EV-SS snrRT-PCR, EV Super-speciation one-step real-time RT-PCR = EV-SS qRT-PCR, PanEV one-step real-time RT-PCR = PanEV qRT-PCR.

2.4 Discussion

The ability to perform rapid testing on clinical samples and isolates is a strength of PCR technology and in particular, real-time PCR methods. The method is arguably less sensitive than cell culture, due in part to the high specificity of primer/probe sequences, which means that any significant divergence from the enterovirus target annealing sequence compared to the oligonucleotide sequence may result in false negative results, due to non-amplification of target sequence. This is not necessarily true for culture methods, with the most commonly used cell lines able to grow a broad range of enteroviruses. The limitations of cell culture include the time required to report the results, the laborious nature of the method, restrictions on rapidly scaling up the procedure and the fact that a number of *Enterovirus* species, such as coxsackievirus A1, are only cultivable in suckling mice. In addition, the cell tropism is unknown for the more than 40 new enterovirus types that were classified based on their nucleic acid sequence.

It is for these reasons that although more prone to the pitfalls of specificity, RT-PCR methods are preferred for the rapid identification of enterovirus infection in humans. The EV super-speciation qRT-PCR method described in this chapter is a sensitive assay that can be scaled up to accommodate hundreds of specimens, with the advantage of being able to differentiate *Enterovirus C/D* species from *Enterovirus A/B* species. In Australia species A/B *Enteroviruses* are more frequently encountered in clinical specimens from cases of AFP (refer to chapter 3).

The importance of discriminating between *Enterovirus* species becomes evident when considering the potential importation of poliovirus (a species C *Enterovirus*). As elaborated in the introduction, in the event of a polio importation, public health laboratories can quickly find themselves inundated with samples, not only from cases exhibiting symptoms of enterovirus infection, but also from close and secondary contacts. The ability to screen hundreds of samples and triage EV-C/D positive samples for priority characterisation by nucleic acid sequencing or cell culture is extremely useful. In Australia, *Enterovirus* species C/D were identified in 17.6% of enterovirus positive AFP cases between 2004 and 2013 (see Chapter 3). In Papua New Guinea and the Pacific Island countries, this proportion increased to 43.3% of enterovirus positive AFP cases, highlighting the benefit of an enterovirus RT-PCR screening assay that can also rapidly screen for *Enterovirus* species C. (refer to section 3.4).

For routine enterovirus surveillance and diagnostic testing, the super-speciation assay is appropriate for small diagnostic runs and will screen for any potential poliovirus that may be present in symptomatic or asymptomatic patients. This allows a public health laboratory to expedite the sample for further characterisation, with the potential to drastically reduce the time taken to notify the relevant public health authorities.

The initial observation that a fragment of the 5'NTR sequence of the *Enterovirus* genus could be grouped into three major phylogenetic clusters-EV-A/B, EV-C/D and rhinovirus, led to the development of a highly sensitive single round pan-EV RT-PCR for *Enterovirus* species A/B and C/D. The super speciation pan-EV RT-PCR assay has a direct application in diagnostic virology laboratories and in particular those involved with the polio eradication program, who may need to rapidly and accurately screen clinical specimens for poliovirus. It was anticipated that the super speciation assay would also remove the inherent cross-reactivity with rhinoviruses that is a feature of many pan-EV assays. The recent reporting of rhinoviruses detected in stool specimens highlights the need to reduce false enterovirus positives associated with AFP cases^{43,299}. While the super speciation assay still detected one of six rhinoviruses known to cross-react with the standard pan-EV RT-PCR, this represented an 83.3% increase in specificity. This particular rhinovirus sample is worth investigating further, as a fragment of the VP1 genomic region was also amplified by the CODEHOP assay used to type enteroviruses. Further assay development may be able to remove cross-reactivity with rhinoviruses altogether, or it could also be indicative of the continuing evolution of the *Enterovirus* genus

The flexibility of the probe configurations also allows for both the A/B and C/D probes to be labelled with FAM as a true pan-EV qRT-PCR, in replacement of the speciation method using different colour fluorophores. This generic pan-EV detection method would have high specificity for all species A/B and C/D *Enteroviruses*, as opposed to the various species of *Rhinovirus*.

Having used bioinformatics methods to improve the detection and initial characterisation of enteroviruses in clinical samples, the next stage is to take positive samples detected using RT-PCR methods and determine if any specific serotype is associated with serious diseases such as AFP and may therefore represent a potential public health threat. This process is described in detail, in the next chapter.

Chapter Three

Characterisation of Enteroviruses Associated with Acute Flaccid Paralysis

3.1 Manuscripts published in association with this chapter:

- **Roberts, J.A.**, Bowden, D.S., Thorley, B. R., Revill, P.A. Chapter: *New Technologies for Viral Diagnosis and Detection, Foodborne Viral Pathogens*, Hansman, G., White, P., (Eds.), CRC Press, 1st Ed. **ISBN-10:** 1466579501. [*In Press*]
- **Roberts, J. A.**, Hobday, L. K., Ibrahim, A., Aitken, T., Thorley, B. R., 2013. Annual Report of the Australian National Enterovirus Reference Laboratory 2012, *Communicable Diseases Intelligence*, 37 (2). [Refer to appendix A8.2.1].
- **Roberts, J. A.**, Hobday, L. K., Ibrahim, A., Aitken, T., Thorley, B. R., 2013. Annual Report of the Australian National Enterovirus Reference Laboratory 2010-2011, *Communicable Diseases Intelligence*, 37 (2). [Refer to appendix A8.2.2].
- **Roberts, J.**, A. Hung, and B. Thorley. Application of Bayesian methods to the inference of phylogeny for enterovirus surveillance. 2013. *Euro surveillance*: 18(9). [Refer to appendix A8.2.3].
- Kesson, A. M., Ming, C. C., Troedson, C., Thorley, B. R., **Roberts, J. A.**, Echovirus 19 Associated with a Case of Acute Flaccid Paralysis. 2013. *Journal of Paediatrics and Child Health*, 49(3): E239–E242. [Refer to appendix A8.2.5].

3.2 Introduction

Since the widespread adoption of PCR detection methodologies for viral diagnostic testing and research, the ability to obtain serotype specific sequence data has permitted the application of phylogenetic analysis techniques to enteroviruses of public health interest^{27,112,115,116}. These data are particularly useful in the investigation of outbreaks and in the determination of enterovirus circulation on a global and national scale, as well as individual case-specific investigations¹¹⁶.

Not all cases of AFP are due to poliovirus infection, with a number of enteroviruses such as CV-A7, Echo-19 and EV-A71 showing evidence of an association with the syndrome^{42,133,134,300}. The identification of non-polio enteroviruses is also important, not only to exclude the presence of poliovirus, but also to delineate sporadic cases from those associated with epidemics^{25,29,143,164}. Factors involved in determining whether or not a virus warrants further investigation relate to the clinical presentation of the case being

investigated, the number of cases from which the virus has been identified, or the emergence of previously undescribed viruses detected from clinical samples. Analysis of retrospective sequence data, combined with contemporary sequence results, can be informative in determining virus transmission within Australia and the Western Pacific region, to assist public health authorities in identifying potential threats of communicable disease.

Between 1993 and 2001, the NERL characterised enterovirus isolates from AFP cases by antibody neutralisation methods, using pooled and monotype specific antisera. From 2001, enterovirus isolates were characterised by sequencing a fragment of the VP1 genomic region, using single round PCR primers directed to amplify either the 5' or 3' aspects of the protein encoding sequence^{112,114}. From 2009 onwards, the CODEHOP semi-nested RT-PCR assay was introduced, with the advantage of being specifically developed to detect enteroviruses directly from clinical samples as a semi-nested assay, as opposed to single round amplification requiring a higher titre of virus derived from cell culture isolates¹¹⁷. The CODEHOP PCR amplification method targets the hyper-variable region encoding the major antigenic determinant (VP1 BC loop) as shown in Figures 1.7 and 1.14 of Chapter 1, section 1.2.5.

This chapter describes the use of bioinformatics methods in a retrospective, systematic analysis of data derived from a regional surveillance system, aimed at the investigation of AFP in children under the age of 15. Given that AFP can indicate one of the most serious presentations of enterovirus infection, it is logical to examine these samples for the presence of enterovirus serotypes that may be associated with increased virulence and any associated morbidity, mortality and enhanced transmission. The identification of viruses displaying these traits is of interest to the public health community and would therefore represent strong candidates for further research including methods that examine virus structure, such as the *in-silico* reconstruction and simulation methods developed as part of this thesis and described in Chapter 5.

3.3 Materials and Methods

3.3.1 Partial VP1 Sequencing of Enterovirus Positive Materials

Enteroviruses for sequencing were derived either from culture supernatant or direct testing of clinical samples. RNA was extracted as described in section 2.2.3. The CODEHOP assay was developed by CDC USA and was performed according to the published protocol ¹¹⁷. The following PCR primers and cycling conditions were used to generate approximately a 350 base pair fragment of the VP1 region (Table 3.1):

	Primer	Sequence 5' to 3'	Position (PV-1 J02281)
1	AN32	GTYTGCCA	3002-3009
2	AN33	GAYTGCCA	3002-3009
3	AN34	CCRTCRTA	3104-3111
4	AN35	RCTYTGCCA	3002-3009
5	SO224for	GCIATGYTIGGIACICAYRT	1997-1996
6	SO222rev	CICIGGIGGIAYRWACAT	2951-2969
7	AN89for	CCAGCACTGACAGCAGYNGARAYNGG	2602-2627
8	AN88rev	TACTGGACCACCTGGNGGNAYRWACAT	2951-2977
9	AN232for	CCAGCACTGACAGCA	2602-2616
10	AN233rev	TACTGGACCACCTGG	2963-2977

Table 3.1. Primers used in the “CODEHOP” assay. Primer locations based on the prototype sequence for Poliovirus type 1 Mahoney, GenBank accession number J02281.

5.0 µL of extracted RNA template was added to a PCR amplification tube containing 5.0 µL of cDNA mastermix:

5xFS Buffer	2.0 µL
10mM dNTP (40mM)	0.25 µL
Nuclease free water	0.25 µL
AN32, AN33, AN34, AN35 (10uM)	0.5 µL
DTT 0.1M (Invitrogen #Y00147)	1.0 µL
RNase inhibitor (Roche #03335402001)	0.5 µL
<u>SuperScript III RT (Invitrogen #56575)</u>	<u>0.5 µL</u>
Total	5.0 µL

The tubes were placed in a thermocycler and incubated for five minutes at 25°C, followed by one hour at 50°C to reverse transcribe single stranded RNA. First round PCR

master mix was prepared as follows and 40 μL added to the each tube containing reverse transcription master mix and cDNA.

Nuclease free water	23.5 μL
10x Taq Buffer (Roche #1127138001)	5.0 μL
(#61) SO224For (10 μM)	5.0 μL
(#40) SO222Rev (10 μM)	5.0 μL
10mM dNTP (40mM) (Promega #24975812)	1.0 μL
<u>Taq Polymerase (Roche #11418432001)</u>	<u>0.5 μL</u>
Total	40 μL

First round PCR conditions consisted of 40 cycles of denaturation at 95°C for 30 seconds, followed by 42°C for 30 seconds annealing and a 45 second extension step at 60°C. Ramping rates for the transition from annealing step to extension step was programmed to occur at a rate of 1.5°C/sec. At the completion of cycling, samples were held indefinitely at a 4°C. Subsequently, 1 μL of first-round amplified product was added to the following master mix:

2x FastStart Mix (Roche #04710444001)	25 μL
Nuclease free water	16 μL
(#136) AN89For (10 μM)	4.0 μL
<u>(#135) AN88Rev (10 μM)</u>	<u>4.0 μL</u>
Total	49 μL

Second round PCR thermal cycler conditions consisted of a six minute DNA polymerase enzyme activation step at 95°C, followed by 40 cycles of denaturation at 95°C for 30 seconds, annealing at 60°C for 20 seconds and a 15 second extension step at 72°C. At completion, the PCR products were held indefinitely at 4°C.

The amplified product size and concentration was determined using an Agilent Bioanalyser microfluidic system that required 1 μL of PCR product or control per well of a Bioanalyser 7500 DNA microfluidic chip. The products were electrophoresed using a fluorescence-based microfluidic system, which allowed for the accurate determination of band sizes and concentrations. The Roche PCR product purification kit was used and 20ng of double-stranded DNA was added to a sequencing reaction consisting of:

Applied Biosystems BDT v3.1	2.0 µL
5x Buffer	3.0 µL
Primer	1.0 µL
<u>Template</u>	<u>20 ng</u>
NFW for a final volume of	20 µL

Sequencing reactions were assigned a unique number and forwarded to “Molecular Diagnostics” (formerly Applied Genetic Diagnostics) at the Centre for Translational Pathology, The University of Melbourne. The sequencing facility was certified to ISO certification standard by the Australian National Association of Testing Authorities (NATA). Sequencing results were received in electronic format via email before undergoing further analysis.

3.3.2 Determination of Enterovirus Type.

Reconstruction of contiguous sequence data was performed using the ChromasPro software package (Technylesium) and analysed using Bioedit version 7.0.1. A reference sequence database consisting of all enterovirus prototypes was constructed using data available from GenBank. Relevant sequences were downloaded via the National Centre for Biotechnology Information (NCBI) website²⁹⁸.

Initial analysis involved a nucleic acid Basic Local Alignment Search Tool (BLAST) using the NCBI website³⁰¹, with the top 10 hits used as an indicator of the relevant *Enterovirus* species, A to D. Subsequent analysis involved the alignment of the subject sequence with all prototypic members of the relevant species. Once aligned, percentage identity at the nucleotide level and amino acid level were calculated and an initial phylogenetic tree constructed using distance methods for nucleic acids and maximum likelihood methods for amino acid sequences.

Enterovirus type was determined using established parameters for identification of a homologous virus when compared with prototype sequence (refer to Chapter 1, section 1.2.5). Specifically, the criteria required for a homologous enterovirus typing required a sequence identity greater than 75% for nucleic acid and 88% for amino acid sequence,

combined with the sharing of a common ancestor or node for each phylogenetic tree generated for both nucleic acid and amino acid sequence. In circumstances where the virus sequence did not fully meet these criteria, complete VP1 sequences were amplified and the sequence reassessed against prototype sequences. In instances when the full VP1 sequence still did not meet the criteria, sequences were forwarded initially to Polio and Picornavirus Laboratory Branch of the Centres for Disease Control and Prevention in the USA, which acts as a sequence repository for enterovirus identification on behalf of the ICTV Picornaviridae Study Group³⁰². Sequence data derived from a putative novel enterovirus sequence was then forwarded to Dr Nick J. Knowles, Executive Committee member and Chair of the Picornaviridae Study Group of the ICTV.

3.3.3 Sequencing of the EV-A71 Capsid Region

Enteroviruses identified as EV-A71 were characterised by sequencing the VP1 region. EV-A71 viruses, identified within genogroup B5, were amplified using primers designed in-house (Table 3.2). For genogroup C viruses, a fragment of VP1 was amplified using published methods¹⁶⁵ (primer and probe sequences are listed in Table 3.2). RNA was extracted as described in section 2.2.3 and the sequencing assays performed as per the manufacturer's recommendations.

	Primer	Sequence 5' to 3'	Position
1	EV71-VP1-634F [§]	GGAGAACACAARCARGAGAAAGA	634-657 ^a
2	EV71-VP1-743R [§]	ACTAAAGGGTACTTGGACTTVGA	712-734 ^a
3	EV71-VP1-TaqMan [§]	FAM-TGATGGGCACGTTCTCAGTGCG-BHQ1	2602-2627 ^a
4	EV71-VP1-3F [§]	AGAYAGGGTGGCRGATGT	3-20 ^a
5	EV71-VP1-703R [§]	CTGAGAACGTGCCCATCA	686-703 ^a
6	EV71B5-F1 [¥]	TCACCGGGTCCTTCATGGC	2042-2060 ^b
7	EV71B5-F2 [¥]	TCATCTGTCACCCTTGTGAT	2167-2186 ^b
8	EV71B5-R1 [¥]	AGCTGACTGGATAGTGCTTTCT	3502-3523 ^b
9	EV71B5-R2 [¥]	TCATTATGAGTAGCGAGGTG	3358-3377 ^b

Table 3.2. EV-A71 amplification and sequencing primers. Primers used for the detection and sequencing of EV-A71. Primer locations based on sequences for Enterovirus A71, ^aGenBank accession number AM490160, ^bGenBank accession number FJ461781.

[§] Denotes primers derived from Khanh, TH *et al.*, (2012)¹⁶⁵

[¥] Denotes primers designed "in-house".

For the EV-71 genogroup C real-time RT-PCR method, the following conditions were used. 5 µL of extracted RNA template was added to the following master-mix.

2x AgPath MasterMix Buffer	12.50 µL
EV71-VP1-TaqMan probe/primer mix (0.12uM/0.4uM) ¹⁶⁵	0.60 µL
NFW	0.60 µL
25x RT-PCR Enzyme Mix	1.00 µL
RNAsin	0.25 µL
<u>Template</u>	<u>10.00 µL</u>
Total	25.0 µL

The thermal cycler profile was 45 minutes at 45°C, followed by ten minutes at 95°C then, 45 cycles of denaturation at 95°C for 15 seconds, followed by an extension step of 45 seconds at 60°C.

Sequencing reactions for the EV-A71 genogroup C viruses consisted of a single round one-step RT-PCR. Master mix was prepared as follows with 10 µL of RNA extract.

2x SSIII Reaction buffer	25.0 µL
Nuclease Free Water	10.5 µL
EV71-VP1-3F (10uM)	1.0 µL
EV71-VP1-743R (10uM)	1.0 µL
SS III Enzyme	2.0 µL
RNAsin	0.5 µL
<u>Template</u>	<u>10 µL</u>
Total	50.0 µL

The initial incubation consisted of 30 minutes at 50°C, followed by a polymerase activation step of two minutes at 94°C. The PCR amplification consisted of 40 cycles of denaturation at 94°C for 15 seconds, followed by 50°C for 30 seconds annealing and a 1 minute extension step at 68°C. An amplicon “polishing” step was performed at 68°C for 5 minutes, after which samples were held at 4°C.

Viruses identified as EV-A71 genogroup B, by either the CODEHOP method or partial sequencing of the 5'NTR, were amplified using the following methods. Sequencing reactions for the EV-A71 genogroup B viruses consisted of a nested RT-PCR. A master mix was prepared as follows, to which 10 µL of RNA extract was added.

2x SSIII Reaction buffer	25.0 µL
Nuclease Free Water	7.5 µL
EV71B5-F1 (10uM)	2.5 µL
EV71B5-R1 (10uM)	2.5 µL
SS III Enzyme	2.0 µL
<u>RNAasin</u>	<u>0.5 µL</u>
Template	10 µL

The reverse transcription step consisted of 30 minutes at 42°C, followed by a polymerase activation step of 3 minutes at 94°C. PCR amplification consisted of 35 cycles of denaturation at 94°C for 30 seconds, followed by 50°C for 45 seconds annealing and a 1 minute extension step at 65°C. At completion the samples were held at 4°C.

2 µL of the first-round amplified product was added to the second round master mix:

2x Roche FastStart	25.0 µL
Nuclease Free Water	21.0 µL
EV71B5-F2 (10uM)	1.0 µL
EV71B5-R2 (10uM)	1.0 µL
<u>Template</u>	<u>2 µL</u>
Total	50.0 µL

The second round PCR amplification conditions were: 2 minutes of enzyme activation at 94°C, followed by 35 cycles of denaturation at 94°C for 30 seconds, 50°C for 30 seconds annealing and a 1:30 minute extension step at 72°C. An amplicon “polishing” step was performed at 72°C for 5 minutes and samples were held at 4°C.

The Agilent Bioanalyser microfluidic system was used to determine the PCR amplified product size and concentration. Products were purified and sequenced as described in section 3.3.1.

3.3.4 Sequencing of Newly Described and Novel Virus Capsid Region.

In order to increase the sensitivity of the methods being used to obtain clean amplicons for sequencing a putative novel enterovirus, a semi-nested method was employed. Forward and reverse primers (Novel_EVA-VP1for, Novel_EVA-VP1rev, listed as primers 3 and 4 in Table 3.3), were designed from the sequence derived from a fragment of VP1 using the CODEHOP fragment primers, which indicated the virus to be a member of species A. An initial amplification was performed using a single round RT-PCR, with primers (PanEV1_OS, No. 1 Table 3.3) and (Novel_EVA_2Crev, No. 5 Table 3.3). 2 µL of the RT PCR assay was added to 2 separate PCR reactions, with the first involving the forward primer (EV2_IS, No. 4 Table 3.3) from the 5' NTR and the reverse primer (Novel_EVA-VP1rev, No. 3 Table 3.3) designed from the CODEHOP sequence data. The second reaction involved using the forward primer (Novel_EVA-VP1for, No. 4 Table 3.3) derived from the CODEHOP sequence data, with the reverse primer (Novel_EVA_2Crev, No. 5 Table 3.3) previously designed from the 2C NTPase consensus data, as highlighted in the multiple sequence alignment in Figure 3.1. This semi-nested RT PCR approach resulted in two fragments 1.8kb and 2.0kb in length that overlapped with the original fragment derived by the CODEHOP method.

	Primer	Sequence 5' to 3'	Position (CV-A16 U05876)
1	PanEV1_OS	CAAGCACTTCTGTTTCCCGG	164-184 (A)
2	PanEV2_IS	TCCTCCGGCCCCTGAATGCG	446-465 (B)
3	Novel_EVA_VP1rev	TCCCCATTTGGAGTAGTAGC	2863-2882 (C)
4	Novel_EVA_VP1for	GGGCTTGTGGGAGTTATAGA	2707-2726 (C)
5	Novel_EVA_2Crev	GTGTTTGCTCTTGAACCTGCATGTA	4414-4437 (D)
6	EV-A120_4235for	AAATTATACCAGCAGCCAAAGA	4235-4256
7	EV-A120_3540rev	GTGCTTCCTTTTAGAATTGCAG	3519-3540
8	EV-A120_3450for	AGTTTCCTCTACCACTGCTCAA	3450-3471
9	EV-A120_2040rev	GTAATAACGGCACAATTGACCT	2019-2040
10	EV-A120_1930for	ATACCGTTAAAGTGCAAACAT	1930-1951
11	EV-A120_1360rev	ATACAGTGCCAGCACATATTC	1339-1360
12	EV-A120_1265for	ATTTTGCATCCACGTTCACT	1265-1284

Table 3.3. Primers used in sequencing of the novel virus EV-A120. Primer locations based on the prototype sequence for Coxsackievirus A16 GenBank accession number U05876, (coloured positions refer to the colouring depicted in Figure 3.1).

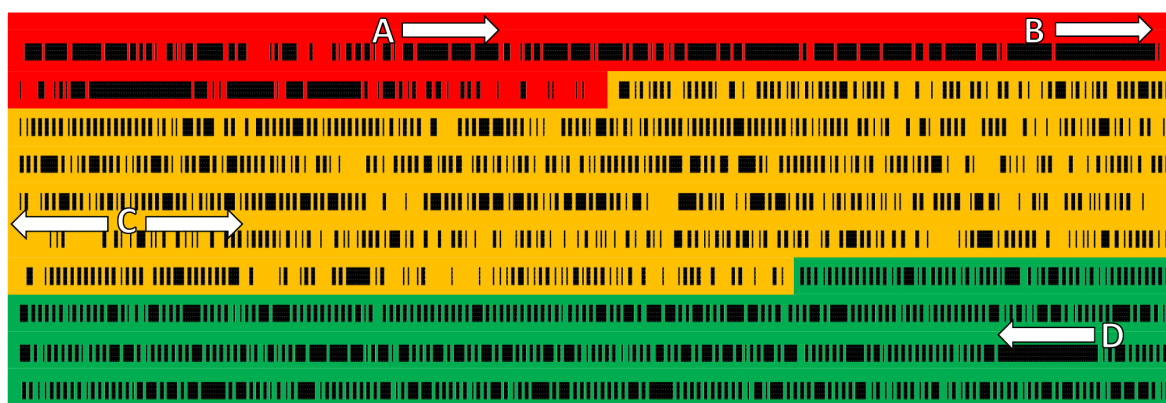


Figure 3.1 Schematic diagram of aligned enterovirus genomes of four prototype species A
Enteroviruses CV-A7, CV-A14, CV-A16 and EV-A71, closely related to EV-A120. Black bars indicate areas of 100% identity across all four prototypes. A) forward primer EV1, B) forward primer EV2, C) forward and reverse primers derived from the CODEHOP fragment and D) consensus reverse primer from the 2C NTPase region. Colour scale: Red – 5′NTR, Orange – P1 (capsid encoding region), Green – P2 (non-structural protein encoding region).

The primers used for PCR amplification were also employed to sequence the respective fragments. Subsequent primers were designed using a primer walking methodology as sequence data were progressively obtained. The final contiguous fragment was 3.7k base pairs in length and incorporated a fragment of the 5′NTR, the entire P1 capsid encoding region and part of the P2 non-structural region including the entire 2A and 2B regions and partial sequence of the 2C region. The location of the primers used for PCR amplification and sequence analysis of the novel enterovirus are depicted in Figure 3.1 and are listed in table 3.3 (see also Appendix A4). Once these methods proved successful, the same process was applied to the design of primers for the amplification of sequence representing the capsid encoding region of EV-C96 viruses (refer to appendix A4).

3.3.5 Phylogenetic Analysis of Enteroviruses

Phylogenetic analysis of enteroviruses was performed using the MEGA 5.2 software package⁵⁰. Sequence alignments were created using the query sequence combined with prototypes representing historical and relevant contemporary sequences for comparison. The respective sequences were either derived from the GenBank repository using the National Centre for Biotechnology Information (NCBI) website interface²⁹⁸, or from members of the Enterovirus Reference Laboratory Network of Australia, or from sequences contained within the database held by Enterovirus Reference Laboratory of Australia.

Raw sequence data were trimmed at the nucleic acid level, in order to ensure that sequences were in-frame before analysis. Sequences were aligned using the MUSCLE algorithm from within the MEGA 5.2 software interface. Model testing was performed for each sequence alignment and the appropriate algorithm was selected on the basis of the lowest log odds score combined with the smallest number of parameters required⁵⁰.

Once an appropriate algorithmic model was determined, phylogenetic inference was performed using the selected model combined with 1000 bootstrap pseudoreplicates. In the case of Bayesian analysis, 10,000,000 states were initially performed using the BEAST software package combined with the BEAGLE GPU extension. Effective sampling size (ESS) was determined for all parameters and in instances where the ESS was below 300, the analysis was increased to 30,000,000 states. Trees were examined using FigTree version 1.3.1 and for Bayesian phylogeography, the program SPREAD³⁰³ was used to map the global distribution in a format compatible with Google earth .kml file extension¹²⁹.

The ancestry of a virus was determined by using phylogenetic methods. Related viruses were selected from prototype sequence data for comparative analysis, in order to determine if any recombination events had occurred. This was done using the program Simplot³⁰⁴, which allowed for a visual representation of sequence identity at the nucleic acid level for the length of each sequence, termed a Simplot³⁰⁵. Subsequent analysis of multiple aligned sequences using this program allowed the inference of ancestral relationships to multiregional aspects of the sequence data, identifying regions of homology between virus species. In this particular instance, complete capsid sequence data representing the query sequence (EV-A120) and all available members of the species A *Enteroviruses* that affect humans were analysed using a 200 base pair window, which was then frame shifted by 20 base pairs, with each window being subjected to analysis using the Kimura-2-parameter model with 100 bootstrap pseudoreplicates.

3.4 Results

3.4.1 Typing of Enteroviruses Associated with Acute Flaccid Paralysis

As part of ongoing AFP surveillance in the WHO Western Pacific region, analysis of 594 acute flaccid paralysis cases from Australia, Papua New Guinea (PNG) and the Pacific Island countries (PIC) was undertaken for the period 2004 to 2013. Prior to 2009, enterovirus positive samples were detected only by cell culture methods, with characterisation performed using sequencing methods targeted towards a fragment of the VP1 genomic region. In 2009, a semi-nested pan EV RT-PCR system was introduced and samples were tested in parallel with cell culture methods (refer to section 2.3.1).

Of the 594 acute flaccid paralysis cases investigated, 109 tested positive for enterovirus by RT-PCR (table 3.4), this represents an enterovirus identification rate of 18% from AFP cases, a figure that is within the expected range of 5-25%, defined by WHO ³⁰⁶. Of all cases of AFP examined from Australia, PNG and the PIC, the dominant virus identified was EV-A71, with 10 cases representing 9.2% of all enterovirus positive cases and 1 in 59 AFP cases (1.68%). This virus was only identified in Australia. This was followed by EV-C96 with 8 cases, representing 7.3% of enterovirus positive cases and 1 in 74 AFP cases (1.35%), detected in the PIC and PNG, but not Australia. In Australia, 11.1% of the 342 AFP cases investigated yielded an enterovirus, while in the PIC, enteroviruses were identified from 24.5 % of 102 AFP cases and 30.6% of 150 cases were positive for enterovirus in PNG (Table 3.4).

In terms of Australian AFP cases investigated during this period, EV-A71 was by far the most predominant virus identified representing 10 of 38 (26.3%) enterovirus positive cases and 10 of 342 AFP cases (2.92%). In the PIC, CV-A24 was the most commonly identified enterovirus at 4 of 25 (16.0%) enterovirus positive cases and 4 of 102 AFP cases (3.92%). In PNG the most commonly identified enterovirus associated with AFP was EV-C96 representing 6 of 46 (13.0)% of enterovirus positive cases or 6 in 150 (4.0%) total AFP cases, (Table 3.5). In addition, three Rhinoviruses were identified in samples received from AFP cases.

Enterovirus Type	Australia	Enterovirus Type	PIC	Enterovirus Type	PNG
<i>EV Negative</i>	304	<i>EV Negative</i>	77	<i>EV Negative</i>	104
EV-A71	10	CV-A24	4	EV-C96	5
CV-A4	3	E-6	3	EV-C96+E-15	1
CV-B5	3	EV-C96	2	CV-A13	4
PV3-SL	2	CV-A13	2	CV-A24	3
E-11	2	CV-B3	2	CV-A2	2
E-30	2	E-1	2	CV-A20	2
Rhinovirus	2	EV-B88	2	E-11	2
CV-A10	1	E-11	1	E-13	2
CV-A16	1	E-14	1	E-21	2
CV-A24	1	E-25	1	EV-C99	2
CV-A7	1	E-7	1	EV-A120	1
CV-B2	1	EV-D68	1	CV-A1	1
CV-B3	1	EV-B74	1	CV-A12	1
E-18	1	EV-B75	1	CV-A4	1
E-19	1	EV-C99	1	CV-A5	1
E-25	1			CV-A6	1
E-3	1			CV-B1	1
E-7	1			CV-B2	1
EV-D68	1			E-1	1
PV1-SL+PV2-SL	1			E-13+EV-B74	1
PV1-wild*	1			E-17	1
				E-19	1
				E-25+E11	1
				E-27	1
				E-6	1
				EV-B74	1
				EV-B76	1
				EV-B97	1
				PV3-SL	1
				PV3-SL+E-3	1
				Rhinovirus	1
Total Cases	342	Total Cases	102	Total Cases	150
EV Detected	38 (11.1%)	EV Detected	25 (24.5%)	EV Detected	46 (30.6%)

Table 3.4: Breakdown of enteroviruses identified by country, a total of 594 AFP cases were investigated during the period 2004 to 2013 with a total of 109 positive enteroviruses (18.4%) identified from the Western Pacific Region. CV-A= Coxsackievirus A, CV-B= Coxsackievirus B, E= Echovirus, EV= Enterovirus, PV= Poliovirus, SL= Sabin-like. *2007 PV1-wild case detected in a person over the age of 15 years (as described in section 2.1)^{137,161,295}. Australian AFP cases and related reporting activity are available via the Commonwealth Government's, Department of Health publication: Communicable Diseases Intelligence³⁰⁷⁻³¹⁴.

EV Type	EV Species	Total (Percent of 109 EV'S)
<i>EV-A71</i>	A	10 (9.2%)
<i>EV-C96</i>	C	8 (7.3%)
<i>CV-A24</i>	C	8 (7.3%)
<i>CV-A13</i>	C	6 (5.5%)
<i>E-11</i>	B	6 (5.5%)
Total		38 (34.9%)

Table 3.5 Top 5 enteroviruses identified from cases of acute flaccid paralysis. The five most frequently identified enteroviruses in cases of AFP for Australia, PNG and the PIC during the period 2004 to 2013 are listed in order of frequency. These five enterovirus serotypes represent 34.5% of all enteroviruses identified from the AFP cases examined.

In Australia the most frequently identified *Enterovirus* species was A, constituting 16 of 38 (42.1%) of all positive cases (Table 3.6). In the PIC, species B *Enteroviruses* were dominant with 15 of 25 (60.0%) enterovirus positive cases, while in PNG species C *Enteroviruses* predominated at 20 of 46 (43.5%) enterovirus positive cases (Table 3.6).

Species	Australia	PIC	PNG	Total
A	16		8	24
B	14	15	17	46
C	5	9	20	34
D	1	1		2
RhV	2		1	3
Total	38	25	46	109

Table 3.6 Enterovirus species identified in Australia, Pacific Island countries and Papua New Guinea. Breakdown by species of 109 enterovirus positive samples from 594 cases investigated during the period 2004 to 2013.

The recent predominance of species A *Enteroviruses* in Australia included a significant EV-A71 outbreak in South-Eastern Australia in 2013, with sporadic identifications of EV-A71 prior to this event (Figure 3.2). Data derived from the ERLNA was compiled, enabling an observation that during the increase in EV-A71 cases, a concurrent increase in HFMD cases due to CV-A6 occurred in conjunction with an increase in meningitis cases due to E6. Between 2004 and 2011, no EV-A71 was identified from an AFP case in Australia. No evidence of AFP cases associated with enterovirus outbreaks was observed in the PIC or PNG from 2004 to 2013.

A non-cultivable enterovirus was identified from an AFP case in PNG in 2009 using pan enterovirus PCR methods and was determined to be a novel enterovirus. Subsequent

investigations in 2013 identified this virus as a novel member of the species A *Enteroviruses*, and was assigned as EV-A120 by the international committee on taxonomy of viruses in that same year (refer to section 3.4.4).

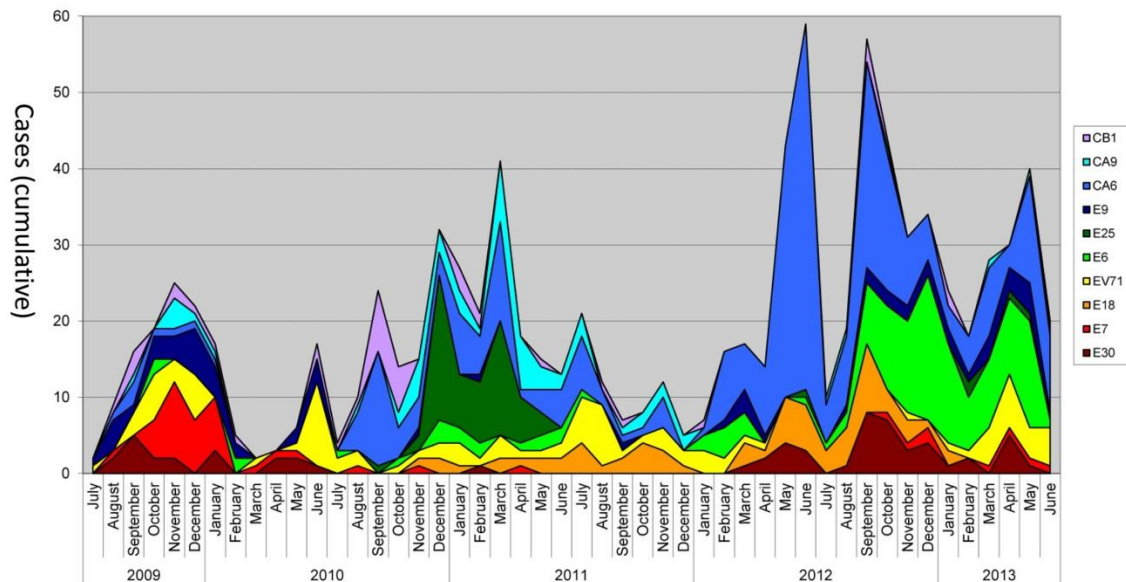


Figure 3.2, cumulative graph of the 10 dominant circulating enterovirus types Australia. Cumulative graph of major circulating enterovirus types during the period surveillance from 2009 to 2013 for Australia, data derived from multiple sources nationally as part of the Enterovirus Reference Laboratory Network.

3.4.2 Characterisation of EV-A71 in Australia

The initial detection in 2011 of EV-A71 subgenogroup C4 in Western Australia, prompted an initial investigation into the molecular epidemiology of a previous Australian EV-A71 outbreak in 1999 to 2001. Retrospective sequence data from 1997 to 2003 were obtained from GenBank and used to create phylogenetic trees for the genogroup B and genogroup C EV-A71 viruses separately. Using Bayesian phylogeographic methods, two trees incorporating sequence data with the temporal and geographic distribution of EV-A71 viruses were created, (Figures 3.3, 3.4 and 3.5). The results indicated that viruses obtained from the initial outbreak in Perth, Western Australia in 1999 shared a common ancestor with viruses circulating in Sarawak Malaysia and Singapore. EV-A71 sequence data from the second stage of the Australian outbreak, which occurred in 2000 to 2001 in Sydney, New South Wales on the eastern coast of Australia, indicated a common ancestor was shared with viruses identified in peninsular Malaysia, (Figures 3.3, 3.4).

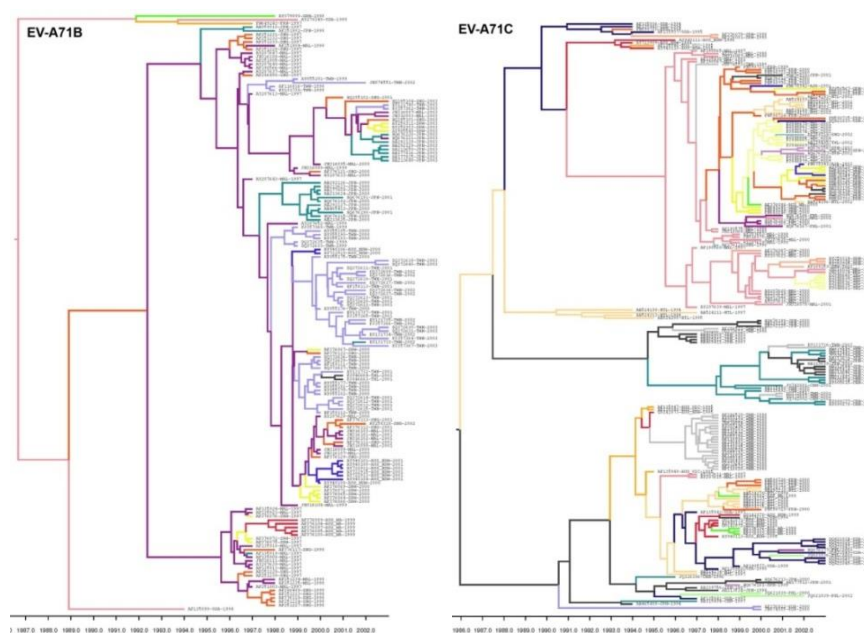


Figure 3.3 Phylogeographic analysis of enterovirus A71 using Bayesian methods 1997-2003*. Analysis of genogroup B and genogroup C EV-A71 sequences during a genogroups shift in Australia from 1997 to 2003. Coloured branches indicate country of origin. x-axis is time in years. *This graphical representation highlights a problem associated with the presentation of phylogenetic data in tree form: data-rich trees can be difficult to read. For a more intuitive and visually simpler version of these data, refer to figure 3.4 below.

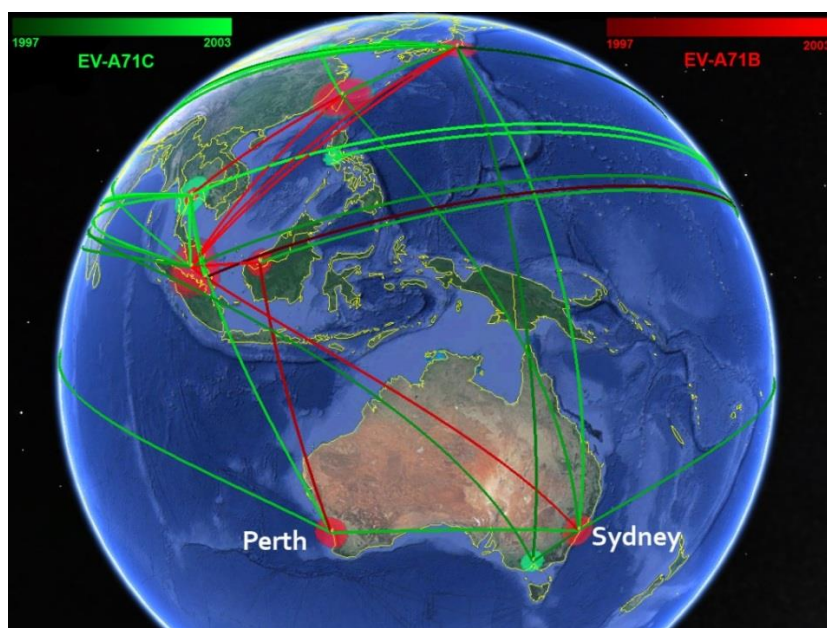


Figure 3.4 Phylogeographic analysis of enterovirus A71 using Bayesian methods overlaid on a global map. Google Earth representation of EV-A71 transmission during the period 1997 to 2003. The image indicates virus transmission patterns over time during a period of genogroup shift in Southeast Asia and Australia, from genogroup C (green) to genogroup B (red). Figure data are derived from the same Bayesian phylogenetic analysis method as Figure 3.3. For an animation of these data, refer to the electronic version³¹⁵.

With the implementation of the Enterovirus Reference Laboratory Network of Australia in 2009, enterovirus typing data was made available by members from different states of Australia. Viruses were identified from a number of different cases presenting with various diseases. Genogroup B EV-A71 viruses had been identified up to 2011, with no cases associated with AFP. Genogroup C4 EV-A71 was identified in 2011 (Figure 3.5), with an ancestral link to viruses circulating in China (Figure 3.6) and exhibiting a high sequence identity (99.8%) to viruses identified from an outbreak in Singapore from 2008 to 2009¹⁶⁴.

The genogroup C EV-A71 viruses were identified in place of the established genogroup B5, which had been introduced during the genogroup shift in 1999 and 2000 originating in Western Australia³¹⁶. By late 2012, genogroup C4 EV-A71 appeared to have become established, with the first case of AFP being reported to the NERL at the beginning of 2013.

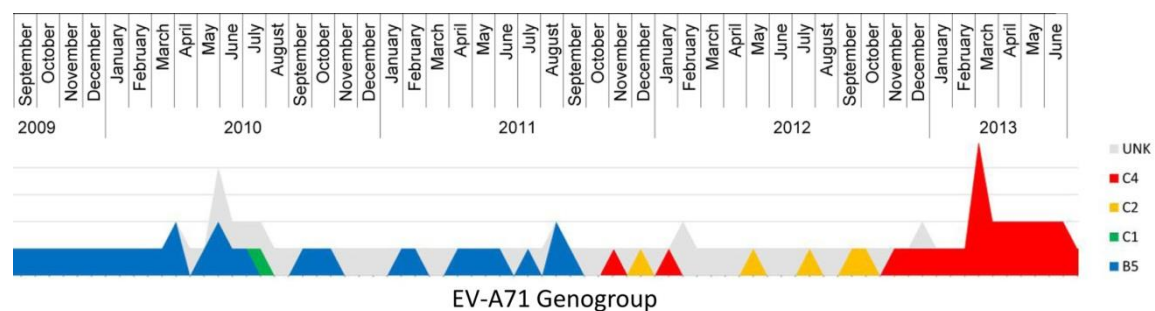


Figure 3.5. EV-A71 genogroup shift during the period of 2009 to 2013 from cases reported by the Enterovirus Reference Laboratory Network of Australia combined with data from AFP cases identified during the period 2012 to 2013 from the Enterovirus Reference Laboratory at VIDRL. Figure legend indicates EV-A71 sub-genogroup by colour B5= blue, C1= green, C2= orange, C4= red and unknown sub-genogroup= grey.

By mid-2013, twelve AFP cases associated with EV-A71 genogroup C4 had been identified through the ERLNA and AFP reporting systems. Samples from patients were only available from ten of these cases. One fatal infection was confirmed by the NERL as EV-A71 sub-genogroup C4, and a further four deaths were reported from cases exhibiting clinical manifestations of pulmonary oedema with a clinical suspicion of EV-A71 involvement. All cases occurred on the south-eastern seaboard of Australia during 2012 and 2013. Phylogenetic inference indicated that the viruses shared a common ancestor with the exception of a single AFP case in 2012, which was identified as belonging to

sub-genogroup C2. The fatal case with a positive EV-A71 sub-genogroup C4 identification also shared the same common ancestor with EV-A71 viruses circulating in China (Figure 3.6).

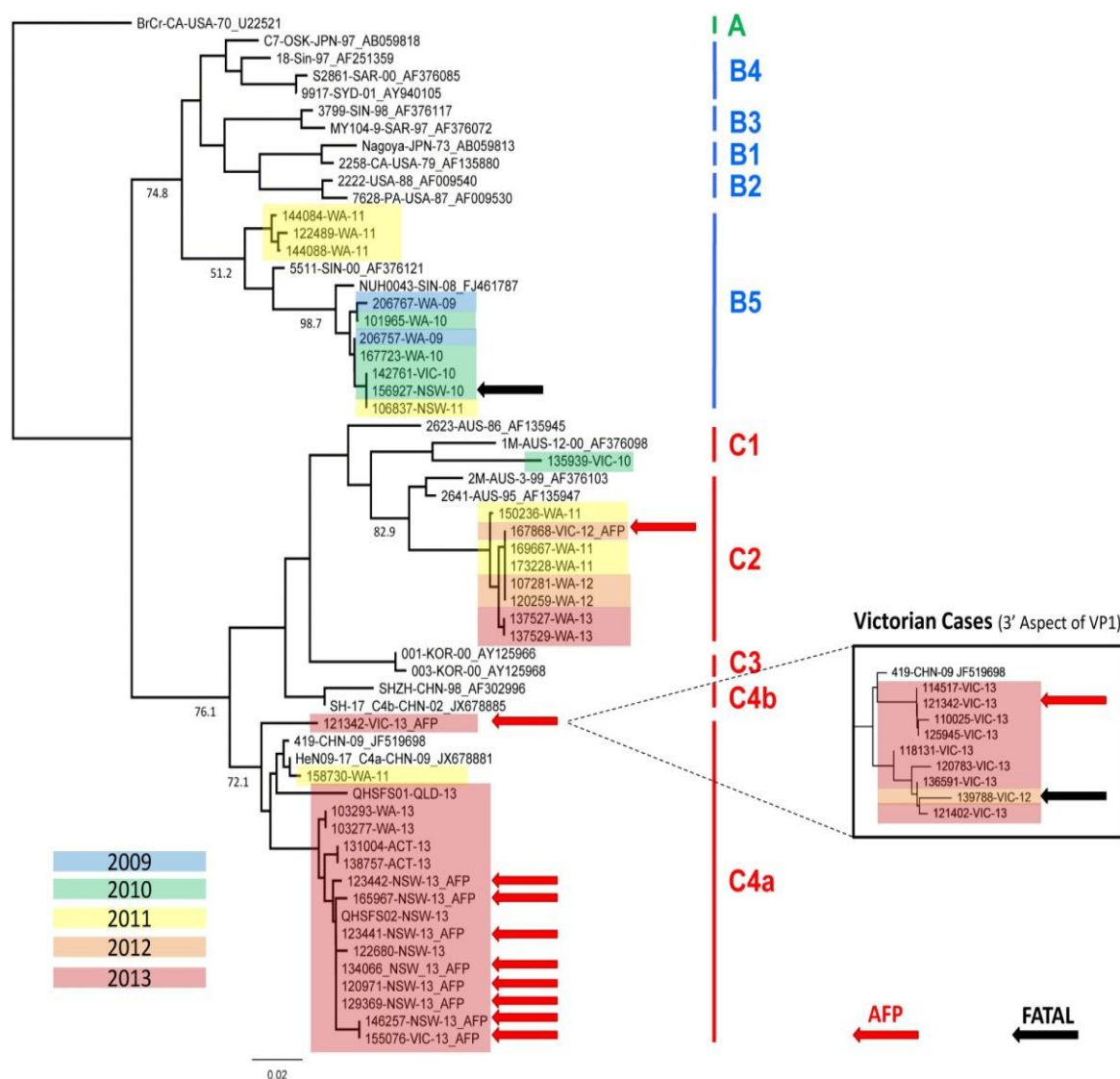


Figure 3.6. Phylogenetic analysis of enterovirus A71 identified in Australia during the period 2009 to 2013. Sequence data representing the 5' aspect of the VP1 genomic region (CODEHOP), derived from members of the Enterovirus Reference Laboratory Network of Australia, and the AFP surveillance system operating under the auspices of the Enterovirus Reference Laboratory. Maximum likelihood tree calculate using an HKY+G+I model with 1000 bootstrap pseudoreplicates. Red arrows indicate cases of AFP, black arrows indicate fatal cases. Breakout figure indicates comparative sequence analysis performed on the 3' aspect of the VP1 genomic region.

3.4.3 Characterisation of Variants of EV-C96

During the course of investigations of AFP cases from the PIC and PNG during the period 2004-2013, it was noted that a number of newly characterised members of the species *C Enteroviruses* were identified. Partial sequence data derived from the VP1 genomic region indicated that these viruses were typed as EV-C96. Retrospective examination of virus growth patterns in cell culture indicated a variation in tropism between different viruses identified (Table 3.7).

Identifier	Year	Country	HEL	HEP2C	RD-A	L20B
107711	2004	PNG	+	+	+	-
110581	2004	PNG	-	+	+	-
124709	2006	KIR	+	+	+	-
009265	2008	PNG	-	+	-	-
009269	2008	PNG	-	+	-	-
124097	2011	PNG	+	ND	+	-
134022	2012	FIJ	ND	+	-	-
173490	2012	PNG	ND	+	-	-

Table 3.7 Cell culture results of EV-C96 isolates. Data obtained for culture of enterovirus C96 in multiple cell lines. Human embryonic lung (HEL), human epidermoid carcinoma (HEP2C), rhabdomyosarcoma (RD-A) and mouse fibroblast expressing the poliovirus receptor is CD 155 (L20B). PNG= Papua New Guinea, KIR= Kiribati, FIJ= Fiji.

Partial VP1 sequences derived using the CODEHOP method from the respective viruses were analysed using phylogenetic methods. The samples from PNG indicated that they belong to a unique and distinct clade, separate from previously described EV-C96 viruses with sequences available in the public domain (GenBank). Distinct also from this group were two viruses identified from Samoa and Kiribati. Although the phylogenetic analysis indicated strong support for the Kiribati, PNG and Samoan EV-C96 viruses as a distinct lineage with respect to prototype species *C Enteroviruses*, there was poor support for clustering within the EV-C96 clade (Figure 3.7).

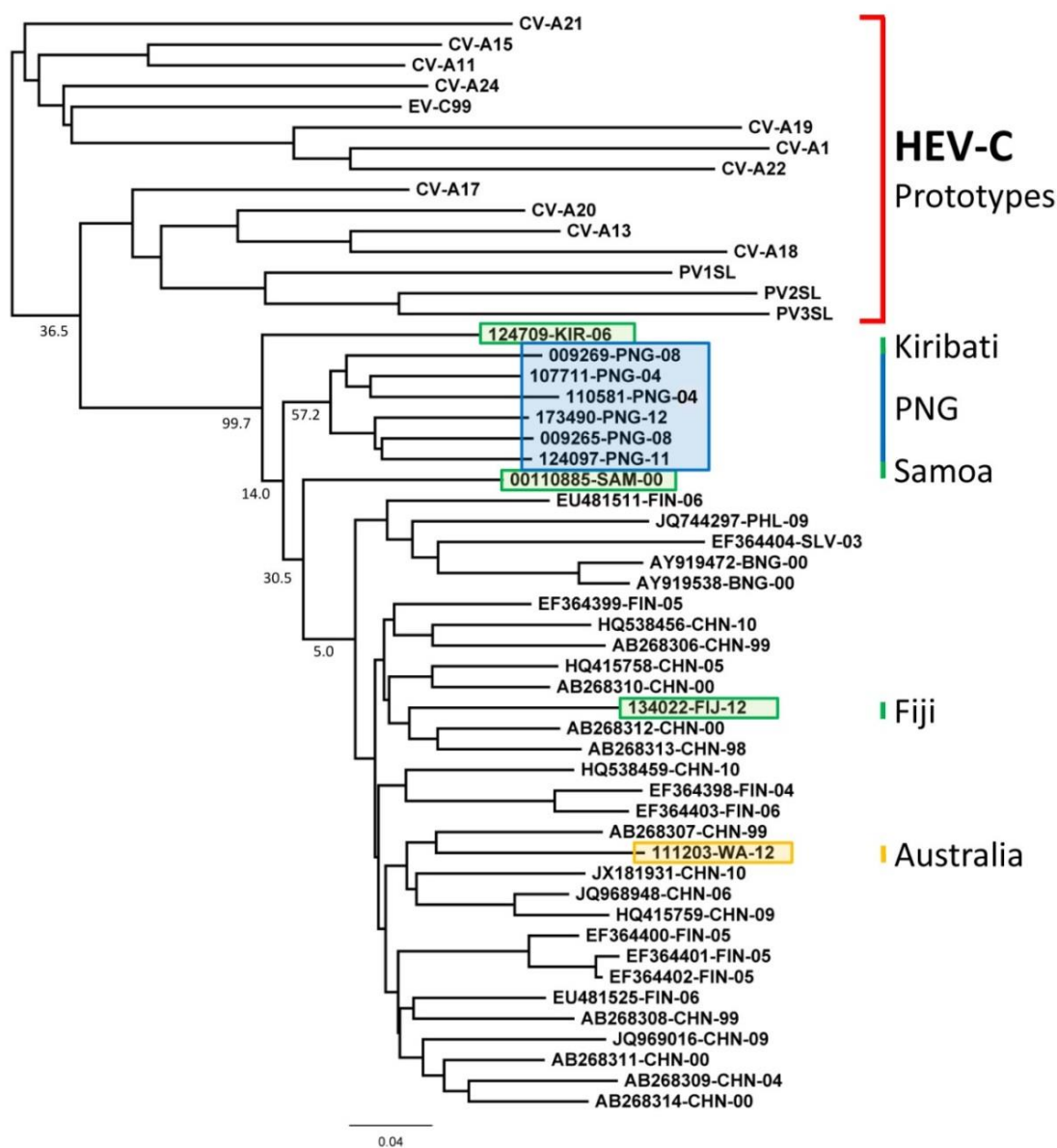


Figure 3.7 phylogenetic analysis of the newly described EV-C96 identified from Australia, PIC and PNG combined with prototype species *C. Enteroviruses*. Partial VP1 sequence encompassing the BC loop region, amplified using the CODEHOP assay. Phylogenetic tree created using DNA distance methods with 1000 bootstrap replicates.

3.4.4 Characterisation of the Novel Virus EV-A120

An enterovirus of unknown type was identified from PNG as part of AFP surveillance in 2009. Initial culture results from inoculation of faecal samples onto multiple cell lines indicated the presence of toxins, and dilution of samples using standard WHO protocols to reduce the effect of toxins on the cell cultures resulted in no growth. Screening using the PanEV rRT-PCR indicated the presence of an enterovirus. Subsequent testing using semi-nested semi-quantitative EV-SS snrRT-PCR (refer to section 2.3.1), indicated a virus titre of approximately 0.9 log CCID₅₀/mL and estimated at 8 infectious particles per millilitre of sample when compared to EV-A71 control virus (Figure 3.8).

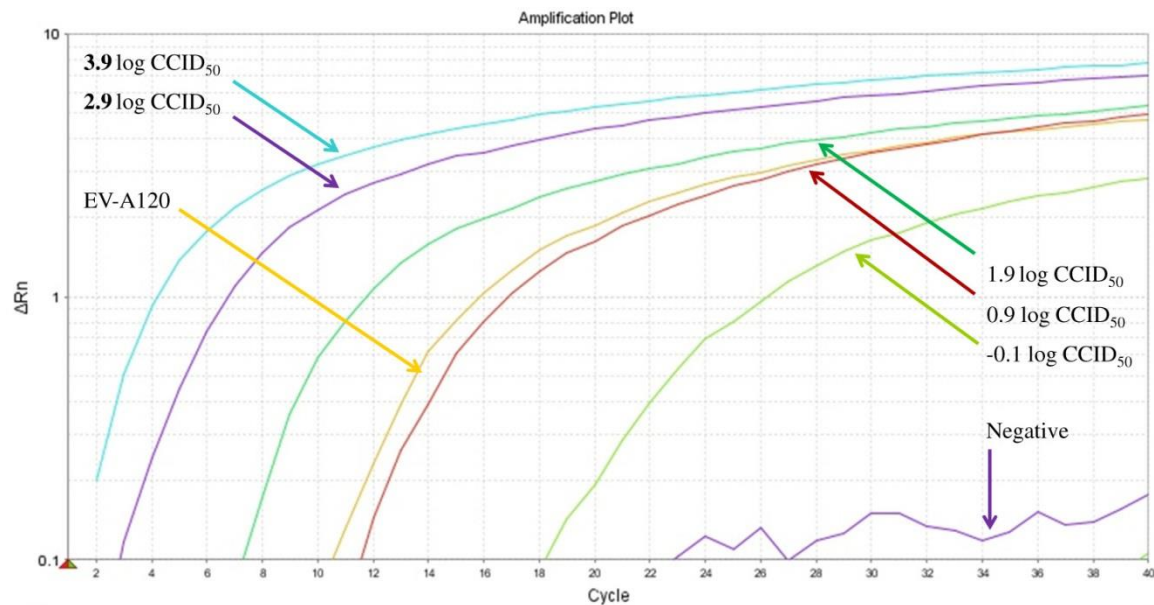


Figure 3.8. EV-A120 virus titre from original faecal sample. Semiquantitative EV-SS snrRT-PCR method incorporating a standard curve derived from serial dilutions (-0.1 to 3.9 log CCID₅₀) of quantitated EV-A71 genogroup B5 control material. x- axis = cycle number, y- axis = ΔRn (fluorescence).

In order to identify the enterovirus, sequencing using the CODEHOP method was performed and a 297 base pair fragment of the VP1 genomic region was obtained. This fragment was used to query the GenBank database online and no significant match was found. Full capsid sequencing was performed as described in section 3.3.3. Comparative analysis against known prototypes of enterovirus was performed, and indicated less than 75% nucleic acid and less than 88% amino acid identity against known prototypes (Figure 3.9) indicating that this virus was a novel type belonging to the species *A Enteroviruses* according to the guidelines examined in section 1.2.5¹¹².

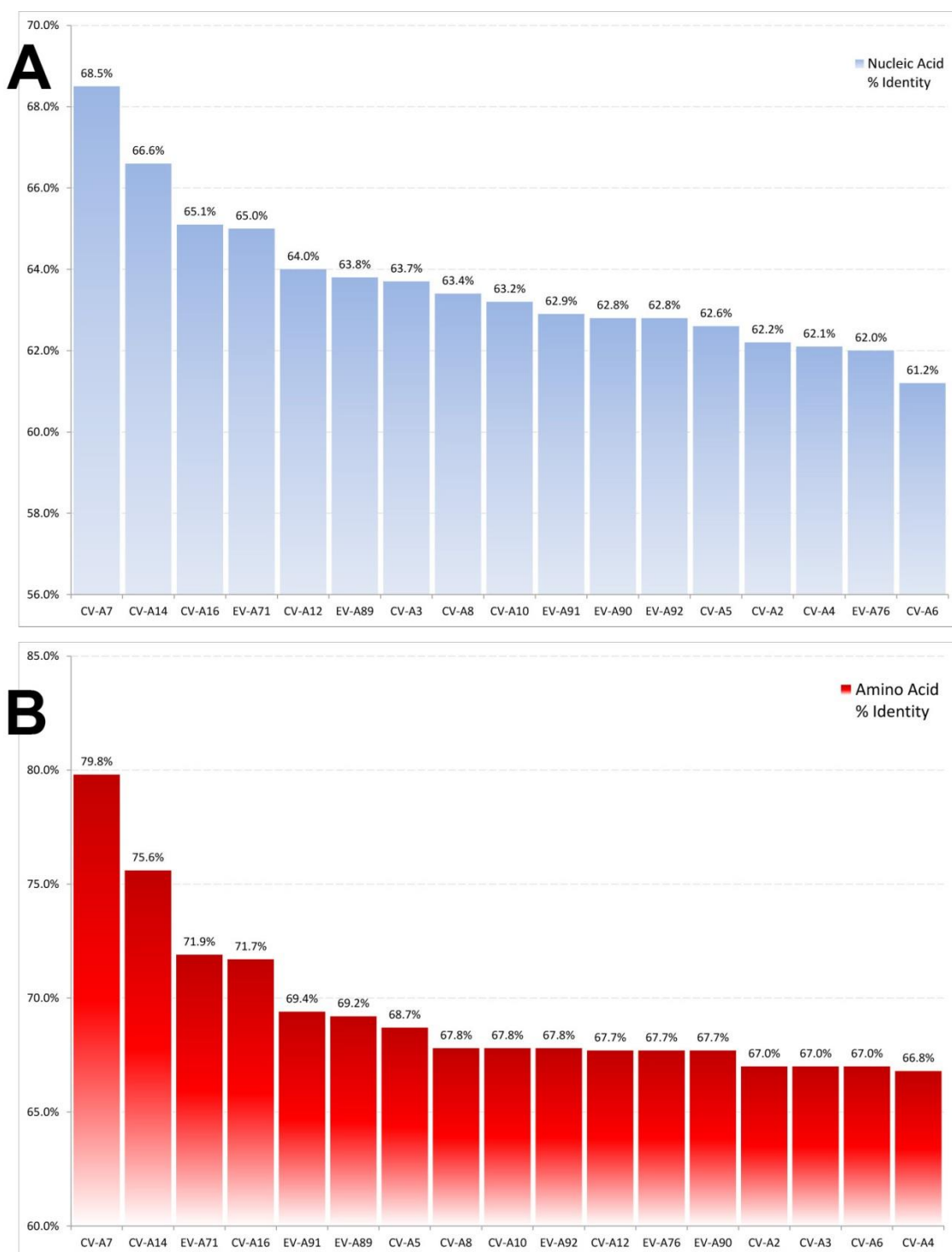


Figure 3.9. EV-A120 sequence identity to species A *Enterovirus* prototypes. Full capsid sequence of EVA 120 compared to all available prototype sequences representing species A *Enteroviruses*. A) nucleic acid % identity, B) amino acid % identity. Cut-off values for homologous enterovirus type are 75% for nucleic acid and 88% for amino acid identity. *x*- axis = prototype, *y*- axis = percent identity (%).

Sequence data representative of the complete capsid encoding region of all known species A *Enteroviruses* affecting humans was aligned with the novel virus sequence data using the Clustal-W algorithm¹²³. Once aligned, model testing was performed in order to determine the most appropriate phylogenetic analysis method using Mega version 5.2.1⁵⁰. Two models were determined to be the most suitable for use in phylogenetic inference; the generalised time reversible (GTR) and the Hasegawa, Kishino and Yano (HKY) models, with a discrete gamma variable incorporating invariant sites (Figure 3.10). Phylogenetic inference using Maximum Likelihood methods were performed using both models with 1000 bootstrap pseudoreplicates. No significant differences between outputs were observed.

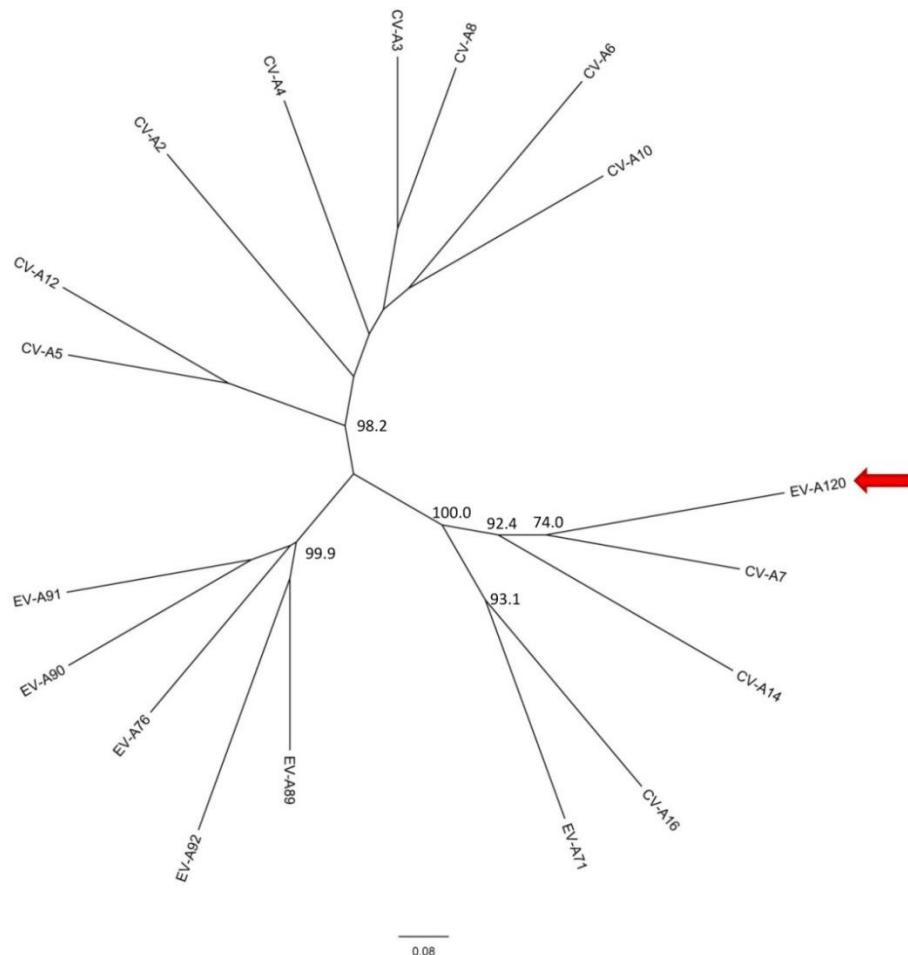


Figure 3.10. EV-A120 Maximum likelihood phylogenetic tree (HKY+G+I) with 1000 pseudoreplicates of *Enterovirus* species A prototypes with sequence data representing the entire capsid encoding region. Scale bar indicates nucleotide substitutions per site.

Analysis using phylogenetic inference provided strong support that the novel virus was related to a clade incorporating CV-A7, CV-A14, CV-A16 and EV-A71. Simplot analysis incorporating bootstrapping methods³⁰⁵ was performed for the entire capsid region using nucleic acid sequence data alignments representing the novel virus and available capsid encoding sequence data for related prototypic members of the species A *Enteroviruses*, (Figure 3.11). No significant re-combination events were detected, indicating that this virus is not a recombinant but given the phylogenetic analysis results, is most likely an evolutionary ancestor of CV-A7 and CV-A14.

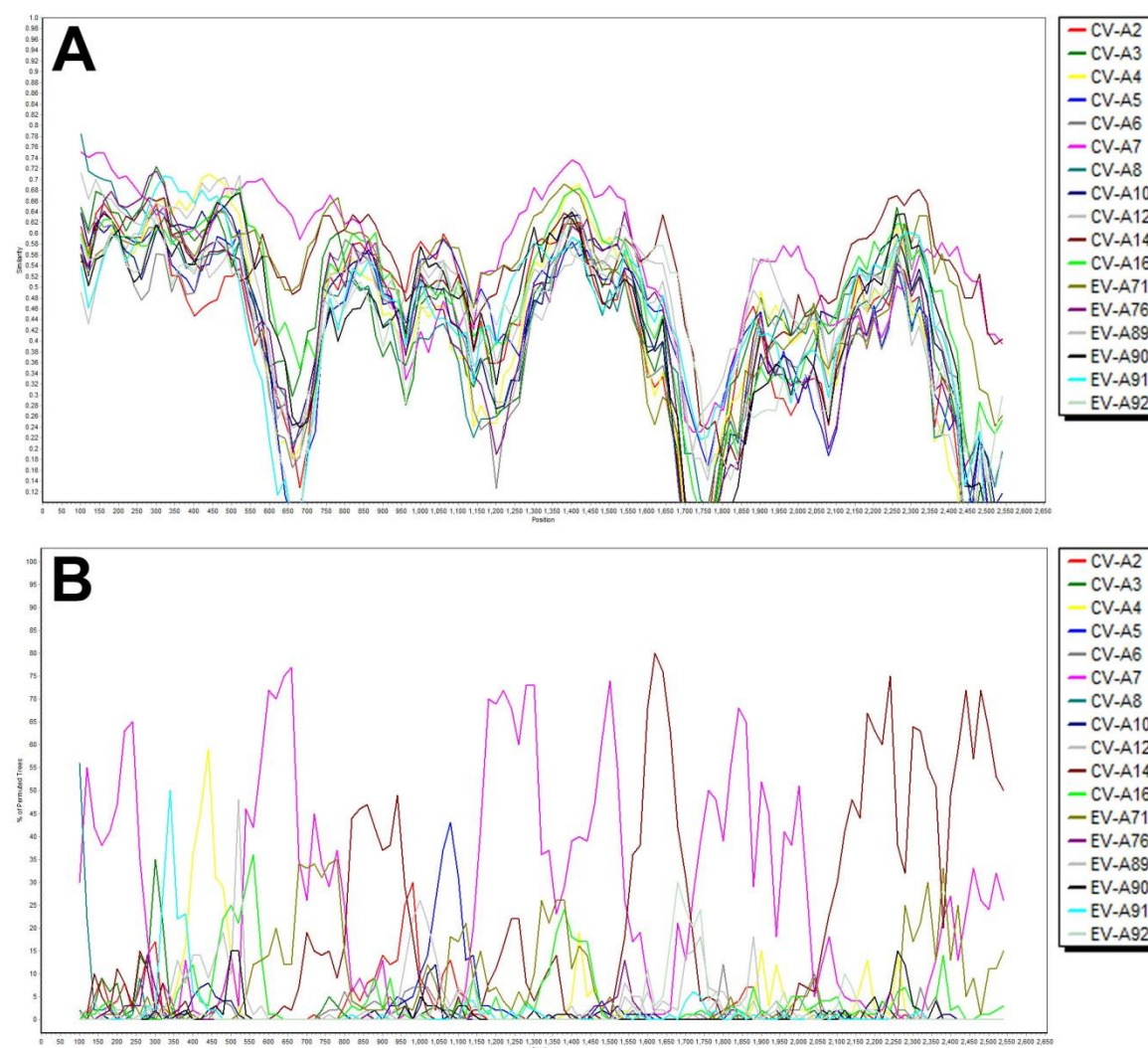


Figure 3.11 Simplot and BootScan outputs for EV-A120 and species A *Enterovirus* prototypes A) Simplot output indicating sequence identity below 75% nucleic acid threshold, B) BootScan analysis indicating heterogeneous matching that of the capsid region species A *Enterovirus* prototypes, the most significant matches are with CV-A7 (magenta), CV-A14) (brown). Simplot and BootScan performed with a 200 base pair window in 20 base pair steps using the Kimura-2 parameter method with neighbour joining, x- axis = nucleotide number, y- axis = A) similarity B) % bootstrap trees.

3.5 Discussion

The use of bioinformatics methods, incorporating algorithmic sequence matching in the form of BLAST searches, sequence alignment and phylogenetic analysis, are critical tools in establishing the molecular characterisation and epidemiology of viral pathogens. This is particularly evident in the case of serious neurological disease typified by the syndrome of AFP and the identification of associated enteroviruses.

The preceding work highlights the usefulness of identifying homologous enterovirus types, allowing the determination of viruses in sporadic cases of paralysis or those associated with outbreaks. A number of non-polio enteroviruses, including coxsackieviruses and echoviruses, have established associations with neurological infection that can lead to paralysis through their identification from CSF, a sterile site, and repeated detections from such cases⁴². However, a number of non-polio enteroviruses have been identified from non-sterile sites, such as stool specimens of AFP cases. More stringent analysis is required to establish a connection with severe neurological disease as it may be an incidental finding since it is estimated that approximately 90% of enterovirus infections are asymptomatic⁸⁰.

Presented in this chapter was the association of a large number of AFP cases with a major outbreak of EV-A71 subgenogroup C4, which also included reports of up to 5 fatalities in young children^{317,318}. The molecular epidemiology of this outbreak follows a similar pattern to what had been observed in 1999 and 2000, when the introduction of genogroup-B EV-A71 viruses in Western Australia replaced the circulating genogroup C viruses³¹⁶. In 2000, a second outbreak of EV-A71 genogroup B was reported in New South Wales. Between the years 2000 to 2011, the predominant circulating EV-A71 belonged to genogroup B^{316,319}. The subsequent detection of subgenogroup C4 in Australia in 2012 was a significant public health concern, since this strain has caused extremely large outbreaks in China and south-east Asia involving a high case fatality rate, serious morbidity and mortality³²⁰. Examination of EV-A71 sequence data from 1997 to 2003 using Bayesian Phylogeographic methods (section 3.4.2, figures 3.3 and 3.4), indicated two separate importation events into Australia. The first was in 1999 to Perth from Sarawak Malaysia followed by an importation event in 2000 to Sydney from

peninsula Malaysia. The detection and sequence characterisation of the first sub-genogroup C4 EV-A71 in Western Australia indicated a genetic similarity to viruses identified in outbreaks in Singapore and Malaysia in the preceding years. The chains of virus transmission were similar to those observed when examining the previous EV-A71 outbreaks in 1999 and 2000 in Australia, using Bayesian phylogeographic analysis.

EV-A71 infection is often associated with HFMD, but can also manifest as severe neurological illness such as anterior horn cell disease and meningo-encephalitis^{134,140}. The EV-A71 epidemic in Australia from 2012 to 2013 resulted in multiple cases of paralysis as well as deaths. Investigation of the outbreak was complicated by the fact that a number of other enteroviruses can cause the same symptoms, specifically coxsackieviruses also classified within species A. During the EV-A71 outbreak, coxsackievirus A6 (CV-A6) was co-circulating and caused the majority of HFMD cases investigated. In addition, echovirus 6 was most frequently associated with viral meningitis during this period (Figure 3.2)³¹⁷. This highlights the need to identify the types of enterovirus circulating concurrent with any cases being investigated and the clinical manifestations with which they are associated. The availability of rapid methods for enterovirus detection and identification, are essential to fulfil this need.

With respect to the enteroviruses identified in PNG and the Pacific islands, EV-C96 viruses were the most frequently associated with AFP. EV-C96 is a newly characterised virus first described in 2006 by Oberste *et al.*,¹¹⁴ and although the virus has no established causal association with AFP and severe neurological illness, the frequency of identification of this virus with AFP cases may indicate such an association. The NERL observed that this virus displayed variations in cell line tropism, indicating the potential for examining structural variations between viruses to identify a potential receptor binding site. The observation of varying cell tropism in EV-C96 isolates has also recently been observed by Lu *et al.*, (2014)³²¹.

The novel species A *Enterovirus* identified from a paralysed child in PNG, indicated via comparative sequence analysis, that it shared a close identity to CV-A7 and CV-A14. Although isolation of the virus in a case of AFP does not constitute a causal association, it is of interest to note that in the past, CV-A7 was referred to as poliovirus type 4 during large outbreaks of poliomyelitis like illness in Russia and the United States¹³³. Sequence data incorporating the 5'NTR complete capsid region and partial P2 nonstructural region

was sent to the ICTV, who have provisionally assigned the novel species *A Enterovirus*, as EV-A120³²².

The identification of the novel EV-A120 enterovirus was especially significant, given that the virus was not identified using conventional culture methods, but by amplification using PanEV snrRT-PCR during parallel testing of samples from AFP cases. The examination of virus titre using the newly developed RT-PCR methods in chapter 2, estimated the virus titre to be on the borderline of what is detectable using cell culture. The low titre of virus, combined with the toxicity of the original sample, complicated the identification and sequencing process, as a high titre virus was not available for further examination. Whole genome methods, such as those used in deep sequencing, have been used to examine novel viruses³²³⁻³²⁵, but the cost involved in using this technology was a barrier to the use of such methods. The use of primers designed from consensus data, combined with nested RT-PCR and primer walking, allowed the amplification of sequence data representing almost the entire 5' aspect of the genome, which importantly, contains the P1 region encoding the virus capsid. The sequencing methods used to derive the EV-A120 sequence data proved so effective, that an identical process was used to determine the capsid encoding region for multiple EV-C96 viruses identified from AFP cases.

In summary, the retrospective analysis of data from 2004 to 2013 identified two viruses (EV-A71 and EV-C96) as having a relatively high association with AFP in Australia, the Pacific Islands and Papua New Guinea. Further investigation of these viruses was considered warranted, and thus, these viruses were selected for 3-dimensional structural characterisation. Additionally, the discovery of EV-A120 as a novel pathogen in specimens derived from a case of AFP provided the opportunity to test whether a model of the capsid structure could be generated purely from sequence data, without the need to grow it in culture. The reconstruction and examination of 3-dimensional structures for these selected viruses are described in Chapter 5.

Chapter Four

Reconstruction and Simulation of Poliovirus

4.1 Manuscripts Published in Association with this Chapter

- **Roberts, J.A.**, Kuiper, M.J., Thorley, B.R., Smooker P.M., Hung. A., 2012. Investigation of a predicted N-terminal amphipathic α -helix using atomistic molecular dynamics simulation of a complete prototype poliovirus virion. *Journal of Molecular Graphics and Modelling*, 38, September: 165-173. [Refer to appendix A8.2.6].

4.2 Introduction

The examination of the molecular structure of viruses has been an important part of understanding processes such as antiviral drug binding or receptor interactions (refer to sections 1.4 and 1.5.1). Studies examining molecular interactions typically rely on atomic coordinate data derived from experimental methods, but such structural data are not always available for viruses that may be of interest. In order to determine the feasibility of reconstructing and simulating structural components of viruses for which no structural data exist, new processes were required to recreate partial and complete virus structures for simulation and comparative analysis. A number of methods incorporating the use of comparative protein modelling can be utilised, to produce models of viruses that have not had structural data reported²⁴⁹⁻²⁵⁵. The availability of atomic coordinate data derived from x-ray diffraction and cryo-EM experiments of closely related viruses can be the basis of reconstructing an unknown virus, using software such as Modeller, SwissPDBviewer and SwissModel^{326,327}.

Model prediction can only be based on structural data that is complete and of reasonable quality^{328,329}. There are predictive modelling methods that use energy minimization (for example Modeller or SWISSmodel) which help to resolve issues related to steric clashes^{330,331}. Clashes may exist in predicted models where amino acid side chains have been incorrectly placed. Rotation of these of side chains may resolve these issues, but not necessarily in the case of deeply buried amino acid side chains, where a combination of spatial positioning, charge and hydrophobicity may, for example, result in unexpected displacement of structures within the immediate vicinity. Furthermore, crystal structures and the models on which they are based, even those of very high quality while providing a “static” view of a protein, may lack detail on atomic motions (refer to sections 1.3 and 1.5).

Crystal structures provide at best a “snapshot” of the protein under experimental crystallisation conditions, which are often far from physiological conditions (refer to section 1.3.1). They give no direct information on the stepwise mechanical processes affecting conformational transitions as a result of pressure, temperature, ionic interactions, or interactions within the context of a biological system. Atomistic MD simulations, which model motions of individual atoms at finite (physiological) temperature and water and ion solvent, are complementary to experimentally derived structures, providing details of protein dynamics of possible functional importance under variable conditions. These behaviours are restricted to conditions under which the model parameters – force fields - have been derived, and thus are determined to be suitable.

To demonstrate the utility of MD to elucidate biologically-relevant motions of viruses, and to demonstrate its appropriateness to model newly described and novel viruses, a series of *in-silico* experiments was required using a virus that has been extensively researched, with a large number of available structures. For this purpose, wild-type poliovirus 1 was chosen, due to the large body of literature dating back over 100 years, and the wealth of research data relating to the investigation of its structure.

The wild poliovirus type 1 (Mahoney strain) capsid was resolved in 1985 to a resolution of 2.9 Å, representing one of the first mammalian viruses to be resolved using x-ray diffraction methods¹⁶. Although the complete structure in the external capsid surface was elucidated, certain regions within the interior could not be resolved from the electron density data, providing an opportunity for MD to fill this gap. Only partial data existed of the N-terminus of VP1, which had been postulated to be in the form of an α -helix³³². This α -helix is considered a critical component in the uncoating process and extrication of RNA across the cell membrane into the cell cytoplasm⁸⁹. Residues 1-7 of the N-terminus of VP2 and residues 234-238 of the C-terminus of VP3 were not present in the 1985 structure, but indications of their placement can be found in related structures such as those of the empty capsid form of the virus elucidated in 1994⁶⁸.

The N-termini of VP1, VP2 and VP4 are each located on the inner surface of poliovirus and it was postulated that resolution of these structures was not possible by x-ray crystallography, due to the spatial orientation compared to the symmetric icosahedral

nature of the virus structure¹⁶. Prior to assembly into a final capsid structure, biologically active protein structures in the form of a pentamer precursor are present, comprising five protomers assembled into a pentameric configuration using VP1, VP3 and a single combined precursor configuration of the proteins VP2 and VP4, termed VP0 (refer to sections 1.2.2 and 1.2.3). VP0 is postulated to undergo cleavage of a scissile bond in the presence of RNA during the assembly of the infectious 160S particle from the 70S form in the process of creating the separate proteins VP2 and VP4, thereby locking the capsid in its final and infectious conformation⁶⁸.

Since the initial atomic description of the poliovirus and rhinovirus capsids in 1985, more than 70 x-ray diffraction or cryoEM derived structures, representing various capsids of members of the family *Picornaviridae*, have been made available on the RCSB protein data bank. As of February 2014, 25 atomic coordinate files exist, representing poliovirus 1, 2 and 3 capsid structures in various states. However, it was noted that the N-terminal region of VP1 was not resolved by any of these investigations for species C *Enteroviruses*. This may have been due to spatial disorientation or molecular movement of the protein chain. The proximity of the helix to the RNA core within the virus particles was not considered to be a contributing factor during the generation of the first crystal structure coordinates published in 1985³³³. The amino acid sequence of the terminal sequence of the VP1 protein in many enteroviruses, including poliovirus and rhinovirus, indicated a periodicity of hydrophobic residues that may form an amphipathic helix with polar residues along one half of the surface and non-polar residues on the opposite surface⁸⁹. The complete VP1 crystal structure of human rhinovirus 16 was resolved in 1985 at a resolution of 2.15 Å and indicated the formation of an amphipathic helix at the N-terminus, supporting the hypothesis that this occurs in a similar manner in other enteroviruses³³⁴.

The manual inclusion of missing structures, whose positioning can be determined from partial data accumulated from a number of different ancestral structures, is a process that is easy to perform using current conventional methods with programs like Visual Molecular Dynamics (VMD)³³⁵, but the manipulation of these data needs to be performed with caution and validated against existing known conformations of structure.

This chapter describes the reconstruction of wild poliovirus type 1, incorporating predicted structures that were missing from the original X-ray diffraction data. The model was subjected to MD simulation and analysis against the original capsid structures, to validate the method. Previous MD simulation methods of the virus capsid utilised symmetry methods, in which the icosahedral rotational symmetry of the capsid were exploited to model the entire capsid. Specifically, a single protomer was modelled within a periodic cell, creating the effect of the protomer being embedded in a full virus capsid, approximated using icosahedral periodic images of the asymmetric unit (refer to section 1.5)²⁵⁴. Exploiting icosahedral symmetry does manage to avoid the high computational overheads associated with large atomistic MD simulation by being only thousands of atoms in size, but it also fails to emulate any unknown event that may break symmetry (such as RNA extrusion)^{91,336}, and thus do not adequately represent the biological state of the complete virus capsid.

A novel protocol for construction of full poliovirus virions is discussed in this chapter, and the simulation and hardware requirements for simulating them at all-atom resolution are established. In particular, this chapter describes: construction of missing virion structural components and a novel procedure for encapsulating the RNA; the establishment of appropriate conditions to simulating virions, including PBC cell shape/size and computational architecture; demonstrating the stability of virions under simulation conditions and thus the applicability of *in-silico* modelling to construct biologically-realistic virion models which can be used to investigate known external factors which influence gross morphological structure. Virion morphology is examined in response to variations in temperature conditions, as are the influences of known capsid binding compounds (palmitate and Disoxaril, respectively). Finally, a number of quantitative and qualitative analyses were performed on the trajectories obtained, to demonstrate the range of properties which may be elucidated by MD, including structural refinement of a region of the virus unresolved by crystallography.

It is anticipated that the simulation protocol and preliminary analyses established in this chapter will form the basis for development of new analysis methodologies, adapted for large biological systems, thus enabling more effective quantitative interrogation of MD simulation trajectories.

4.3 Materials and Methods

4.3.1 *In-Silico* Reconstruction of Poliovirus Capsids

The reconstruction of poliovirus capsid and complete virions were based on GenBank sequence ID NC002058. Database searches using the SwissModel website determined that of the available enterovirus crystal structures, the x-ray co-ordinate data of poliovirus type 1 Mahoney strain (PDB file 1HXS) at a 2.2 Å resolution, was the most suitable template³³⁷.

Poliovirus capsid reconstruction was performed using the Visual Molecular Dynamics (VMD) software³³⁵, incorporating the nVidia GPU-specific Compute Unified Device Architecture (CUDA) extensions. A significant element lacking from the original crystal data was an α -helix located at the amino terminus of the VP1 protein. The 19 missing residues constitute a region containing an α -helix with an amino acid motif common to members of the genus *Enterovirus*. PSIPred predictions^{338,339} derived from amino acid sequence data and comparative information determined from available enterovirus crystal structures in the Research Collaboratory for Structural Bioinformatics protein data bank (RCSB PDB), assisted in the spatial alignment of the helix, which was reconstructed using the molefacture plug-in of VMD.

Icosahedral symmetry data required to reconstruct the 60 virus protomers into a full capsid structure were available in the form of metadata within the original crystal structure³³⁷. Using the "mono2poly" Tcl script, a single protomer was positioned according to the biological parameters to assemble the 240 proteins and 120 lipids required for a full poliovirus capsid³⁴⁰.

An additional 3 million atoms were required to replicate a 0.154mM sodium chloride suspension, with cell edge padding at 12 Å on each 3-dimensional axis. The final atom count for the entire cubic representation of the empty capsid simulation was approximately 3.9 million, with cubic cell dimensions of 344.8 Å on each axis. Additional model configurations involving complete virions, pentamers and a detached helix were constructed in preparation for MD simulation (Table 4.1).

Simulation	Detached helix	Pentamer	Empty Capsid	Complete Virion
Total Atoms	8,586	388,521	3,938,950	3,301,378
PBC	Cubic	Cubic	Cubic	Rhombic-dodecahedron
Water molecules	2,760	106,629	1,038,857	745,287
Amino acids	20	4,390	52,800	52,822
Nucleotides (RNA)	0	0	0	7,640
Fatty Acids	0	10	120	120
Magesium ions	0	0	0	3,799
Sodium ions	10	370	3,728	2,785
Chloride ions	8	309	3,011	2,267

Table 4.1 Table outlining specific details of each simulation. Table indicates the models created for simulation, with the periodic boundary conditions (PBC), the type of molecules and the total atoms used.

In addition to the wild type Mahoney poliovirus, a homology model was attempted to construct the related Sabin vaccine strain. As no crystal structure data existed for the Sabin strain of poliovirus type 1, the Mahoney wild poliovirus strain, which shares 98.5% capsid amino acid sequence identity with the vaccine strain, was used as a template for comparative protein modelling using the SWISS-model server based on the 1HXS Mahoney PDB data.

Topology files and CHARMM parameter data for the associated myristate and palmitate molecules were determined using the SWISSparam server^{341,342}.

4.3.2 Integration of Wild Poliovirus RNA using Steered Molecular Dynamics.

As there were no structural data available for the reconstruction of the internal core of the virus containing genomic RNA and the covalently-linked 22 amino acid VPg protein, a simple approach was used, involving 10 RNA fragments created using the "make-na" server³⁴³ and genome fragments compressed (from 2000 Å to 300 Å) using *in-vacuo* steered MD, were arranged in a simple helical structure (Figure 4.1). Each of the 10 RNA fragments was 750 bp, and a 20 bp poly-A tail was attached to the 3'NTR of the final RNA fragment. The VPg protein was replicated using the PDB file 2BBP and was covalently attached to the 5'NTR of the genome. Appropriate angles and partial charges

were determined using the NMR data derived from the published work on the VPg protein and its association with the 5'NTR of the genome^{237,238}. Once constructed, the VPg and the first nucleotide of the Wild Poliovirus 5'NTR were uploaded to the SwissParam server to obtain appropriate topology and parameter files for simulation (refer to appendices 5.6 and 5.7 for topology and parameter files).

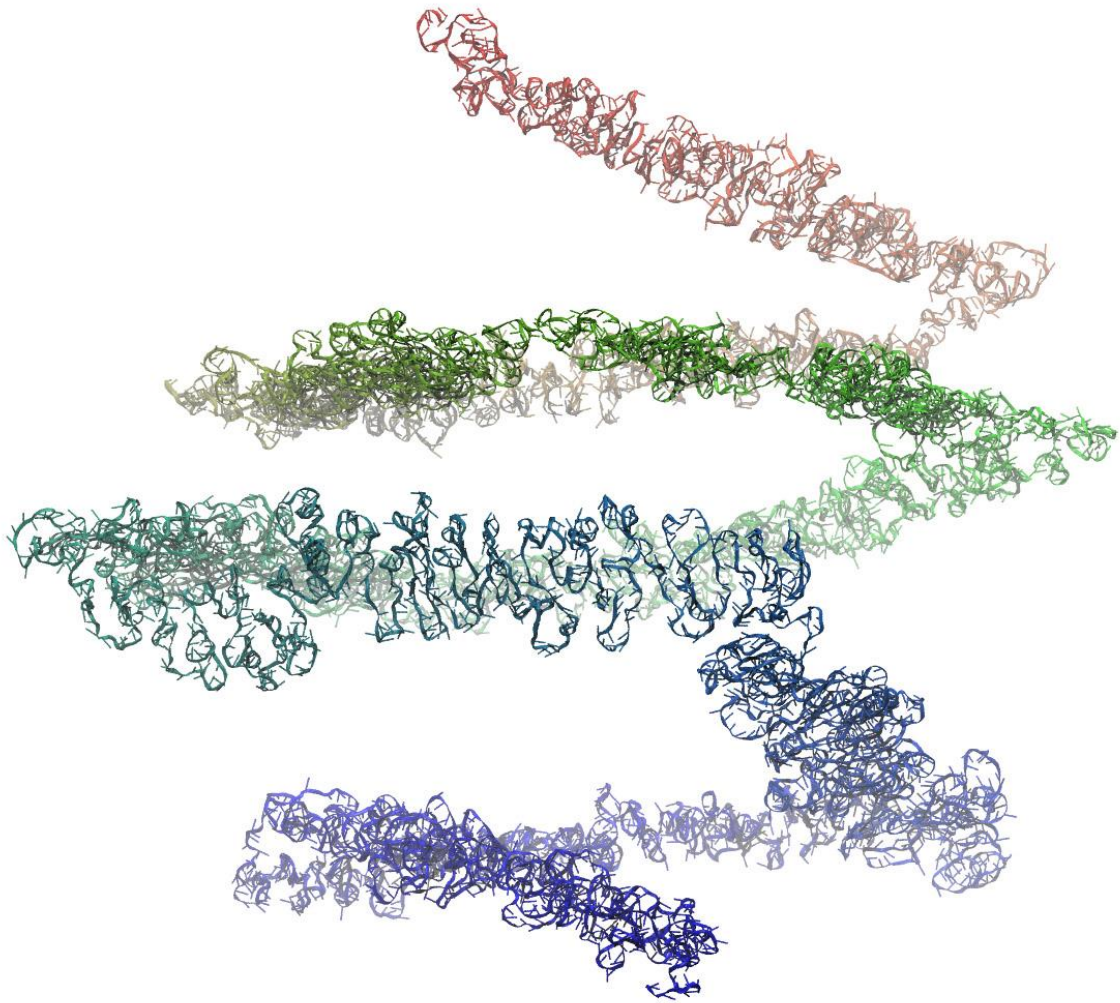


Figure 4.1 Helical arrangement of complete genome prior to Steered MD. Simplified “cartoon” representation of the supercoiled structure of the Wild Poliovirus RNA genome, prior to compression and placement in the core of an empty capsid structure. Colour scale: Red (nucleotide position number 1) to Blue (poly-A tail to 7640bp). Image generated using the VMD software package³³⁵.

Steered MD was then used to direct the super coiled RNA into a sphere with of diameter of approximately 300 Å, *in-vacuo* (refer to appendix A5.1 for Tcl scripts). The now spherical RNA was then solvated and neutralised in a magnesium chloride solution. Using a constant atom number, constant pressure and constant temperature ensemble (NPT) with steered MD, the RNA was compressed into a smaller sphere with a 184 Å diameter and then transposed into the core of the wild poliovirus capsid, creating a final structure totalling 1.05 million atoms.

4.3.3 Molecular Dynamics Simulation of a Complete Wild Poliovirus Virion.

MD simulation was performed using the 64-bit Nanoscale Molecular Dynamics (NAMD) software package³⁴⁴, which incorporates the CHARMM force fields. A number of different in-house computer systems and supercomputers at VIDRL and the Victorian Life-Sciences Computation Initiative (VLSCI) were used. Simulations were run on 256 cores of an SGI Altrix XE supercomputer (Intel x86 architecture) running NAMD version 2.7, 512 nodes (2048 cores/threads) of a BlueGene/P running NAMD 2.8 beta 1 with memory optimisation extensions and 256 nodes (4096cores, 16,384 threads) and on a BlueGene /Q running NAMD 2.9 with memory optimisation extensions. In-house simulations were performed on a Dell T3500 workstation incorporating and Nvidia Quadro 5000 and Tesla C2070 CUDA-based GPU accelerator, using version 2.8 beta 1 and a 2.9 of NAMD, which included the GPU-specific CUDA software enhancements.

Initially, 50,000 steps of energy minimisation were performed. This was followed by equilibration using an NPT for 0.5 ns at 310 degrees Kelvin with a pressure of 1013.25 mbar (one atmosphere). Once the process of energy minimisation and equilibration had been completed, the virus simulation was moved to a BlueGene/P supercomputer running NAMD 2.8 beta 1, with memory optimisation extensions on 2048 cores (refer to appendix A5.8 for specific parameters). From August 2012 onwards, simulations were performed exclusively on a BlueGene /Q supercomputer, using 4096 cores (16,384 threads). NPT simulations of a detached α -helix (8,586 atoms), pentamer (388,521 atoms) and empty capsid (3.94 million atoms), were run for a minimum of 10 ns each. Full virion simulations totalling proximally 3.3 million atoms using rhombic dodecahedral periodic boundary conditions were run for 0.1 μ s. Trajectory data for each

simulation were then transferred locally and analysed on site at VIDRL using a Dell T3500 workstation.

Refer to appendix A2 for further details of equipment used.

Once an appropriate workflow was established to create stable full virus simulations, the methodology was applied to a series of computational experiments investigating the effects of temperature variation on gross virus morphological changes. The complete virion experiments performed as above were repeated, but at temperature settings of 4°C, 37°C, 60°C and 95°C for 10 ns each, in order to observe any structural changes in capsid conformation.

4.4.4 Examination of trajectory data using VMD

The wild poliovirus MD simulations (and subsequent simulations presented in chapter 5) were performed with atomic position data (trajectories) collected at 0.1 ns time-steps and output into the NAMD “DCD trajectory file” format. The Visual Molecular Dynamics (VMD) software package was used to analyse these data sets. VMD is capable of not only examining static structural data and creating files for simulation, but is also a useful tool for the examination of large trajectory data sets, limited in size only by the amount of Random Access Memory (RAM) available on the computer system used for the analysis, in this instance 24GB for the Dell T3500 computer. VMD is also capable of running analysis scripts written in Tcl or the python languages. A number of Tcl scripts were created and used in the examination of data derived from the NAMD simulations (refer to appendices A5.2, A5.3 and A5.4).

One of the most widely used and accepted methods for the examination of particle displacement over time is the Root Mean Square Deviation (RMSD) method³⁴⁵. Calculated using the formula (4.1):

$$\sqrt{\frac{1}{N} \sum_{i=1}^N |x_i - y_i|^2} \quad (4.1)$$

RMSD is the square-root of the averaged distance (squared) between the two atoms (x and y) being compared. This results in a numerical value, usually presented in Å, when used for comparative measurement of biological systems. RMSD calculations are used frequently in the following chapters of this thesis, as a coarse measurement of movement of protein or nucleic acids within an MD simulation, or when comparing x-ray diffraction data with MD simulation outputs.

Another method for the examination of structural deviation over time is the Root Mean Square Fluctuation (RMSF) method, which examines the fluctuation of a particle around a defined average position and is calculated using the following equation (4.2):

$$\sqrt{\frac{1}{T} \sum_{t_j=1}^T (x_i(t_j) - \tilde{x}_i)^2} \quad (4.2)$$

The analysis methods performed in this thesis will focus on RMSD observations. RMSD observations over simulation timespan, can be a good indicator of structural stability, as shown by Freddolino *et al.*, (2006), in the examination of stable and unstable states of STMV¹. Stable structures were indicated by a modest initial rise in RMSD in the first 100 ps of the simulation, followed by a plateau phase that resulted in a 1 Å rise in RMSD over a 10 ns time-frame. Unstable structures continued on a gradient following the initial 1 Å rise in RMSD and continued to between 4 Å and 6 Å depending on the simulation temperature (298K or 310K respectively).

In addition to the calculation of RMSD values to examine stability, the external virus capsid dimensions were determined over the course of the simulations. This was achieved by calculating the distance of the farthest α -carbon at each of the 6 axes (x , $-x$, y , $-y$, z and $-z$) from a defined point at the centre of the capsid structure (refer to appendix A5.2.1). These values were taken from the trajectory data at 0.1 ns time points over the duration of the simulation. The aggregation of these data allowed for the observation of capsid geometry over time. Substantial changes to this geometry would indicate an unstable structure¹.

4.4 Results

4.4.1 Optimization of Simulation Cell Morphology and Size.

A cubic periodic cell, for the simulation of an icosahedral structure, is inefficient. When encompassing what is essentially a spherical object in appearance, cubic cells require an excessive number of water molecules to fill the cell, resulting in the placement of a large number of redundant water molecules. These water molecules would be beyond the interaction range of the icosahedral-shaped virion and therefore would exert little direct effect on its behaviour. For this reason, rhombic-dodecahedral periodic boundary conditions were chosen to reduce the amount of water required to solvate the virus, with the advantage of better reproducing the *in vitro* aggregation of multiple picornavirus particles (Figure. 4.2).

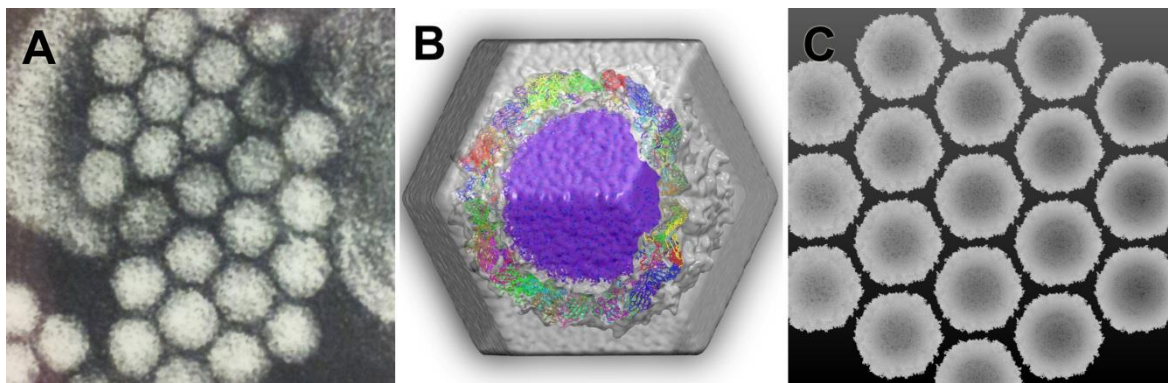


Figure 4.2 Rhombic-dodecahedral periodic boundary conditions A) Negative-stained transmission electron micrograph of a picornavirus sample showing honeycomb-like aggregation of virus particles. Image credit: Electron Microscopy Department, VIDRL. B) MD simulation of full wild poliovirus virion suspended in saline solution with transparent density representation of water used to show periodic cell boundaries. C) Transverse section of periodic cell reflections indicating the similar "honeycomb-like" appearance of the MD simulation of a full wild poliovirus virion to electron microscopy data. Figure reproduced in Roberts *et al.*, (2012)².

Initial *in-silico* reconstructions of the solvated wild poliovirus virions were in the order of 3.94 million atoms using a cubic PBC with 12 Å padding (Table 4.2). The adoption of a rhombic-dodecahedral PBC allowed the atom count to be reduced to 3.30 million atoms with a 16 Å padding, allowing a greater separation of structures contained within periodic cells, thus preventing any potential interactions of the viruses between periodic cells. A simulation time of 1.92 days/ns (0.52 ns/day) was achieved when using 1024 cores (1024 threads) of a BlueGene/P supercomputer . In 2012, the supercomputer was

upgraded to a BlueGene/Q system and simulation times of 0.332 days/ns (3.01 ns/day) were achieved when using 4096 cores (16384 threads), with a maximum throughput of (0.09 days/ns or 11.1ns/day) using 16,384 cores (32,696 threads). Beyond this point, inherent restrictions relating to available memory within the BlueGene hardware configuration, prevented an increase in core usage and therefore throughput.

Simulation*	Empty Capsid	Complete Virion	Complete Virion
Total Atoms	3,938,950	2,754,142	3,301,378
PBC	Cubic	Rhombicdodecahedral	Rhombicdodecahedral
PBC Padding	12 Å	12 Å	16 Å
ns/Day (1024cores)*	0.52	0.66	0.59
Speed-Up	N/A	1.27x	1.13x

Table 4.2 Performance comparison of different periodic boundary conditions. Table outlining specific details of each simulation including periodic boundary conditions (PBC) and total atoms used.

*Statistics for comparison derived using the BlueGene/P supercomputer.

4.4.2 Examination of NAMD on Different Processor Architectures

During the *in-silico* reconstruction of the capsid models, it was discovered that the 32-bit architecture of the Microsoft Windows XP operating system was unable to process simulation data, due to an inherent 4 GB memory restriction. Structures which totalled more than four million atoms when solvated in a virtual saline solution exceeded this 4 GB memory restriction, especially when examining trajectory data greater than 10 ns in length.

In order to circumvent the restrictions experienced within the Windows operating system, and also to preserve the established computing environment used in the laboratory, a 64-bit "live Linux" distribution³⁴⁶ was utilised, in which the operating system (including the necessary programs for model construction, atomistic MD simulation and output analysis) were able to "boot" on any of the 64-bit capable computer systems tested. In order to enhance the portability of the device, the "live" operating environment in the form of Ubuntu version 10.04 LTS was installed on a MicroSDHC device.

Concurrent with this issue relating to the operating system environment, three different iterations of NAMD were available. Preliminary investigations had been performed using the NAMD 2.7 software package on an x86 architecture, but a new version of NAMD

became available and incorporated enhanced memory optimisation features for the BlueGene architecture as well an increased feature set for the CUDA GPU architecture.

To determine whether there were any statistically relevant impacts of using different versions of NAMD, a series of experiments using empty capsids representing wild and Sabin-like strains of type 1 poliovirus were performed. Each simulation was run for a total of 1 ns and comparisons of each architecture for x86, CUDA and BlueGene were made. Gross assessment of capsid stability was performed using RMSD calculations for each virus protein, VP1 VP2 VP3 and VP4 (Figure 4.3). Structural motifs for each viral protein were assessed using the RMSF method (Figure 4.4). No significant difference was observed between different versions of NAMD software or hardware architecture.

At the conclusion of these experiments, the Dell T3500 workstation was converted to a dual boot operating system, comprising the original windows environment and 64-bit Ubuntu version 10.04 Linux distribution. Refer to appendices A2 and A3 for further details of equipment and software used.

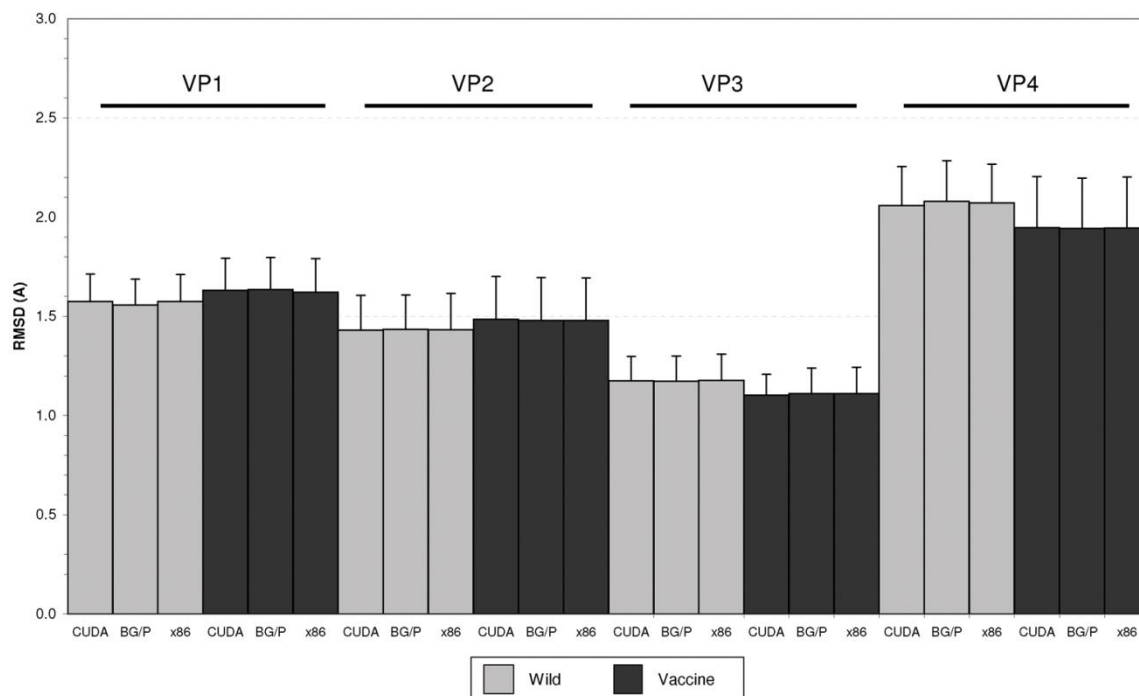


Figure 4.3 Comparison of Computer Architectures for Simulation Environment. RMSD values for proteins VP1 to VP4 (wild and vaccine strains of poliovirus type 1), computer architecture comparisons. Mean RMSD value 60 virus protein replicates found within a capsid, error bars indicate standard deviation. CUDA = NAMD 2.8b1 CUDA, BG/P = NAMD 2.8b1 "mem-opt" feature enabled, x86 = NAMD 2.7.

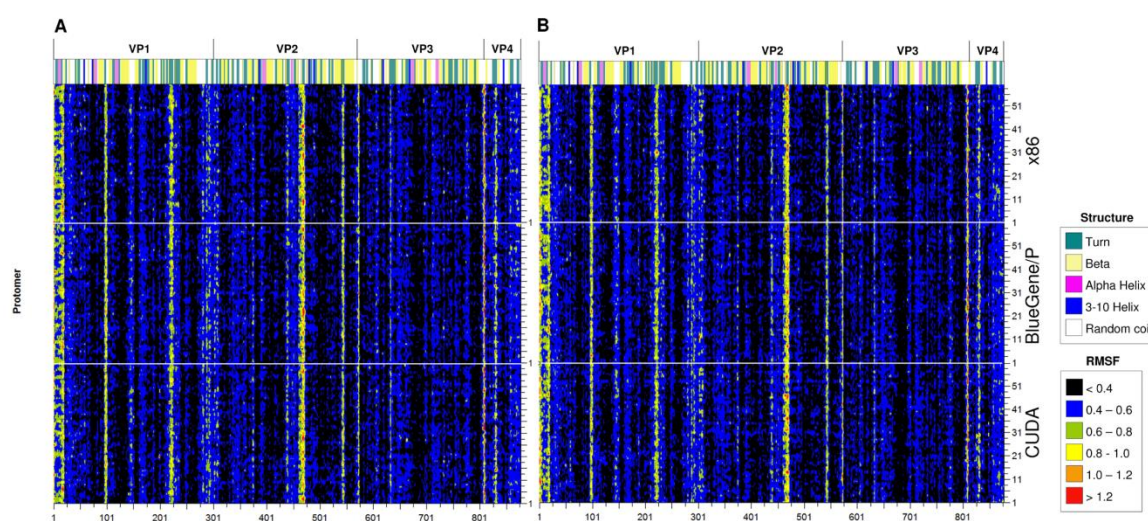


Figure 4.4 Comparison of Computer Architectures for Simulation Environment (RMSF). RMSF for each of the 60 virus capsid protomers calculated in 10ps steps for 0.1ns. y- axis, 60 protomer set for each capsid grouped by architecture: x86, BlueGene/P, CUDA. A = wild poliovirus, B = Sabin-like vaccine poliovirus. X-axis, amino acid count for a single protomer. Each graph contains a secondary structural index header created using the STRIDE algorithm for secondary structural analysis³⁴⁷. x- axis = residue number, y- axis = protomer replicate number (n=60).

4.4.3 Analysis of model stability

After 10 ns of simulation, each of the four simulations (α -helix, pentamer, empty capsid and full virion) were transferred to the local workstation for on-site analysis. To provide a gross measurement of the stability of the simulations, RMSD values for α -carbons were calculated for pentamer, empty capsid and the full virion (Figure. 4.5). It is important at this point to explain the observed phenomenon of “solvent shock”, which is a result of simulation cell size changes due to water molecules (initially in a non-equilibrium state) undergoing a sudden reorientation after the minimization step, leading into the equilibration phase of the MD simulation. The non-equilibrated water molecules are oriented differently to those in an equilibrium fluid state at 310K. Equilibration of the water molecules is normally achieved after a short (<20 ps) simulation time, resulting in an abrupt but minor change in simulation cell size and associated perturbation of capsid structure at the start of the simulation. The “solvent shock” is observed as a sharp increase in the first 20 ps of a simulation. This is followed by a modest increase in RMSD value as the simulation equilibrates; leading to a plateau phase, indicating that the

simulation is stable for the simulation time frame, with no substantial unfolding or disruption in capsid integrity (Figure 4.5).

The variation in radius of the empty capsid and the complete virion were calculated for the duration of the 10 ns simulations (Figure 4.6). The empty capsid showed an average radius of 145.5 Å compared to the full virion at 147.5 Å, representing a 1.4% difference in capsid radius between the two structures. These observations are in line with the observed diameter of enteroviruses, being in the range of 30 nm (300 Å), refer to section 1.2.2.

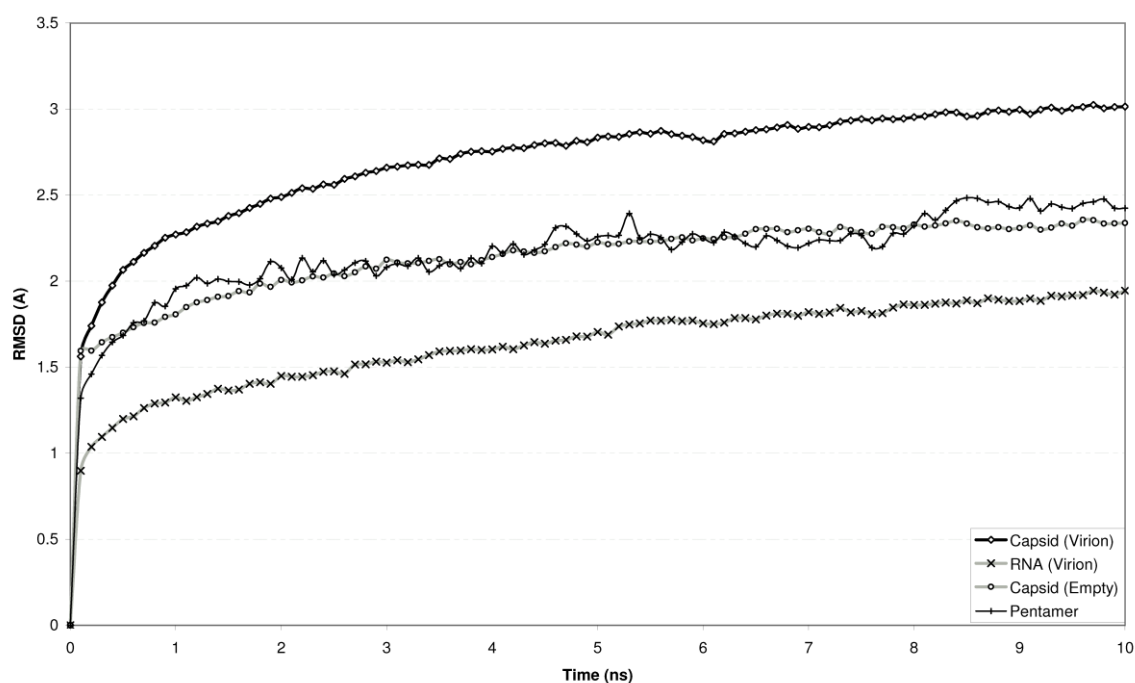


Figure 4.5 RMSD of poliovirus after 10ns atomistic MD simulation. Wild Poliovirus empty capsid, full virion, spherical RNA alone and an individual pentamer over the 10 ns simulation time. Figure reproduced in Roberts *et al.*, (2012)². *x*- axis = time (ns), *y*- axis = RMSD (Å).

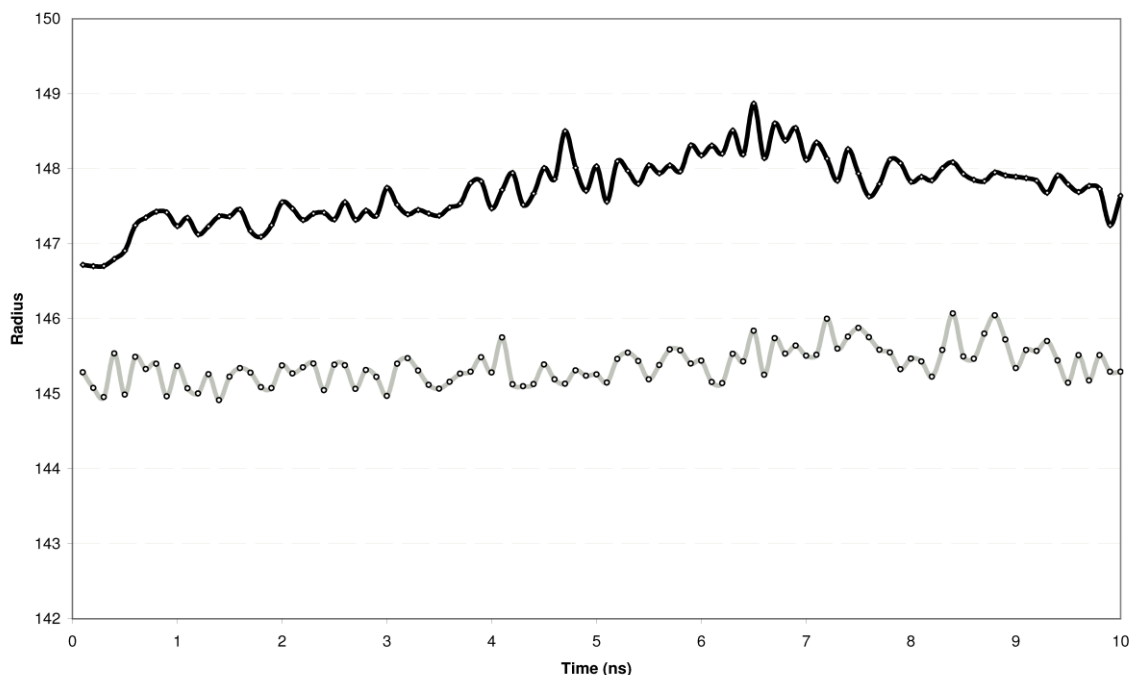


Figure 4.6 Average capsid radius (Å) over 10ns simulation time. Change in radius as measured from the outermost α -carbons of the capsid for the x, y and z axes over time Empty capsid (grey) and full virion (black). Figure reproduced in Roberts *et al.*,(2012)². x- axis = time (ns), y- axis = Radius (Å).

4.4.4 Analysis of Virus Structure after Molecular Dynamics Simulation.

The flexibility of the structural proteins VP1-4 was performed using the RMSF method over the final 2.5 ns of the initial 10 ns simulations, for empty capsid and full virion structures (Figure 4.7). Differences were noted on the five-fold axis of symmetry, encompassing the five covalently bound myristate molecules of VP4 and the interlocking N-termini of the VP3 and VP4 structural proteins. Increased flexibility was observed in the empty capsid structure compared to the full virion, indicating that flexibility in this region was reduced in the presence of the RNA core.

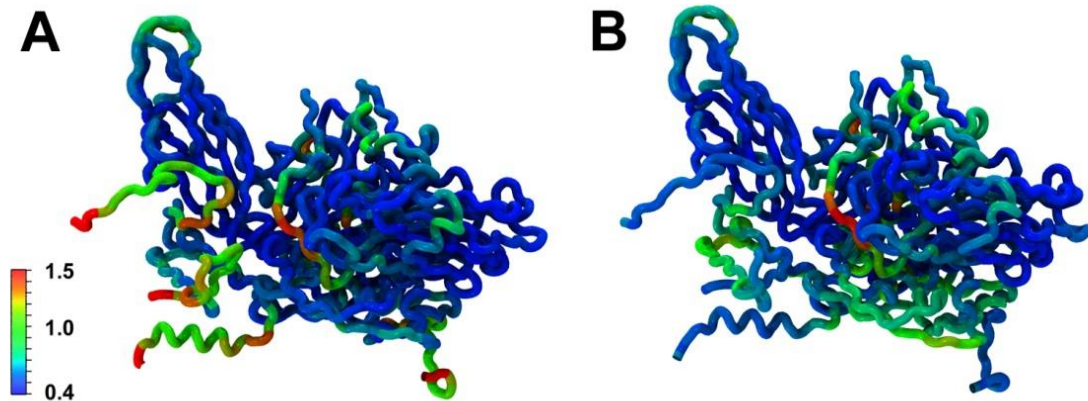


Figure 4.7 RMSF values calculated over 7.5 to 10 ns time period for all protomers in-situ A) empty capsid and B) full virion. RMSF values are indicated by the legend blue 0.42 red 1.5. Figure reproduced in Roberts *et al.*, (2012)².

Also noted was increased flexibility in the regions that define a pore-like structure which appears at the base of the canyon, at the entrance of the hydrophobic pocket. Significant positive ion (sodium) movement was observed in this region through the two-fold axis of symmetry and the pore-like structure at the base of the canyon (Figure 4.8). This suggests that ion permeation may play a role in the biological function of poliovirus. Further rigorous identification and quantitative analysis of the free energy of ion permeation and selectivity through these channels, will be of interest in future work to fully characterise their electrostatic properties and may have implications for virus-host cell interactions, as indicated by Smith *et al.*, for rhinovirus³⁴⁸.

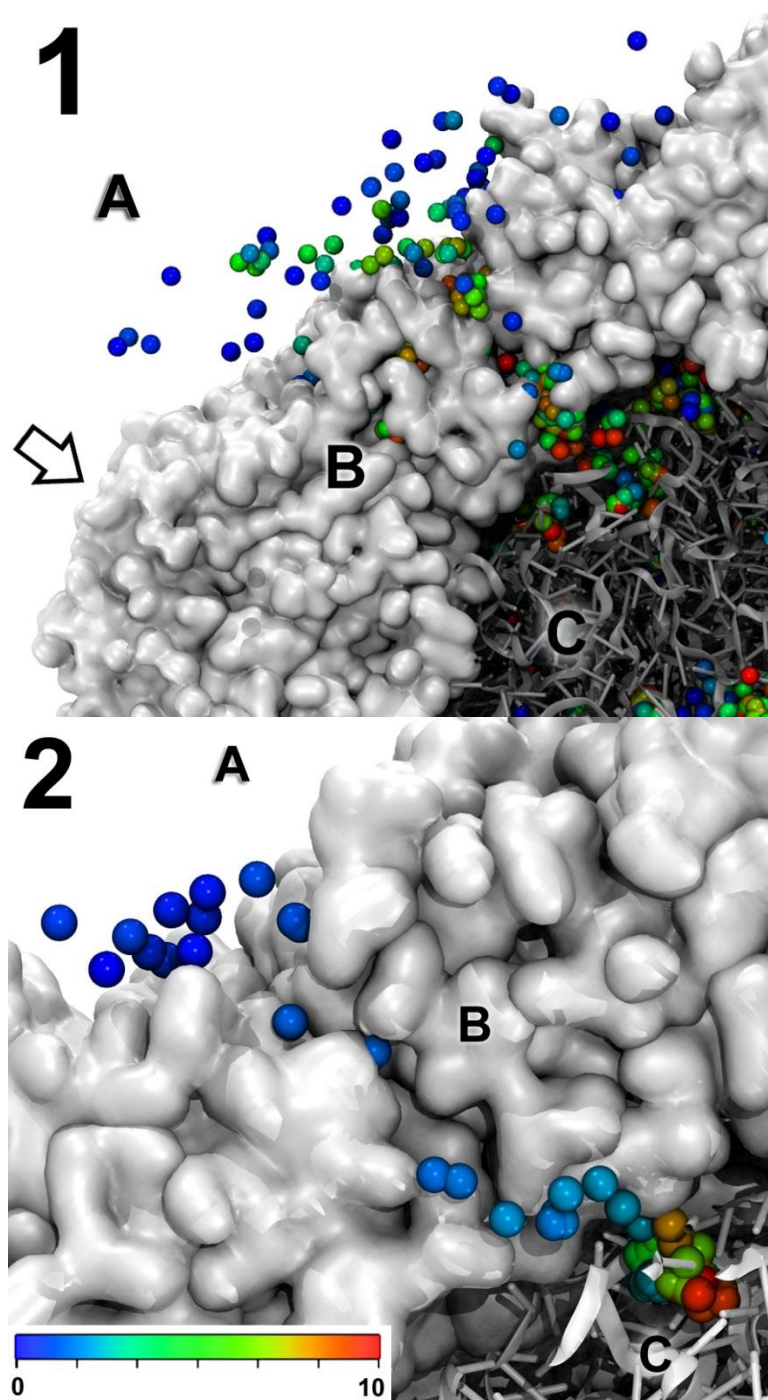


Figure 4.8 Representation of ion movement in a 10 ns equilibration of wild poliovirus virion using molecular dynamics. Panel 1 indicates multiple ions traversing the 2 fold axis of symmetry with the 5-fold axis of symmetry indicated by an arrow. Panel 2 indicates a single ion traversing the virus capsid via the pore-like structure located at the base of the 5 fold axis of symmetry. In each panel the solvent space (A) is indicated in the top left of the figure, whilst a coronal section of the virus capsid is presented (B) as a 3-dimensional surface rendering in white at a resolution of 2Å. Sodium ions are shown traversing the virus capsid from the solvent space to the interior of the capsid and interacting with the viral RNA (C), ions are coloured according to time point in 0.1ns increments. Scale bar represents ion colouring as time in ns.

At the 10 ns time point, each of the 60 individual protomers that comprised the virus capsid were transposed to the initial protomer coordinates, as defined by the 1HXS template crystal structure. These 60 protomers, which were overlaid to occupy the same space, were then averaged to give a single "fit-averaged" structure, Figure 4.9 (A Tcl script used for the fit-averaging process is shown in appendix A5.4).

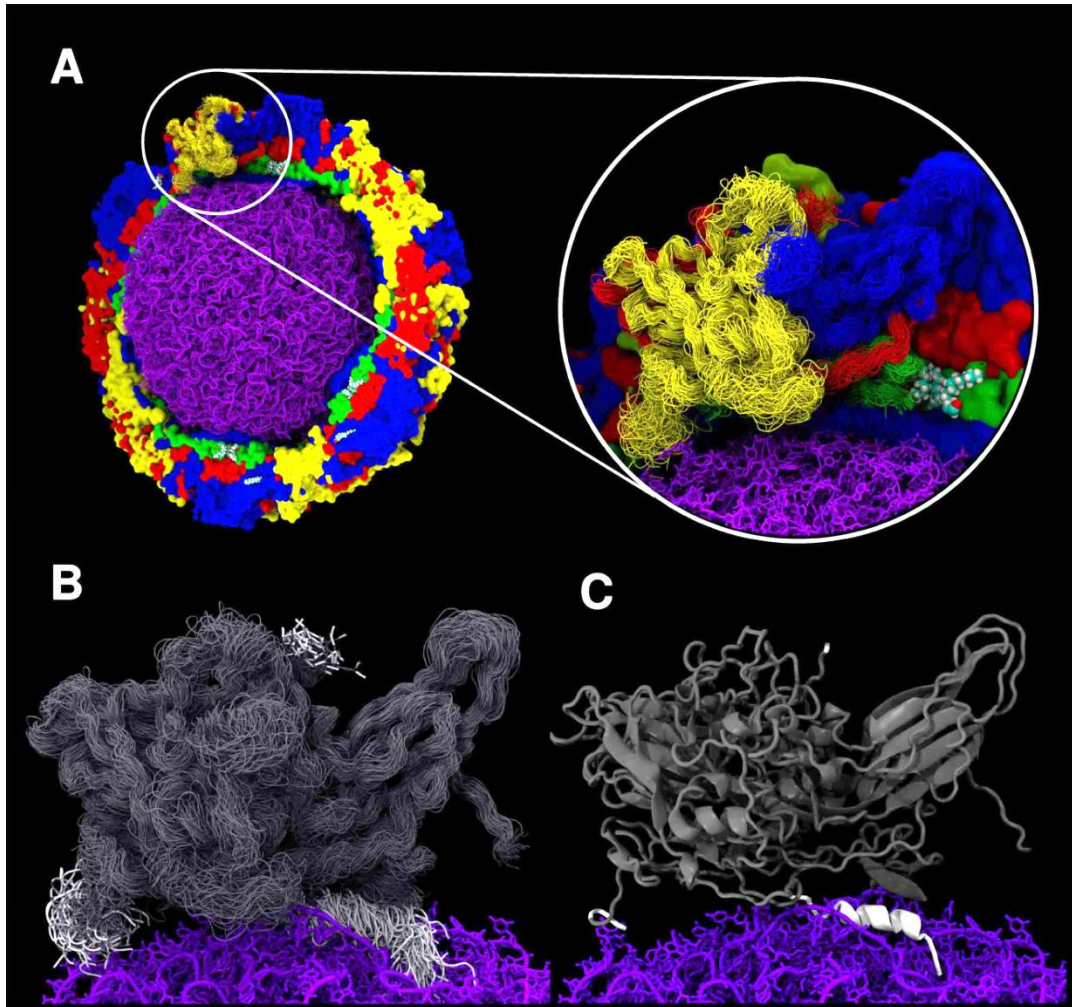


Figure 4.9 Examination of protomer structural motifs using protomer fit-averaging. A) cutaway representation of the full wild poliovirus virion highlighting relative positioning of a protomer in relation to the RNA genome (purple). The fit averaged protomer (n=60) is coloured by chain as in previous Figures and represented as a backbone structure at the 10 ns time point. B) Overlaid backbone structure of all 60 fit-averaged protomers, structures missing from original x-ray crystal data, (1HXS), are highlighted in white. C) Cartoon representation of the structure derived from the fit averaging procedure. Regions that were missing from the original x-ray diffraction data, (1HXS), are highlighted in white. Figure reproduced in Roberts *et al.*, (2012)².

The individual fit averaged protomer structures for the pentamer, empty capsid and full virion were compared to the original 1HXS template PDB file, which incorporates B-factor data. RMSD values were calculated for the VP1 protein within the 60 fit-averaged protomers of the simulated structure, and then compared to the 1HXS template B-factors for VP1 (Figure 4.10). The resulting RMSD values ranged from 0.873 Å to 1.099 Å when compared to the original template structure, 1HXS, and from 1.125 Å to 1.604 Å for ten other x-ray diffraction-derived structures for poliovirus (1AL2, 1AR6, 1AR7, 1AR8, 1AR9, 1ASJ, 1PO1, 1PO2, 1POV, 1VBD) available from the RCSB PDB (Figure 4.11).

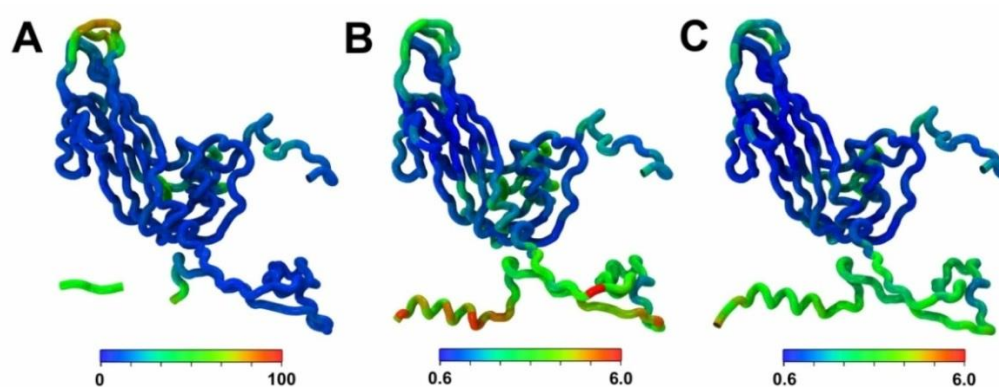


Figure 4.10 Thermal Comparison of VP1 structural proteins. A) VP1 structural protein of the 1HXS template crystal structure coloured by beta value supplied with the original coordinate data. Figures B and C are coloured "RMSD by residue" of fit averaged VP1 structural proteins (n=60) at 10 ns time point B) empty capsid and C) full virion. Figure reproduced in Roberts *et al.*, (2012)².

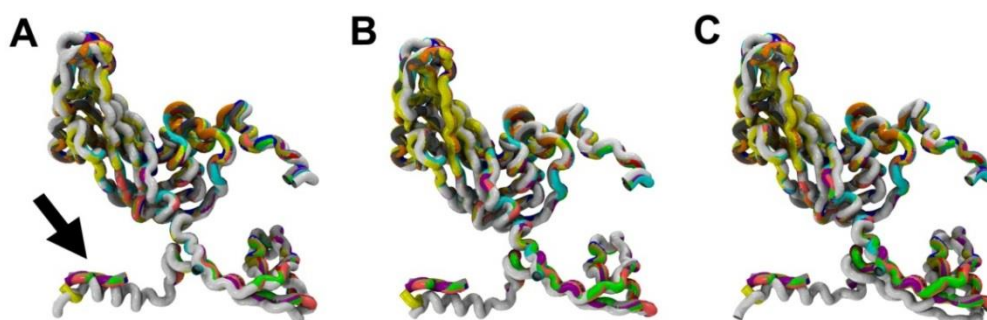


Figure 4.11 View of fit-averaged VP1 structural protein reconstructed VP1 protein (White) overlaid compared with available wild poliovirus type 1 crystal structure coordinates for VP1 (coloured) obtained from the RCSB PDB. Position of the reconstructed amphipathic α -helix (black arrow) after 10ns simulation of A) Pentamer (fit averaged protomers, n=5), B) Empty capsid (n=60) and C) Full virion (n=60). Figure reproduced in Roberts *et al.*, (2012)².

Stability of the predicted α -helix of the N-terminus of the VP1 protein was determined in isolation from the remainder of the capsid structure. This enabled the determination of whether the helix possesses intrinsic structural stability, or whether interaction with neighbouring structural components in the full virion is required to maintain helical structure. The detached protein, incorporating residues one to twenty of VP1, maintained a helical structure for the duration of the 10 ns simulation (Figure. 4.12). Residues ARG15-ASP11 and ARG15-GLU16 formed salt bridges, which were significant stabilising elements of the α -helix. The cyclic distribution of amino acid residue types is indicative of an α -helix that is amphipathic in nature and matching previous descriptions for rhinovirus³³⁴.

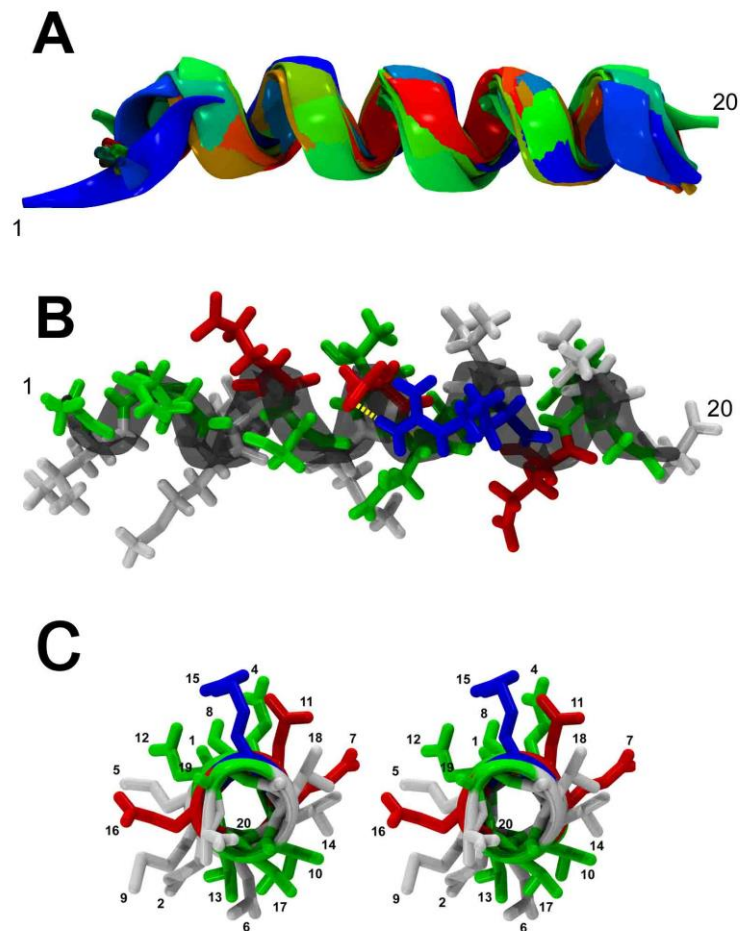


Figure 4.12 N-terminus of VP1 after 10 ns simulation. Detached α -helix from the N-terminus of VP1 after 10 ns simulation. A) Conformational indicated at 1 ns time points, (1ns = blue, 10ns = red). B) Liquorice representation of amino acid side chains, showing salt bridge between ARG and ASP (centre), (hydrophobic = white, polar = green, basic = blue, acidic =red, yellow = Hydrogen-bond. C) Stereo representation of side chain charge distribution around the amphipathic helix. Figure reproduced in Roberts *et al.*, (2012)².

4.4.5 Examination of Viral RNA Interactions with the Capsid.

Individual amino acids of all 60 capsid proteins were analysed, for their proximity to the genomic RNA of the virus model. The RNA interactions were predominately confined to the N-terminal aspects of the VP1, with the majority of interactions occurring within the first 40 amino acids and a second series of interactions occurring between amino acids 64 and 71. VP4 RNA-to-protein interactions occurred primarily in four regions distributed across the entire protein. Indications of capsid interactions for VP2 occurred primarily within the first 10 residues of the N-terminal aspect of VP2, with a second subset of interactions occurring between amino acids 35 and 55 (Figure. 4.13). There was limited interaction of VP3 with the genomic RNA, with less than 10% of residues found in proximity to the RNA for the region between amino acids 29 and 43.

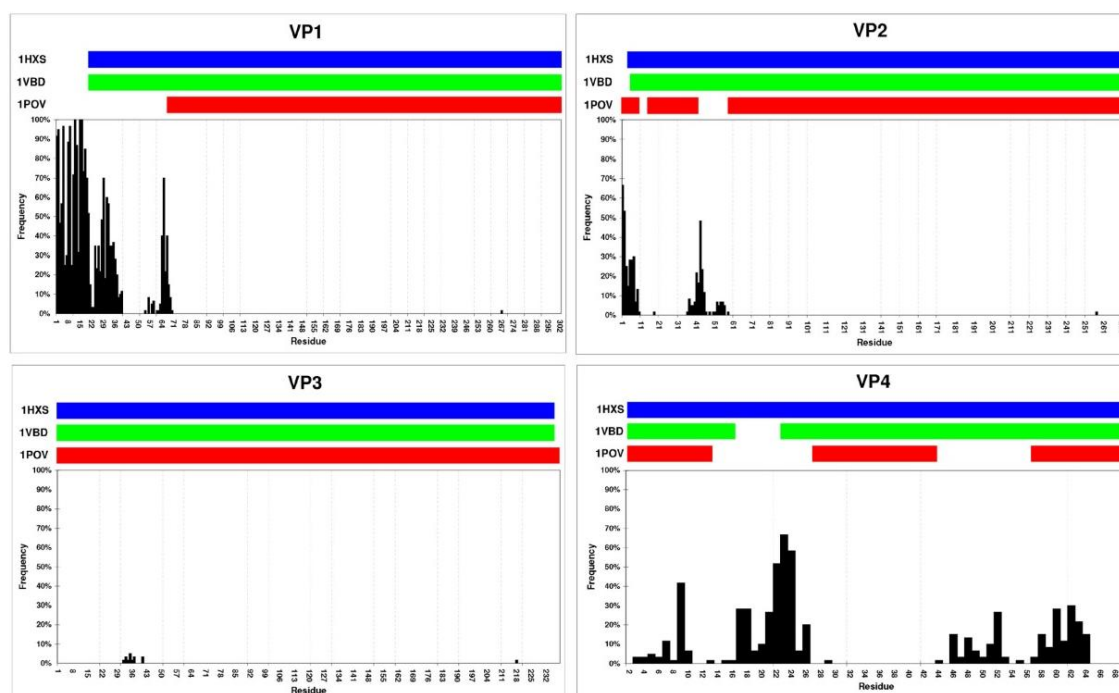


Figure 4.13 Bar graph depicting the frequency of amino acids occurring within a distance of 5 Å relative to the viral RNA for all 60 capsid protomers. Coloured bars at the top of each chart indicate available coordinate data by residue for the full virion template files 1HXS (blue) and 1VBD (green) and empty capsid 1POV (red). Figure reproduced in Roberts *et al.*, (2012)². *x*- axis = residue number, *y*- axis = Frequency % (n=60).

When comparing the RMSD values of the VP1 protein, particularly with respect to the original 1HXS template structure, it was shown that the α -helix displayed marginally higher RMSD values in the empty capsid than that of the full virion (Figure. 4.14). This would indicate that the α -helix in the empty capsid exhibits increased flexibility when the RNA is not in close proximity.

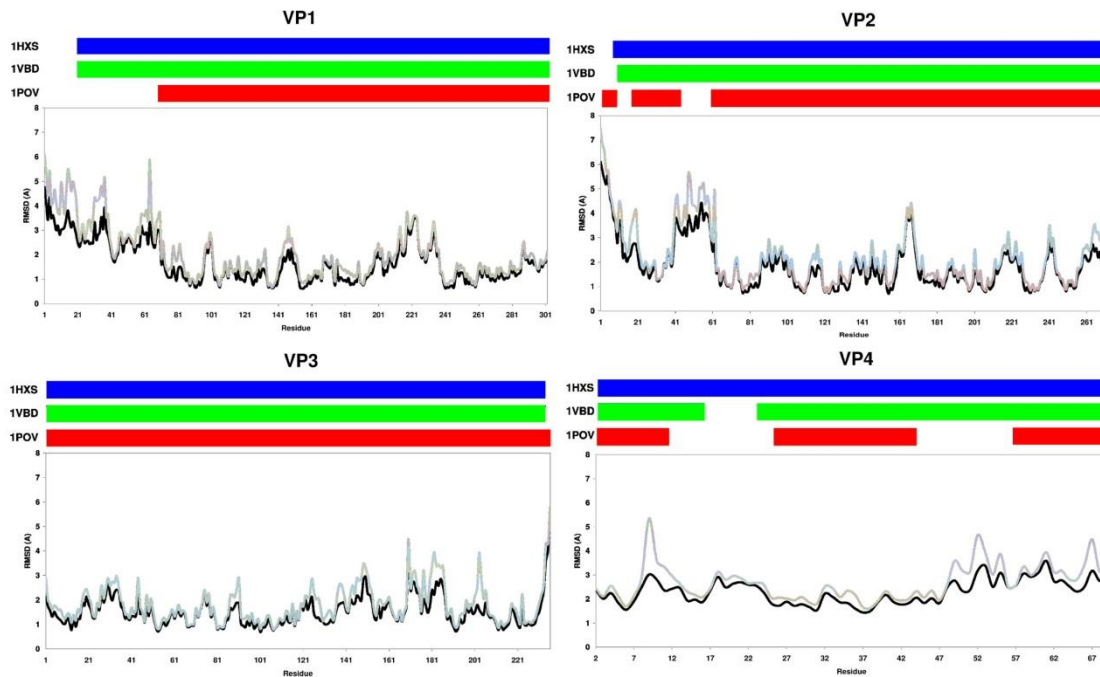


Figure 4.14 Examination of capsid chain RMSD by residue. RMSD of individual residue for fit-averaged protomers ($n=60$) at the 10 ns time point for empty capsid (grey) and full virion (black). Coloured bars at the top of each chart indicate available coordinate data by residue for the full virion template files 1HXS (blue) and 1VBD green) and empty capsid 1POV (red) x - axis = residue number, y - axis = RMSD (Å). Figure reproduced in Roberts *et al.*, (2012)².

4.4.6 Disruption of Virus Capsid Morphology.

Investigations of the thermal effects of full virus simulations on gross virus morphology indicated subtle differences at the four temperatures tested, 37°C, 4°C, 60°C and 95°C. Morphological changes can be observed when comparing the 37°C simulation to temperature reductions, notably in the reduction in size of the opening of the two-fold axis of symmetry at 4°C. At 60°C, the opening of the two-fold axis of symmetry maintains prominence and is accompanied by significant size increases for the pore-like structures present at the base of the canyon, proximal to the hydrophobic pocket of VP1. At 95°C, these two features undergo substantial rearrangements and result in a complete destabilisation of the canyon and a significant increase in the size of the opening at the two-fold axis of symmetry (Figure 4.15). These results are similar to those indicated *in-vitro* for wild poliovirus and coxsackievirus A7^{91,234,336}.

It has been suggested that capsid binding antiviral drugs inhibit the conformational shifts in virus structure required for uncoating^{349,350}. If this is the case, then it is fair to assume that examination of a drug bound in the hydrophobic pocket using MD simulation may result in the ability to observe inhibition of large-scale conformational shifts. The use of long timescale simulations (in the order of 0.1 μ s) to determine the effect of the capsid binding compound WIN51711 (Disoxaril) was performed, and movement proximal to the centre of the capsid structure was examined. Movement was measured by calculating the centre of mass for each of 60 protomers in full virion simulations, and by measuring changes in radial distance from the centre of the capsid structure. Simulations representing a virion structure with the native lipid (palmitate) occupying the hydrophobic pocket (Figure 4.16), a virion with no pocket factor, and a virion bound with Disoxaril, were performed. Coordinates for the positioning of Disoxaril were determined using the PDB coordinate data for poliovirus (PDB ID 1PIV)³⁵¹.

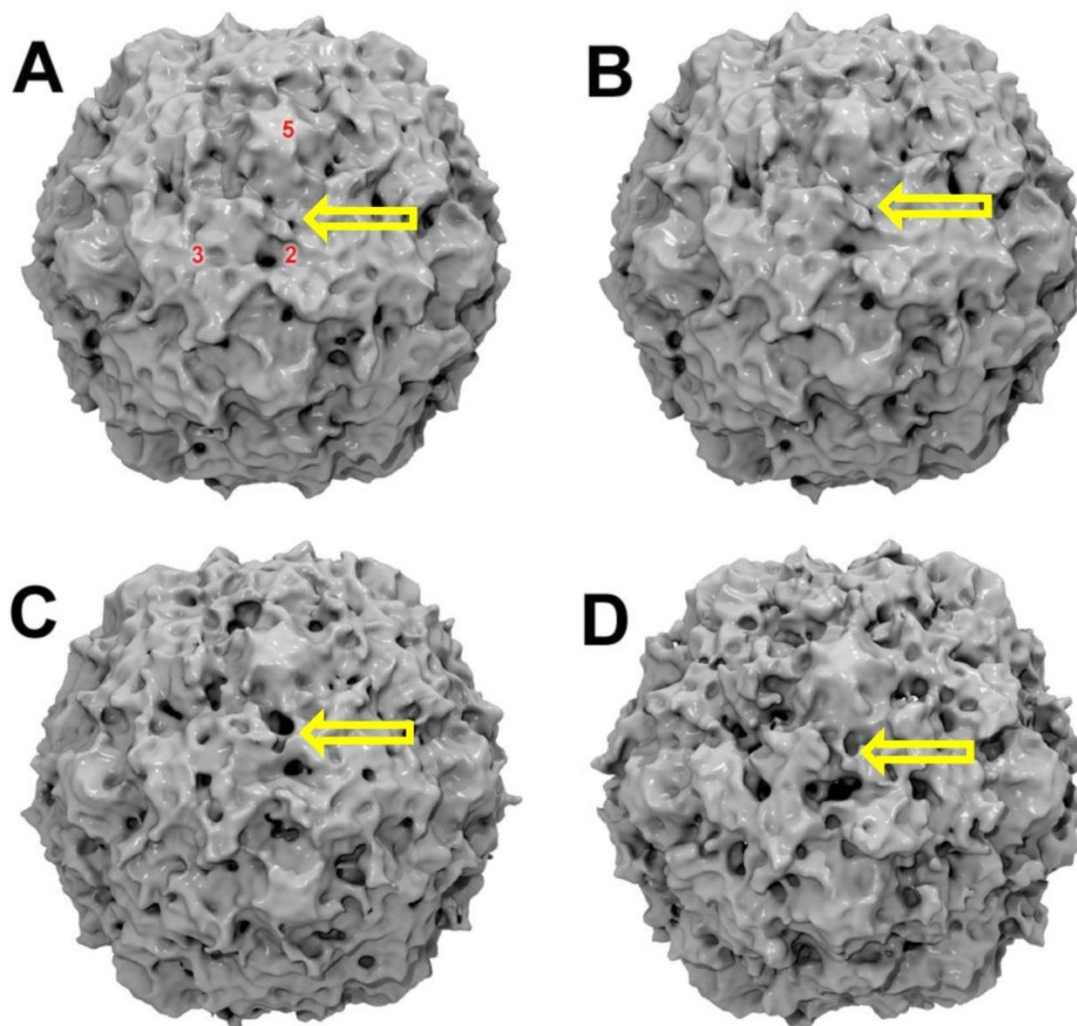


Figure 4.15. 10 ns simulations of wild poliovirus virion at different temperatures. A) 37°C after 10 ns relevant landmarks of virus symmetry are indicated by numbers 5– five-fold axis of symmetry, 2– two-fold axis symmetry, 3– pseudo three-fold axis of symmetry, Pore-like structure (Yellow arrow). B) 4°C after 10 ns, C) 60°C after 10 ns, D) 95°C after 10 ns. All representations are shown in the same orientation. Figures representing an 8 Å generated using the tachyon renderer in VMD.

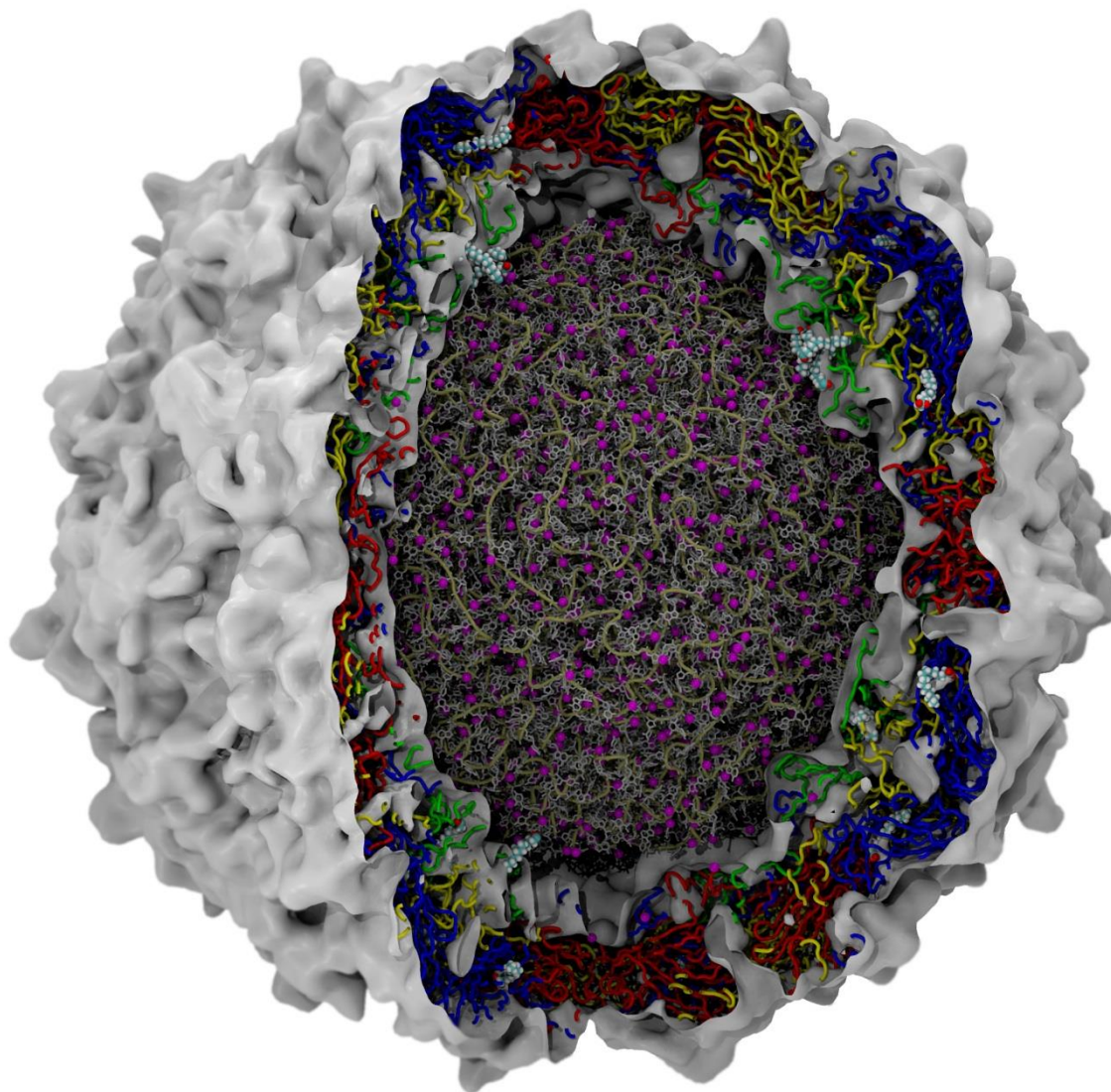


Figure 4.16 Wild Poliovirus type 1 (Mahoney strain) after 0.1 μ s of atomistic MD simulation. Total simulation size approximately 3.3 million atoms. Water and surrounding ions have been removed from this representation. Volumetric (density) map representation is represented (white) with protein backbone shown as tubes and coloured by subunit VP1 = Blue, VP2 = Yellow, VP3 = Red, VP4 = Green. Lipids (palmitate and myristate) are represented as Van der Waals models with colouring by name of element. RNA is shown using a “liquorice” representation (grey- side chains, tan- backbone) with associated magnesium ions (Purple).

The MD simulations for the capsid bound with palmitate, Disoxaril and no pocket factor over the 0.1 μ s timescale, indicated identical RMSD values for the entire simulation (Figure 4.17-A). A pattern of virion diameter changes were observed in the simulations containing palmitate and no pocket factor (Figure 4.17-B, C), whereas the simulation with Disoxaril bound to the hydrophobic pocket indicated less variation in capsid diameter (Figure 4.17-D), being particularly evident during the initial recovery from the “solvent shock” phase of the MD simulation. The simulations incorporating no anti-viral drug required in excess of 50 ns to equilibrate. The simulation involving bound Disoxaril showed little evidence of change in the first 30 ns simulation, with a gradual, but slight variation in capsid diameter. The final visual representation of the complete virion containing palmitate after 0.1 μ s atomistic MD simulation appears in Figure 4.16.

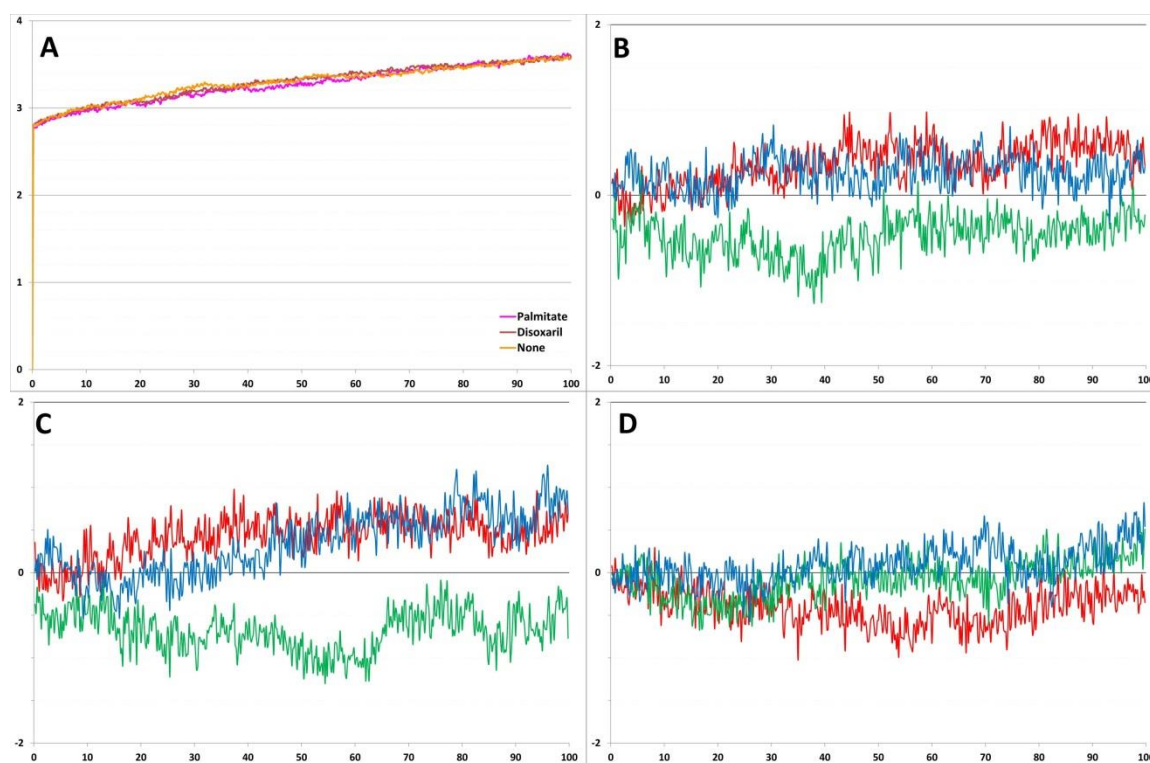


Figure 4.17 RMSD and virus capsid diameter variation for simulations incorporating different pocket factors. A) RMSD calculated over a 0.1 μ s timescale for protein backbone of capsid structures with no pocket factor (orange), palmitate (magenta) and Disoxaril (Brown). Percent (%) variation in virus capsid diameter for the x , y and z axes calculated in relation to time point 0 of a series of 0.1 μ s simulations, B) no pocket factor, C) palmitate and D) Disoxaril. x -axis (blue), y -axis (green) and z axis (red). The x - axis for graphs A, B, C and D are measured in nanoseconds, the y - axis for graph A is RMSD, and the y - axis in graphs B, C and D are in percent (%) variation.

4.5 Discussion

The computational power available from supercomputers now enables MD simulations of multi-million atom systems, facilitating the *in-silico* investigation of macromolecular structures including viruses². Utilising MD software to perform atomistic simulations of a complete wild poliovirus including the native RNA sequence, it was possible to successfully simulate the virus for a period of up to 0.1 μ s and approximately 4.2 million atoms. The x-ray crystallographic co-ordinates of wild poliovirus type-1 published in 1985¹⁶ and further refined in 2001³³⁷, were used to reconstruct the molecular structure and the missing virus protein polypeptide sequences were reconstituted according to the published prototype nucleic acid sequence (GenBank NC002058).

This virus was chosen due to the abundance of x-ray crystallography and biochemical analyses published over more than thirty years. It was noted that specific regions of the virus internal structure could not be resolved by x-ray crystallography, specifically the N-terminus of VP1 and the RNA genome. Although the virus capsid exhibits icosahedral symmetry, which allows for the refinement of protomeric subunits, the underlying structure of the genomic RNA has not been determined; most likely due to the relatively amorphous nature of the virus core. Although initially the intention was to only simulate the virus capsid structure, an attempt was made to incorporate the RNA in this structure to better reproduce an infectious particle. The resulting complete virus structure simulation represents the most complete and published atomistic MD simulation of a viral pathogen to date (Figure 4.16)².

The advantage of full virus simulations, despite the significant increase in the total atom count (up to 4.2 million atoms), compared to the simulation of a single protomeric subunit (up to 0.2 million atoms), is that an economy of scale comes into play with 60 protomers available in the full virus (or capsid) simulations as independent entities within a biologically relevant assembly. In order to replicate the data obtained from a complete virion or capsid simulation, 60 separate protomer simulations would be required and even then the molecular interactions from neighbouring protomers and the native RNA would not be accounted for in an independent manner. This would be the equivalent of running sixty 0.2 million atom simulations, requiring the same computational resources as a single 12 million atom simulation. Clearly, the simulation of a complete 60 protomer structure is an economic proposition, which becomes more compelling when incorporating

rhombic-dodecahedral periodic boundary conditions, reducing the total simulation size to approximately 3.3 million atoms, representing a saving of over 70% of the computational resources required to run 60 simulations of a single protomer.

Validation of the atomistic MD simulation of a complete enterovirus was performed by constructing wild poliovirus type 1 (Mahoney strain) and by comparing structural elements with published atomic co-ordinate data. An additional advantage of simulating the entire virion was the ability to avoid potential artefacts that may have arisen by simulating only a portion of the capsid through symmetry. This could be especially relevant if there is an event that breaks capsid symmetry such as the gross morphological changes observed in the virus capsid prior to RNA release^{91,336}. Inclusion of the prototype viral RNA genome may also serve to assist researchers studying RNA-capsid interactions, viral uncoating mechanisms and subsequent RNA release.

The computational experiments involving the simulation of full viruses and associated individual components (pentamer, empty capsid, spherical RNA) performed in this chapter, indicated that protocols so established, had resulted in simulations that were stable when assessed using RMSD measurements. Hypotheses exist that indicate an N-terminal α -helix located on VP1 of poliovirus type 1; this structure was also hypothesised to be amphipathic in nature based on the periodicity of the amino acid sequence⁸⁹. In the MD simulation performed in this chapter, the putative amphipathic helical structure was intrinsically stable when the final 20 amino acid residues were simulated as an independent entity. When examining the structure using a fit-averaged approach, the 60 protomer units indicated an increase in flexibility of the N-terminal region of VP1. This was particularly evident when simulated as a part of an empty capsid structure compared to the full virion with native RNA.

When using complete capsid or virion structures for MD simulation, the individual hydrophobic pockets represent 60 independent drug binding sites, incorporated in a single simulation. In terms of statistical and biological relevance, this is important for future experiments that examine antiviral drug-binding. The core of the five-fold axis of symmetry proximal to the RNA indicated variations in flexibility when compared to empty capsids, as shown in Figures 4.7 and 4.10.

With respect to the observed interactions of the genomic RNA with the inner surface of the virus capsid, the N-terminal aspects of VP1 and VP2 showed significant instances of

contact with the genomic RNA (Figure 4.13). Observations involving the interaction of the N-terminal aspect of VP2 with genomic RNA have been previously observed⁹¹, but the significance of these interactions is unknown. With respect to the interactions of the N-terminal aspect of VP1 with the genomic RNA, it is not surprising that a large internal structure such as an amphipathic helix would make contact with the RNA. This structure has not been observed in poliovirus using *in-vitro* experimentation, most likely due to mobile nature of the helix moving within the core of the virus, or exposed on the external surface of the capsid during virus breathing. It can be assumed that such mobility is critical to virus function.

The N-terminus of VP1 is reversibly displayed on the exterior of the virus capsid through virus breathing, a feat that would necessitate molecular flexibility as indicated in the simulations performed in this chapter. It is hypothesised that binding of poliovirus to the cellular receptor, CD155, results in the irreversible externalization of the N-terminus of VP1. This is thought to occur during the process of cell membrane penetration, involving the formation of a pore-like complex, thus facilitating egress of the genomic RNA^{89,90}.

The orientation of the N-terminus of VP1 on the internal surface of the capsid prior to receptor binding, has been the subject of much speculation^{89,234,334}. The amphipathic helix of the N-terminus of VP1 of human rhinovirus 16 has been reported to be located on the interior surface of the capsid along the icosahedral five-fold axis, the orientation used in simulations in this chapter³³⁴. Recent published data of EV-A71 indicate the positioning of the VP1 α -helix as being oriented toward the 2 fold axis of symmetry²¹⁵. Experiments involving the use of antibodies directed to the N-terminus of the VP1 in poliovirus were performed using cryo-EM analysis and found to be orientated towards the two-fold axis of symmetry³⁵². An advantage of the atomistic models is that components, such as the VP1 N-terminal helix, can be repositioned to account for new data, or to investigate alternative theories and compared as new MD simulations.

Enteroviruses, in particular polioviruses, are relatively stable under a number of conditions that would result in inactivation of other viruses; an example is exposure to ethanol which does not denature or destroy the proteinaceous virus capsid as it does with the lipid containing envelope of other viruses, such as influenza. One way that polioviruses can be inactivated is by exposure to elevated temperatures. Exposure to increases in temperature has been observed to result in the expulsion of RNA across the 2-fold axis of symmetry^{91,336}. The *in-silico* experimentation described in this chapter

involved the gradual increase of simulated temperature on poliovirus virions which yielded evidence of significant morphological changes. Gross morphological distortion occurred in the canyon of the virus at 60°C and 95°C (Figure 4.17) and an opening of the 2-fold axis of symmetry was observed.

As a result of these observations, it was hypothesised that subtle conformational shifts may be observed during the initial “solvent shock” and equilibration phase of a virion structure undergoing MD simulation. To examine if this effect was observable, and affected by the binding of an antiviral drug, a series of experiments were performed involving the presence or absence of a native lipid compound or antiviral drug WIN51711 (Disoxaril). Observations of capsid deformation in the form of expansion and contraction during “solvent shock” and equilibration, were observed in simulations that contained either native lipid or no small molecule in the hydrophobic pocket. No significant alteration in capsid dimensions in response to “solvent shock” was observed in the *in-silico* experiment that contained the antiviral drug WIN51711, indicating that the bound state of the drug may interfere with the rigidity of the capsid, affecting the viruses’ ability to respond to changes in environmental conditions.

A number of preliminary semi-quantitative and qualitative analyses and observations were performed on the simulation trajectories obtained, including the overall “shape” of poliovirus capsids, by monitoring the time-series evolution of its radius, averaged conformations of protomers, capsid-RNA contacts, and secondary structure of previously unresolved N-terminus helix. Further examination of the specific interactions of the RNA with the internal surface of the capsid is warranted. In addition to the examination of the virion structure itself, it is worth noting that MD trajectories derived from these simulations also produce vital information on the interactions between the virus, solvent and ions. Of particular interest is the potential existence of ion channels throughout the capsid, which may serve important functions relating to virion-host cell interactions. These pores have been described previously in rhinovirus capsid structures determined by x-ray diffraction methods, but not poliovirus^{15,348,353}. While the work presented in this thesis identified that the pores may be cation-selective (sodium-ions), future work could focus on extracting quantitative free-energy profiles of ion permeation through these pores.

This chapter has demonstrated a range of preliminary, simple analyses on virion MD trajectories, forming a platform for further work with poliovirus and other members of the genus *Enterovirus* (refer to chapter 5). It is anticipated that new analysis methods and protocols will be developed to keep pace with the increasing sizes of biological systems which may be simulated. Below, three potential properties of virions which merit further investigation are listed.

1. Identify solvent/ion permeable pores around virus capsid and examine their activity using methods involving the extraction of potentials of mean force (PMF) profiles for cations and anions (including complex ions). Such methods have been used in the past to examine potassium, calcium and chloride ion movement through ion channels.^{354,355}
2. Employ statistical techniques to quantify and identify concerted virion motions. Such motions are analogous to vibration motions of small molecules, and low- or high-pass filtering techniques could be used to extract large-scale (slow) motions and small-scale (localised, rapid) motions, both of which may have implications for virus function.
3. Analyse capsid structural activity in a manner that characterises the concerted motions of large segments (“plates”, representing protomers or pentamers) of the capsid, analogous to the study of planetary tectonic plate motions. Studies of these phenomena in viruses have been performed in the past using experimentally derived data^{356,357}. Examination of MD trajectory data in a similar manner would prove useful in the elucidation of the underlying molecular mechanics of these motions.

MD trajectory data collected at an appropriate time-frame resolution (1 ns per frame or less) would allow the investigation of these properties, although a side effect of this process is a larger data set for analysis, beginning at ~40GB of trajectory data for a 0.1 μ s simulation collected at 1 ns time points. Given the sheer size and complexity of virions and the associated data files generated by MD simulation, it will be necessary to develop novel methods to enable efficient use of supercomputing architectures, to fully “mine” the wealth of biophysical information embedded within all-atom MD simulation trajectories. In this chapter, an initial approach to these complex problems is proposed via a protocol for the production and preliminary analysis of stable all-atom MD simulations of poliovirus. These protocols were developed and validated against crystallographic data, laying the foundation for the modelling of other viruses exhibiting icosahedral symmetry, as discussed in the following chapter.

Chapter Five

Comparative Protein Modelling of Enteroviruses

5.1 Manuscripts Published in Association with this Chapter

- **Roberts, J.A.**, Bowden, D.S., Thorley, B. R., Reville, P.A., 2014. Chapter: *New Technologies for Viral Diagnosis and Detection, Foodborne Viral Pathogens*, Hansman, G., White, P., (Eds.), CRC Press, 1st Ed. **ISBN-10:** 1466579501. [*Inpress*].

5.2 Introduction

The methodologies developed for successful *in-silico* reconstruction and atomistic MD simulation of poliovirus type 1, as described in the preceding chapter, would be amenable to other members of the genus *Enterovirus* that share fundamental structural characteristics. In order to take advantage of comparative protein modelling methods in conjunction with MD simulation of novel virus structures, adequate atomic coordinates must be available to be used as suitable templates for the homologous virus sequence. Given that unknown structural components will be modelled on known structural templates, it is important that the coordinate data are derived by reliable methods such as X-ray crystallography, nuclear magnetic resonance or high resolution cryo-EM.

If there are no coordinate data available to exactly match the sequence of the virus under investigation, then the use of comparative modelling methods using structures that match with a high enough degree of sequence identity will be required. This assumes that there are proteins structures available in the Protein Data Bank with a suitable degree of amino acid sequence similarity. But what is an acceptable limit for sequence similarity, which may permit the accurate reconstruction of virus components? The answer is dependent on the type of protein to be modelled, and whether they are structural proteins, such as those found in the enteroviruses with known highly conserved extended-beta structures, or globular non-structural proteins such as proteases and polymerases, where the requirement for a high degree of sequence identity may vary.

As a general rule, an amino acid sequence identity of greater than 30% is considered acceptable when selecting structural templates for comparative protein modelling³⁵⁸, although others have stated that a minimum of 50% sequence identity is required with less than 1 Å deviation from backbone atoms³⁵⁹. There have also been indications that the minimum percentage identity acceptable for some models can be as low as 20%³⁶⁰.

Either way, it appears that through common usage, the figure of 30% identity has been considered by many as representative of the lower limit for the successful reconstruction of structural models using comparative modelling methods. Whilst these guidelines for sequence identity may be appropriate for the modelling of small structures, no substantial work has been performed to determine if these limits are appropriate for the *in-silico* prediction of extremely large macromolecular structures such as viruses.

As described in chapter 1 sections 1.4 and 1.5, a significant proportion of the molecular modelling of virus structures is devoted to the study of antiviral interactions with specific virus features, with a large focus being the development or examination of antivirals for the treatment of influenza infection and HIV.

It is important to understand the limits associated with sequence identity and model accuracy. Assumptions based on a flawed model can lead to a significant waste of resources. This is especially true, when considering that modelling a full virion through a 0.1 microsecond simulation, can from personal experience, take in excess of two million supercomputer hours to perform and a week in “hands-on” time to construct and analyse the simulation.

This chapter investigates issues associated with the reconstruction of structural proteins representing enterovirus capsids, using comparative protein modelling methods with templates of varying degrees of sequence identity. This was achieved by selecting prototype species C and species A *Enteroviruses* that have structural co-ordinate data derived using x-ray diffraction methods. Comparative protein models were created using two commonly used modelling software packages: Modeller^{326,361} and Swiss-PDB viewer^{327,362}. Both packages were selected due to their ease of use and for their capability to allow manual input during the threading process enabling combination of query sequence data with template structural data.

Once the most accurate and precise modelling method was established using manual threading methods, the reconstruction of the species C, EV-C96 and the species A, EV-A120 capsids, were performed. All capsids constructed were examined for stability, using the methods described in chapter 4. The stability of the model virions during the simulation time frames, strongly support the validity of the comparative modelling approach employed.

5.3 Materials and Methods

Sequences used in the reconstruction of prototype, newly described and novel enterovirus capsids were derived from GenBank, while novel species A and newly described species C *Enterovirus* sequences, were derived in-house. The species C *Enterovirus* prototype selected for reconstruction was coxsackievirus A21 (GenBank sequence accession number AF546702)³⁶³. The species A *Enterovirus* prototype selected for reconstruction was enterovirus A71 (GenBank sequence accession number FJ461781)¹⁶⁴.

Comparative protein modelling methods for the reconstruction of a single protomer used in the reconstruction of a complete capsid, were performed using manual threading methods in the programs Modeller and Swiss-PDB viewer. The models were constructed using the recommended protocols for each program, with five protomer homology models generated by default using Modeller³⁶¹, and a single protomer model generated by Swiss-PDB viewer³⁶². In the case of the Swiss-PDB viewer generated homology model, the final step involved uploading the predicted protomer to the SWISS-MODEL website for refinement.

After generation of the final protomer structures, virus capsids were reconstructed, accessory lipids added and a final structure solvated and ionised, using the methods described in chapter 4. Scripts used for the extraction and compilation of data for analysis were the same as those used for the examination of wild poliovirus as described in chapter 4 and as set out in appendix A5.

Capsid structures were solvated and ionised, then subjected to atomistic MD simulation using NAMD on a BlueGene/Q supercomputer housed at the VLSCI. Simulations were performed using a constant number, pressure and temperature (NPT) environment maintained at 37°C and one atmosphere for 10ns each, as described in chapter 4 (refer to appendix A5.8). These simulation parameters were selected in order to allow the predicted structures to equilibrate, providing opportunity for any incorrectly placed side-chains to rotate into a more energetically favourable position. Once equilibrated, the virus capsids were deconstructed using a Tcl script (appendix A8.4), which placed all protomers in the same 3-dimensional space aligned by their structural motifs, specifically the residues representing the beta-strand formations comprising the Swiss roll motifs.

Model quality assessment was performed using RMSD comparison of backbone (C, C α , and N atoms) or all atoms (not including hydrogen “NoH”) for the original template, compared with the predicted model where indicated. Additional model assessment was performed using the web-based “Protein Quality Prediction” ProQ³⁶⁴ method and two values obtained: LGscore, which determines the longest non-contiguous segment and measures the similarity using the structural P-value and MaxSub, which calculates the largest number of residues found with distances between structures shorter than 3.5 Å^{328,329}. Ranking of LGscore is divided into the following qualitative interpretations: >1 = fairly good model, >2.5 = very good model, >4 = extremely good model. Max Sub values are ranked as follows: >0.1 = fairly good model, >0.5 = very good model, >0.8 = extremely good model.

5.4 Results

5.4.1 Comparative Modelling of Complete Virus Capsid Structures

In order to examine the limitations of comparative modelling methods for the reconstruction of novel or newly described virus capsids, a series of experiments were performed to determine the amount of variation compared to a predetermined series of structures elucidated using x-ray crystallographic methods. The aim of these experiments was to determine the most appropriate methodology to accurately and precisely reconstruct prototype virus capsids, and then to apply these methods to the reconstruction of the newly described EV-C96 and the novel enterovirus EV-A120.

5.4.1.1 Coxsackievirus A21

To validate a methodology for the reconstruction of EV-C96 using comparative protein modelling, first CV-A21 was reconstructed based on wild poliovirus crystal co-ordinates and then compared to actual CV-A21 crystallography data. The reconstruction of CV-A21 was performed using automated methods via the SWISS-Model website, with a manual template selection method forcing the use of PDB file 1HXS (the Mahoney strain of poliovirus 1HXS at 2.2 Å). This template was used previously in chapter 4 and contains the most complete x-ray crystallographic data of poliovirus. Amino acid sequence identity between CV-A21 and poliovirus capsid proteins ranged from 57.4% to 76.8%, with an overall identity of 68.1% for the entire protomer structure (Table 5.1). Manual threading methods were attempted using the Swiss-PDB viewer program version 4.04 and the command-line interface-based program Modeller.

CV-A21	VP1	VP2	VP3	VP4	Protomer
PV1-Wild	57.4%	73.1%	73.9%	76.8%	68.1%

Table 5.1 CV-A21 protomer amino acid percentage identity to wild poliovirus type 1 template structure. Comparison of CV-A21 prototype sequence AF546702 to sequence data derived from the PDB structure file 1HXS for wild poliovirus.

The use of the program Modeller involved the alignment of the query CV-A21 sequence data with the PDB file data for wild poliovirus. The output consisted of five structures (PDB coordinate files representing the homology models), were all five structures bore

little resemblance to the x-ray diffraction-derived structure for CV-A21 with RMSD values approaching 5 Å (Figure 5.1A and Table 5.2). Using the protomer averaging method based on the five models, resulted in a greatly improved structure that was within 4 Å of the known coordinates for the CV-A21 virus capsid (Figure 5.1B and Table 5.2). Methods using manual threading via the Swiss-PDB viewer program proved to be relatively straightforward and required a manual alignment process followed by uploading of the data to the SWISS-model server for completion. Resultant data were within 2.1 Å of the experimentally derived VP1 structural data (PDB file 1Z7S)³⁶⁵ for CV-A21 using x-ray diffraction methods (Figure 5.1C and Table 5.2).

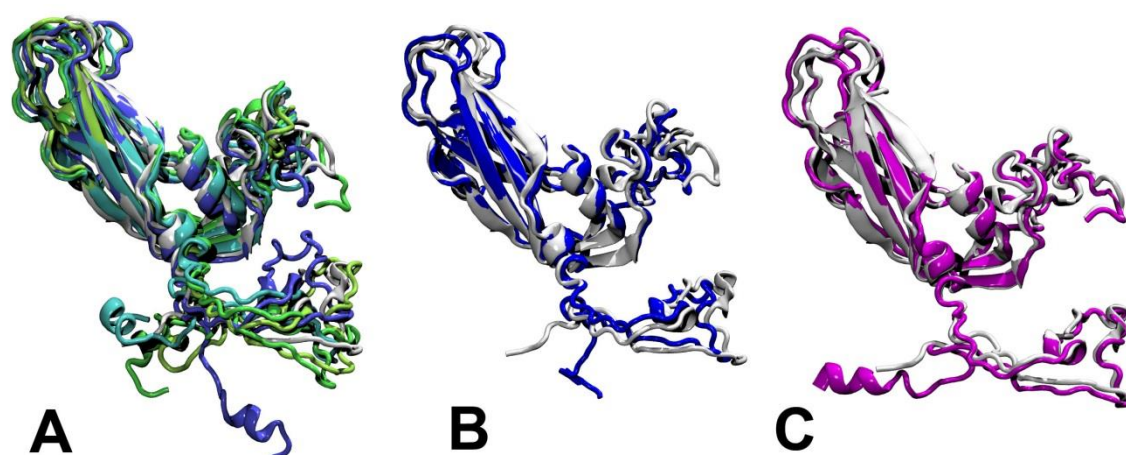


Figure 5.1 Predicted models of coxsackievirus A21 VP1 protein. VP1 protein predictions for CV-A21 overlaid on the original x-ray crystal structure 1Z7S (white), A) five outputs from the Modeller program (coloured green, cyan, violet, yellow and grey), B) fit averaged outputs from Modeller (blue), C) output from Swiss-PDB viewer in manual threading mode (purple). RMSD values, compared to original crystal structure are given in table 5.1 below.

	Modeller No. 1	Modeller No. 2	Modeller No. 3	Modeller No. 4	Modeller No. 5	Modeller Averaged	SWISS-PDB viewer
RMSD	3.382	3.983	4.658	3.281	3.382	2.593	2.097

Table 5.2 Backbone RMSD values for predicted models of VP1 for CA21 using comparative modelling methods. Values given for five outputs from the modeller program, protomer fit averaged output from modeller, protomer fit averaged model from Swiss PDB viewer project.

The model from SWISS-MODEL was used for subsequent MD simulation. After 20 ns of atomistic MD simulation using NAMD (Figure 5.2) with a constant number of atoms,

pressure and temperature (NPT), the stability of the capsid structures were examined using RMSD calculations for the entire protein (Figure 5.3). Virus expansion and contraction was determined using measurements of the virus capsid across the x, y and z axes for the α -carbons, representing the furthest loop structures which are subjected to variability in the solvent accessible space. Virus diameter measurements were also made to the extended beta sheets located towards the centre of mass of each protomer on the x, y and z axes. The β -sheets represent the major structural components of the virus capsid and are therefore not expected to undergo substantial variations in 3-dimensional space when compared to the solvent accessible α -carbons of the external surface of the capsid (Figure 5.4).

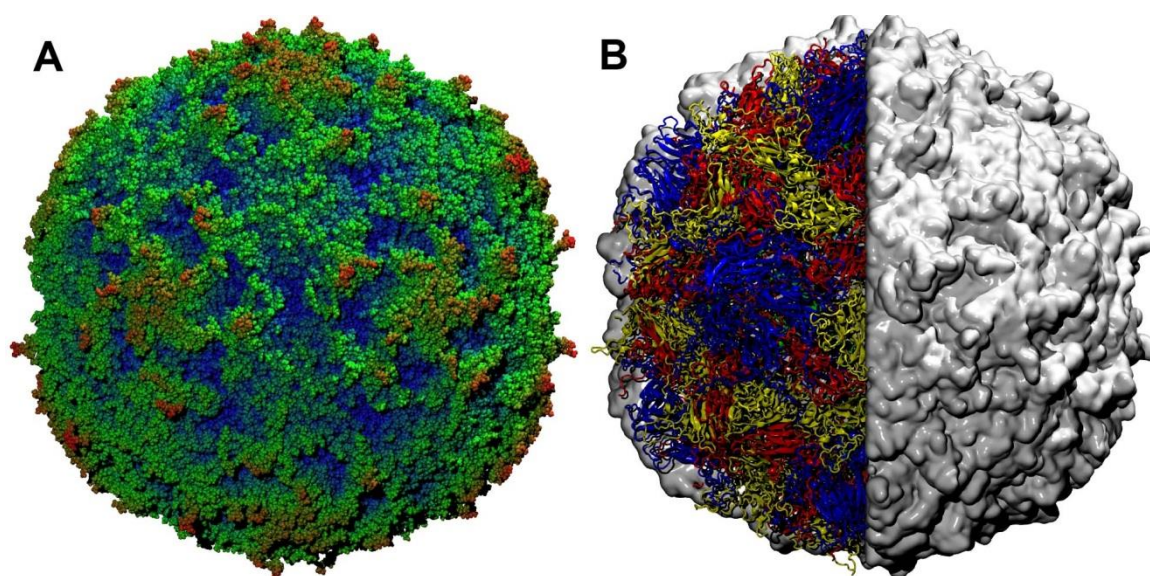


Figure 5.2 Final structure output after 20 ns of atomistic MD simulation for CV-A21. A) Radial depth shaded Van der Waals representation of the CV-A21 capsid measured from the origin (blue = 133 Å to Red 166 Å), B) cutaway density model calculated at 3.0Å, the +x and -z axes. -x and +y axes are represented in new cartoon format and shaded by protein chain VP1 (blue), VP2 (yellow), VP3 (red) and VP4 (not visible in this representation).

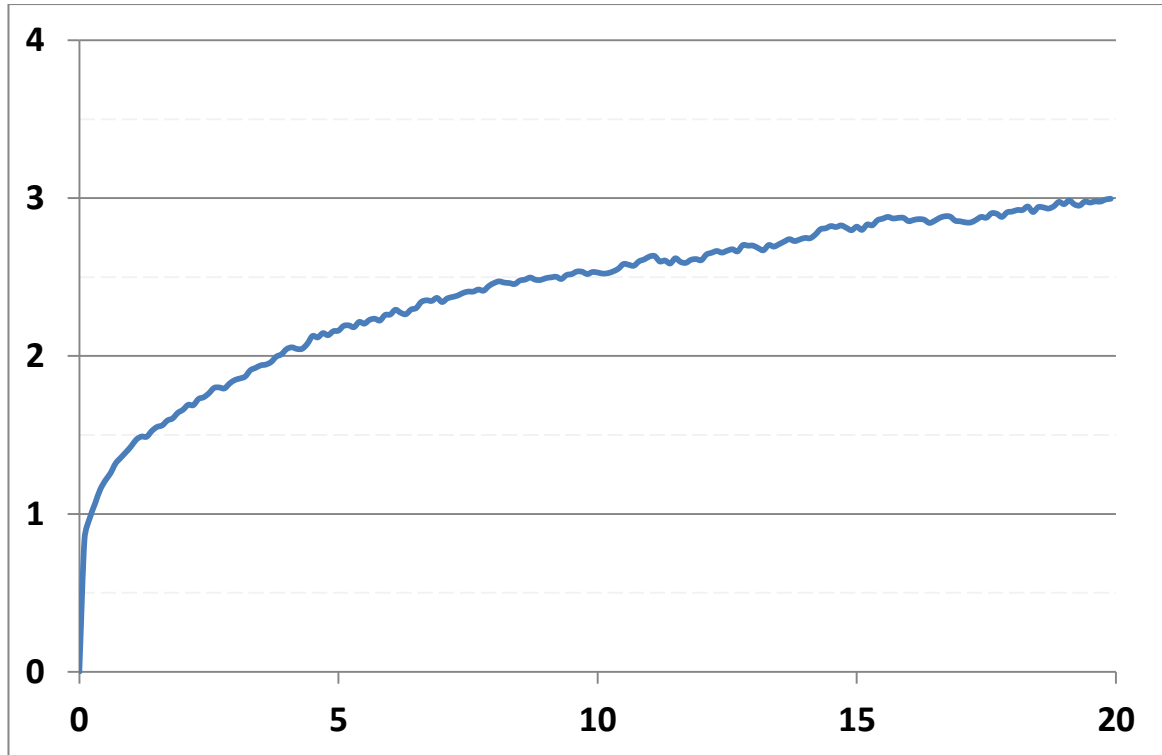


Figure 5.3 RMSD values for whole capsid simulation of CV-A21 homology model after atomistic MD simulation for 20 ns after 50,000 steps of minimisation. *x*- axis = time (ns), *y*- axis = RMSD (Å).

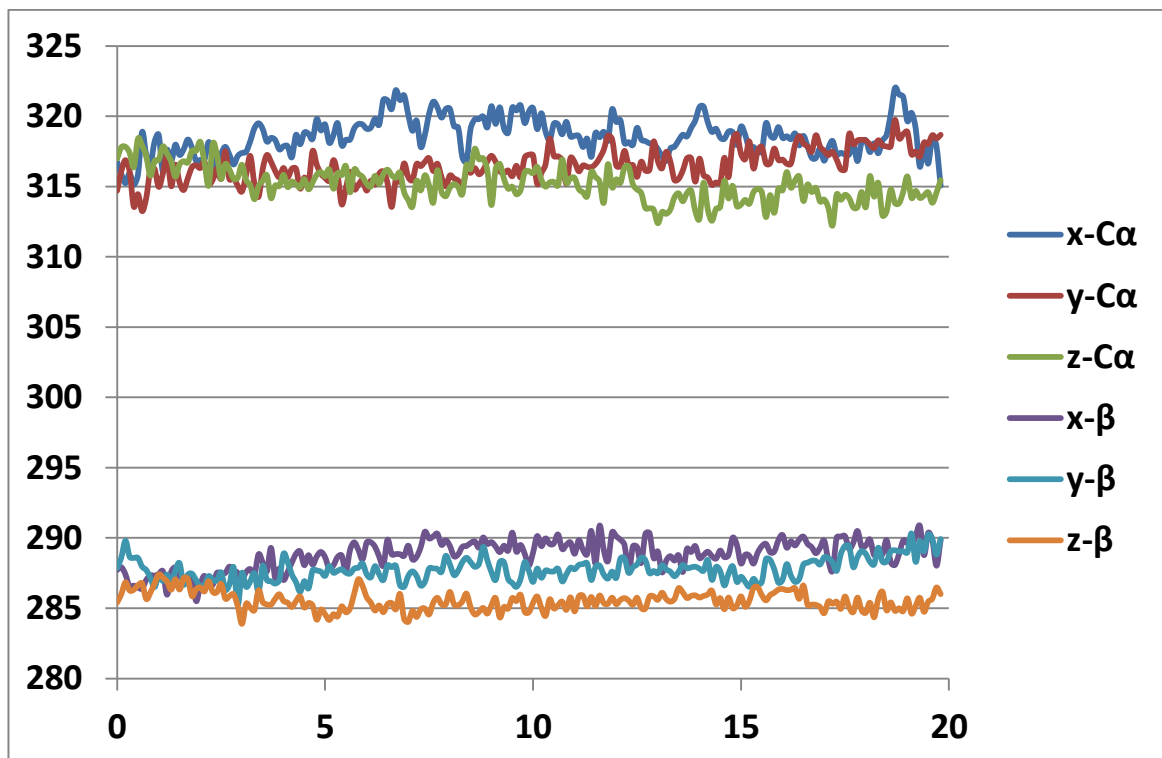


Figure 5.4 Virus diameter measurements for the CV-A21 capsid over the *x*, *y* and *z* axes in Å measured for the α - carbons ($C\alpha$) of the protein backbone and the extended-beta sheets comprising the primary structural regions of the Swiss jellyrolls (β). *x*- axis = time (ns), *y*- axis = diameter (Å).

After the initial “solvent shock”, the capsid structure remained stable as indicated in Figures 5.3 and 5.4. Upon completion of the simulation, all 60 protomers were positioned in the same space as protomer number one and an averaged structure was obtained from the overlay (Figure 5.5).

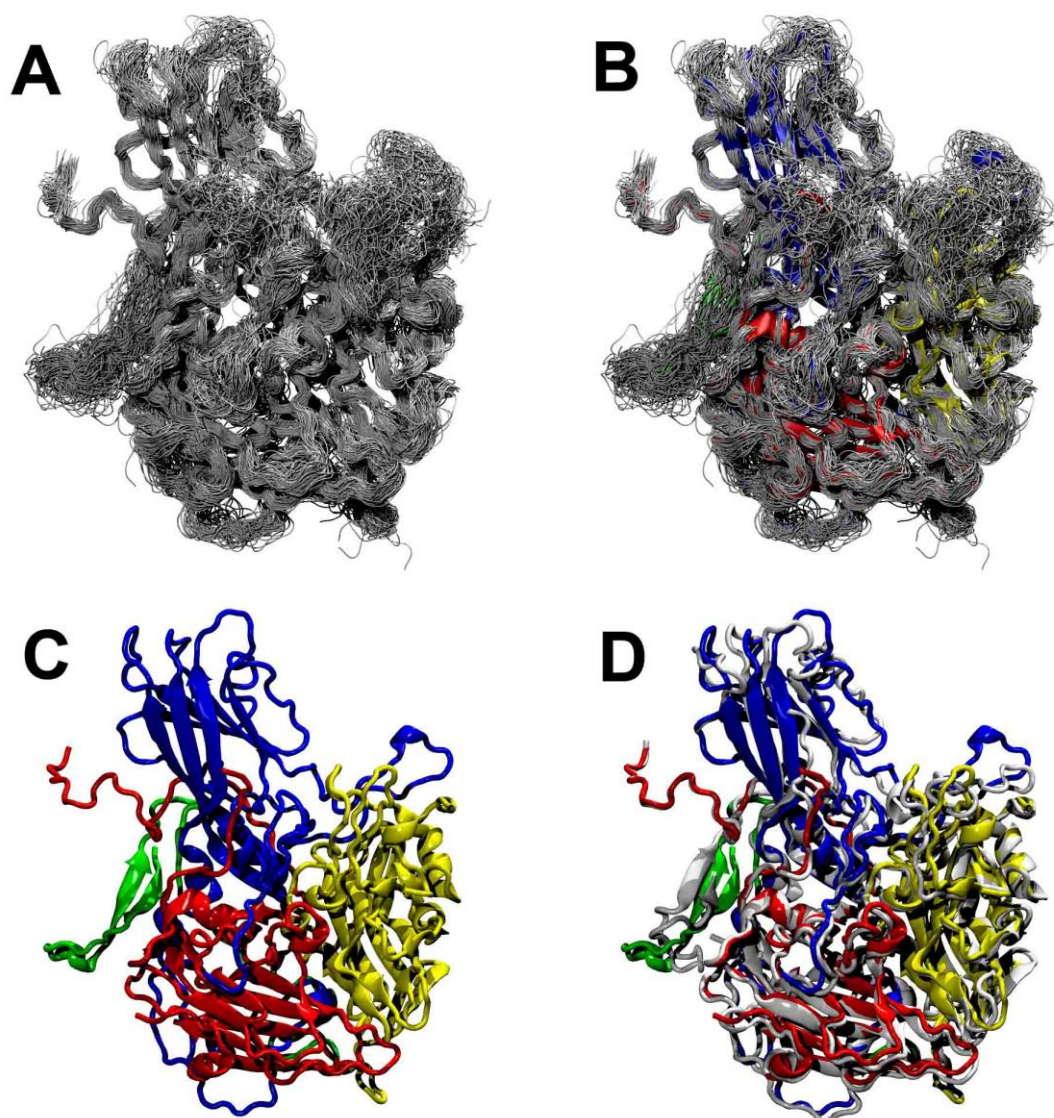


Figure 5.5 Protomer fit-averaged structure output after 20 ns of MD simulation. A) spatial overlay of 60 protomers comprising the virus capsid (grey), B) spatial overlay of the 60 protomers with fit averaged protomer coloured by chain VP1 (blue), VP2 (yellow), VP3 (red) and VP4 (green), C) fit averaged protomer coloured by chain with spatial overlay of 60 protomers removed, D) fit averaged protomer coloured by chain overlaid on original crystal structure for CV-A21 1Z7S, indicating areas of mismatch between *in-vitro* and *in-silico* experimental data.

Examination of the averaged virus proteins VP1, VP2, VP3 and VP4 that comprise the CV-A21 protomer, indicated backbone RMSD values ranging from 1.8 Å to 3.6 Å for the minimised structure and an overall RMSD for the protomer of 2.6 Å when compared to the original crystal structure. Subsequent to the 20 ns of atomistic MD simulation, the RMSD values dropped to a range of 1.6 Å to 3.8 Å for the viral capsid proteins and 2.5 Å for the protomer (table 5.3). The experimental error for the original x-ray data for CV-A21 was 3.2 Å³⁶⁵.

	VP1	VP2	VP3	VP4	Protomer
Minimized	2.707	1.816	2.578	3.850	2.609
20ns MD simulation	2.737	1.614	2.231	3.807	2.524

Table 5.3 RMSD values for individual proteins and protomer for CV-A21 post minimisation and post atomistic MD simulation.

ProQ quality assessment indicated that the original crystal structure for CV-A21 was ranked as an “extremely good model” (LGscore) and “fairly good model” (MaxSub). The models created using threading methods ranked initially as “Very good” for the LGscore but increased marginally and crossed the threshold to “extremely good”, once the fit-averaging method was utilised. MaxSub scores indicated a similar phenomenon (Table 5.4).

LG Score			MaxSub		
1Z7S Template	Protomer No. 1	Averaged Protomer	1Z7S Template	Protomer No. 1	Averaged Protomer
5.902	3.945	4.055	0.423	0.342	0.362

Table 5.4 Quality assessment of CV-A21 predicted model using the ProQ Protein quality prediction website. LGscore and MaxSub values obtained for the original crystal structure, a single protomer derived from position 1 of the predicted capsid structure after 20ns of atomistic MD simulation and the fit-averaged protomer structure derived from all 60 protomers of the capsid after 20ns of MD simulation. For information regarding LGscore and MaxSub ranking see section 5.3.

Based on an RMSD variation of 2.524 Å from the experimentally derived structure and ProQ values that indicated that the predicted model was a “good model”, manual threading methods using SwissPDBviewer were seen to be an appropriate technique. Subsequent models were generated using the SwissPDBviewer program and the validated methods described above were used to examine the quality of the final structures.

5.4.1.2 Enterovirus A71

Enterovirus A71 (EV-A71), which is considered an emerging pathogen, has gained a significant amount of attention from the scientific community^{134,152,164,165,317,366}. During the initial reconstruction attempts made during the PhD candidature period, no x-ray crystal structure data existed. As a result, three sub-genogroups of EV-A71 were reconstructed using manual methods via the SWISS-PDBviewer program, with sequence identity as low as 32.1% for individual virus capsid proteins and overall identity of 43.0% for the protomer structure (table 5.5). The three sub-genogroups constructed were EV-71 A, B5 and C4, as these represented the types most likely to be of scientific interest, with serotypes B5 and C4 co-circulating in the Asia Pacific region. A protomer was constructed using the 1HXS PDB template structure of wild poliovirus. The virus capsid structure for EV-A71 C4 was constructed as per the methods used for wild poliovirus and equilibrated for 1 ns using a live installation of Ubuntu Linux on a microSD card running on a Dell T3500 workstation using VMD and NAMD with CUDA GPU extensions.

EV-A71	VP1	VP2	VP3	VP4	Protomer
PV1-Wild	32.1%	50.9%	41.3%	57.9%	43.0%

Table 5.5 EV-A71 protomer amino acid percentage identity to wild poliovirus type 1 template structure. Comparison of EV-A71 prototype sequence AB550338 to sequence data derived from the PDB structure file 1HXS for wild poliovirus.

After the simulation of the EV-A71 comparative model was determined to be stable for 1 ns using Dell T3500 workstation, it was transferred to the BlueGene Q supercomputer and run on 4096 cores (65,536 threads) for 3 days, (3.3ns per day), resulting in a stable 10 ns atomistic MD simulation of a complete EV-A71 capsid structure (Figure 5.6). The simulation parameters used a constant number of atoms, pressure and temperature (NPT) ensemble as described in appendix A5.8. The stability of the capsid structures were examined using RMSD calculations for the entire protein over the 10 ns simulation time (Figure 5.7). As described for the examination of CV-A21 (section 5.4.1.1), virus expansion and contraction were determined using measurements of the virus capsid across the x, y and z axes for the α - carbons and extended beta sheets located towards the centre of mass of each protomer (Figure 5.8).

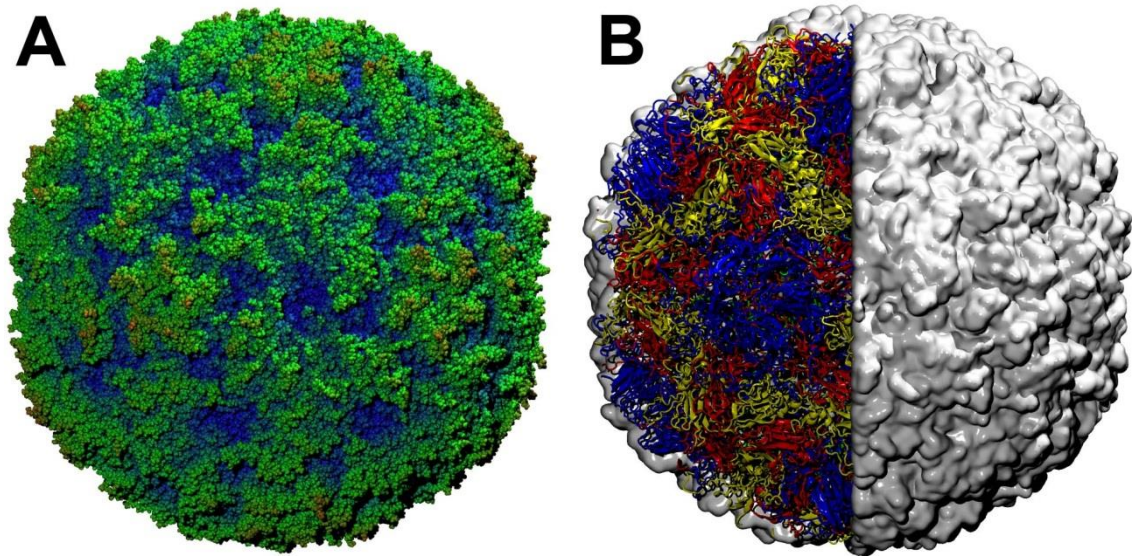


Figure 5.6 Final structure output after 10 ns of atomistic MD simulation for EV-A71. A) Radial depth shaded Van der Waals representation of the capsid measured from the origin (blue = 133 Å to Red 166 Å), B) cutaway density model calculated at 3.0Å, the +x and -z axes. -x and +y axes are represented in new cartoon format and shaded by protein chain VP1 (blue), VP2 (yellow), VP3 (red) and VP4 (not visible in this representation).

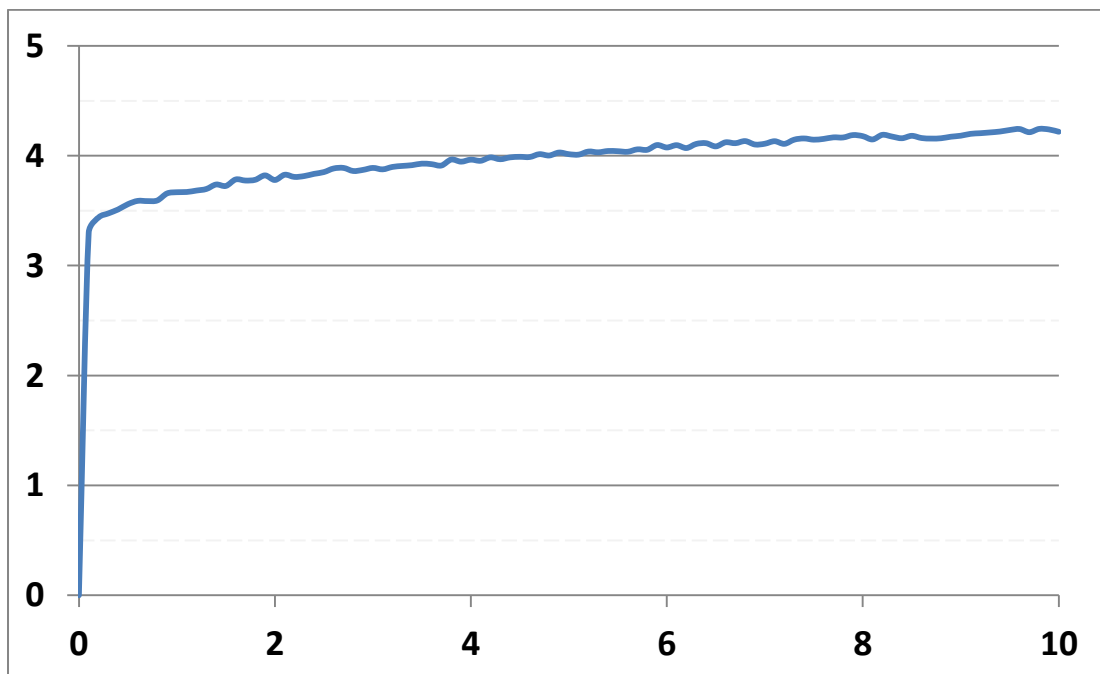


Figure 5.7 RMSD values for capsid simulation of EV-A71 homology model after atomistic MD simulation for 10 ns after 50,000 steps of minimisation. x- axis = time (ns), y- axis = RMSD (Å).

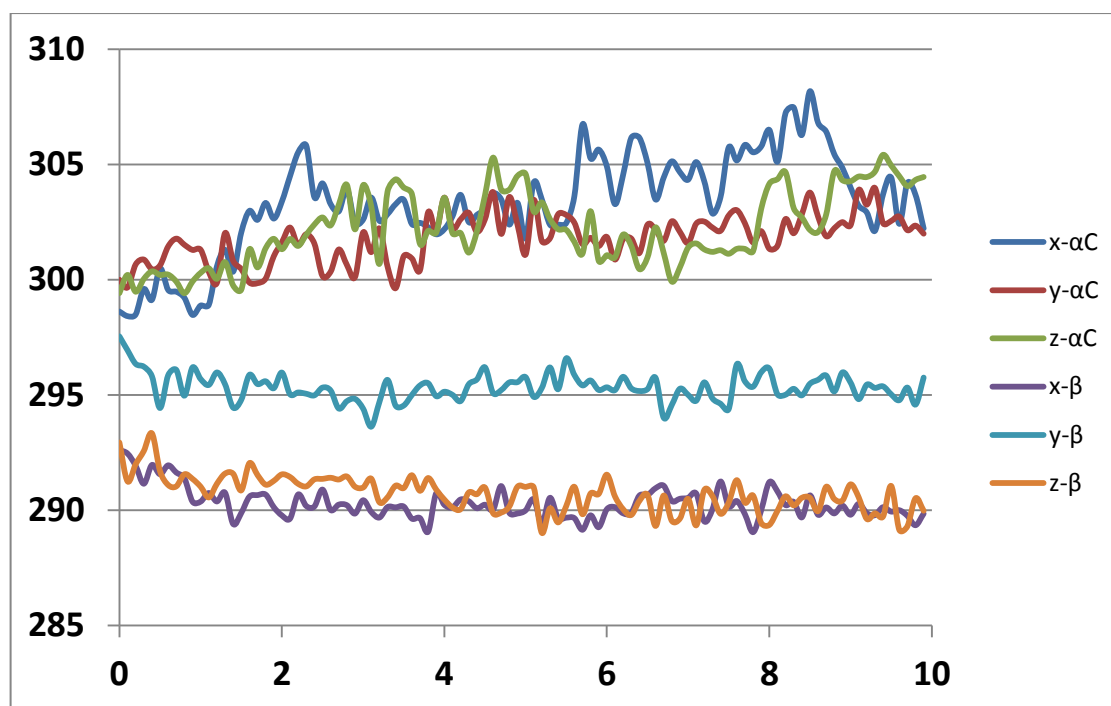


Figure 5.8 Virus diameter measurements for EV-A71 over the x, y and z axes in Å measured for the α -carbons ($C\alpha$) of the protein backbone and the extended beta sheets comprising the primary structural regions of the Swiss jellyrolls (β). x-axis = time (ns), y-axis = diameter (Å).

After the completion of the reconstruction and simulation process for the EV-A71 capsid structure, experimentally derived coordinate data for the virus in differing states using x-ray diffraction methods were released into the public domain, allowing the comparison of the model to a number of x-ray diffraction-derived EV-A71 structures (Figure 5.9, Table 5.6 and 5.7).

Chain	3VBF	3VBH	3VBO	3VBR	3VBS	3VBU	4AED
VP1	2.338	2.368	3.076	3.055	2.327	3.075	2.379
VP2	2.163	2.162	2.374	2.342	2.171	2.343	2.149
VP3	1.558	1.556	3.469	3.437	1.564	3.506	1.471

Table 5.6 “Backbone” RMSD values calculated for predicted models compared to available EV-A71 structures. Predicted model reconstructed using comparative modelling methods, simulated and then compared to all available reference structures, RCSB protein data bank (PDB ID codes indicated in header row).

	Predicted	3VBF	3VBH	3VBO	3VBR	3VBS	3VBU	4AED
Predicted	0	2.156	2.164	3.255	3.248	2.153	3.277	2.120
3VBF	2.156	0	0.147	2.326	2.313	0.121	2.343	0.506
3VBH	2.164	0.147	0	2.322	2.311	0.180	2.341	0.505
3VBO	3.255	2.326	2.322	0	0.266	2.324	0.301	2.423
3VBR	3.248	2.313	2.311	0.266	0	2.312	0.232	2.411
3VBS	2.153	0.121	0.180	2.324	2.312	0	2.341	0.507
3VBU	3.277	2.343	2.341	0.301	0.232	2.341	0	2.440
4AED	2.120	0.506	0.505	2.423	2.411	0.507	2.440	0

Table 5.7 Pairwise comparison of “Backbone” RMSD values calculated for predicted models compared to available EV-A71 structures. Pairwise comparison of reference protomer structures, chains VP1, VP2, and VP3. VP4 has not been included in this analysis as only 3VBF, 3VBH, 3VBS and 4AED contain coordinate data for VP4. RCSB protein data bank (PDB ID codes indicated in header row).

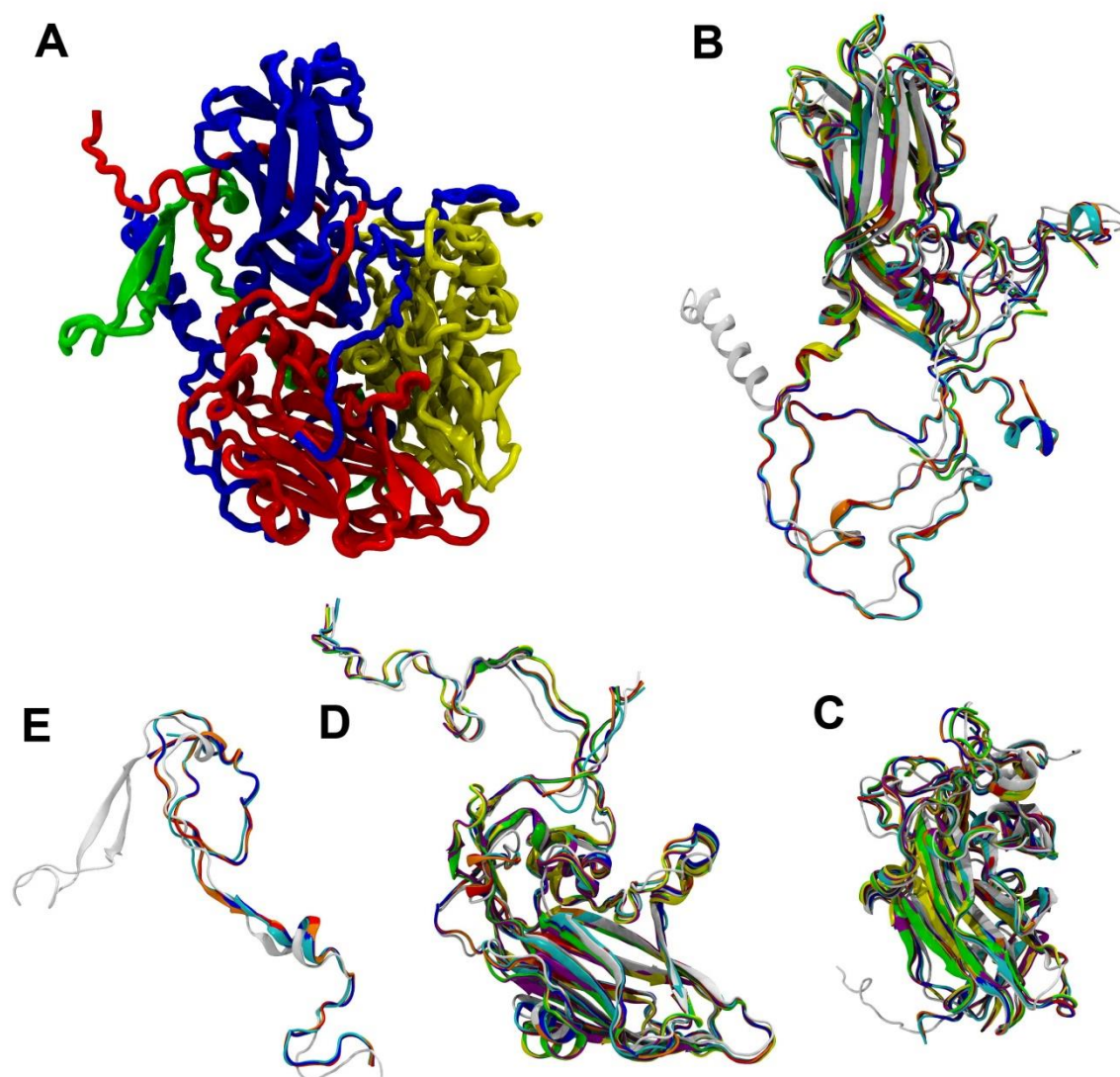


Figure 5.9 Structural comparisons of enterovirus 71 homology models with X-ray crystal data. A) EV-A71 protomer fit-averaged structure coloured by chain VP1 (blue), VP2 (yellow), VP3 (red) and VP4 (green), B) EV-A71 predicted VP1 structure (white) overlaid on the seven available crystal structures for EV-A71 C) EV-A71 predicted VP2 structure (white) overlaid on all available crystal structures for EV-A71 D) EV-A71 predicted VP3 structure (white) overlaid on all available crystal structures for EV-A71 and E) EV-A71 predicted VP4 structure (white) overlaid on all available crystal structures for EV-A71. Reference structure colours are as follows: Red – 3VBF, Yellow – 3VBH, Orange – 3VBO, Green – 3VBR, Blue – 3VBS, Purple – 3VBU, Cyan – 4AED.

Quality assessment using the ProQ website, ranked the experimentally derived structure for EV-A71 (PDB ID 4AED) as an “extremely good model” (LGscore) and “fairly good model” (MaxSub). The models created using threading methods ranked initially as “Very good” for the LGscore, with a score approaching 5 for the fit-averaged protomer. As indicated in section 5.4.1.1, the MaxSub scores indicated that all models were “fairly good” with values within the 0.1 to 0.5 range (Table 5.8).

LG Score			MaxSub		
4AED Template	Protomer No. 1	Averaged Protomer	4AED Template	Protomer No. 1	Averaged Protomer
6.110	4.113	4.819	0.392	0.302	0.375

Table 5.8 Quality assessment of EV71 predicted model using the ProQ Protein quality prediction website. LGscore and MaxSub values obtained for the original crystal structure, a single protomer derived from position 1 of the predicted capsid structure after 10ns of atomistic MD simulation and the protomer-fit averaged structure derived from all 60 protomers after 10ns of MD simulation. For LGscore and MaxSub ranking see section 5.3.

Using the most complete experimentally derived EV-A71 structure (4AED) with the lowest resolution (3.8 Å) as a reference³⁶⁷, the protomer-fit average model was evaluated and RMSD figures obtained for all viral capsid proteins for the backbone atoms and all atoms (excluding hydrogen) compared to the reference (Table 5.9).

	VP1	VP2	VP3	VP4*	Protomer
Not Hydrogen	3.903	3.761	3.358	4.357	3.716
Backbone	2.433	2.319	1.538	2.746	2.120

Table 5.9 RMSD values for predicted models compared to the reference structure 4AED. Values calculated for the backbone, and all atoms excluding hydrogen, of the capsid proteins for the predicted EV-A71 model after 10ns MD Simulation compared to the experimentally derived structure (PDB accession number 4AED).

Results indicated that the comparative model, although only sharing an overall identity of 43% with the wild poliovirus template model (1HXS), was within 3.716 Å of the experimentally derived structure when comparing all atoms except hydrogen. This closely mirrors the resolution of the x-ray diffraction data for EV-A71 at 3.8 Å (PDB ID 4AED). It should be noted that the N-terminal α -helix of VP1 was excluded from this analysis, given the uncertainty of the appropriate positioning of this particular motif, as was noted for wild poliovirus in chapter 4.

In order to further assess the quality of the predicted model and its potential utility in the examination of antiviral drug binding, the amino acid residues lining the hydrophobic pocket were examined. Amino acids within 4 Å of the native laurate molecule in PDB file 4AED were selected for comparison with the EV-A71 model generated using comparative protein methods (Figure 5.10).

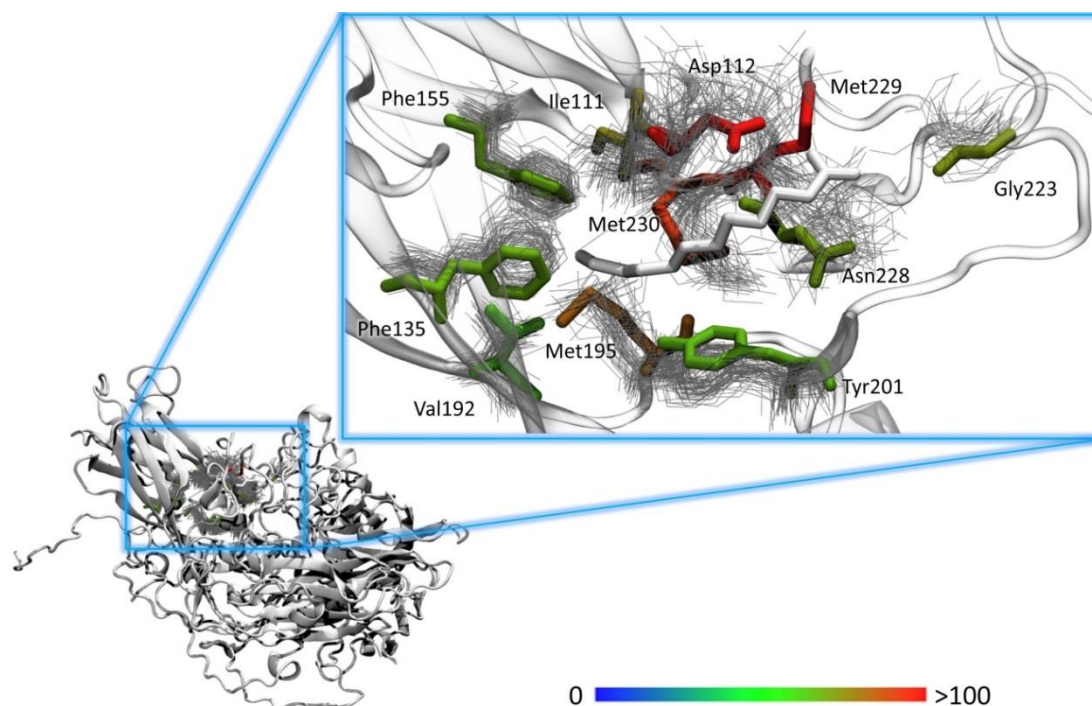


Figure 5.10 Comparison of hydrophobic pocket x-ray diffraction data and the predicted structure for EV-A71. showing residues lining the hydrophobic pocket and within 4 Å of the native lauric acid molecule. Left (white) a single protomer derived from the crystal structure of EV-A71 C4 (PDB file accession number 4AED). Break-out figure shows the amino acid residues (“liquorice” representation) coloured by beta value from the crystal structure of EV-A71 C4 (PDB file accession number 4AED) within 4 Å of lauric acid (white) in the hydrophobic pocket. Corresponding individual amino acid residues derived from 60 protomers after 10ns of MD simulation of the predicted EV-A71 model are identified as a grey “line” representations and numbered by position within the VP1 amino acid sequence and three-letter amino acid abbreviation. Scale-bar = B-factor range 0 (blue) to >100 (red).

Observation of the interior lining of the hydrophobic pocket showed a good correlation between the predicted model and the experimentally derived structure. Where significant deviation occurred between the two structures, a corresponding increase in B-factor value was also observed, indicating that significant thermal fluctuation was present. This was reflected in the predicted model by a corresponding increase in spatial disorder (Figure 5.10).

	60 protomers	Fit-averaged
RMSD (Backbone)	1.410 (SD 0.259)	1.089
RMSD (Not Hydrogen)	3.513 (SD 0.114)	3.292

Table 5.10 RMSD values calculated for hydrophobic pocket x-ray diffraction data and the predicted structure for EV-A71. Values calculated for the backbone structure and all atoms excluding hydrogen between the experimentally derived structure (PDB accession number 4AED) and the predicted structure for residues lining the hydrophobic pocket and within 4 Å of the native lauric acid molecule, for all protomers of the capsid after 10ns of MD simulation and the fit-averaged protomer structure.

Comparison of the hydrophobic pocket indicated that there was approximately 1 Å variation from the experimentally derived structure when calculating RMSD for the backbone coordinates of the fit-averaged structure. When calculating the variation of individual protomers, this difference increased to 1.4 Å (Table 5.10). The exterior of the predicted structure when compared to PDB file 4AED showed a deviation (RMSD) as low as 2.1 Å for the backbone structure and 3.7 Å for all atoms excluding hydrogen, when compared to x-ray crystallographic data (Table 5.9). The interior aspects of the capsid deviated significantly due to the rearrangement of the VP4 protein chain, which differs to structures observed in wild poliovirus. This observation is discussed further in section 5.5.

5.4.2 Comparative Modelling of Novel and Newly Described Viruses

In order to reconstruct the newly described and novel enteroviruses, EV-C96 and EV-A120 respectively, the manual methods determined in section 5.4.1.1 were applied. Sequences representing the capsid region for each virus were obtained using a combination of primer walking and primer design by alignment of homologous serotypes. A combination of the SwissPDBviewer manual threading method and the web-based SWISS-Model was a simple and accurate method for the prediction of protomeric structures for capsid construction, and was used for the subsequent experimental reconstruction of newly-described and novel enteroviruses.

5.4.2.1 Enterovirus C96

The template selection for the reconstruction of the EV-C96 structures was based on homologous structures available as x-ray diffraction data on the RCSB. In this case, two templates of species C *Enteroviruses* were deemed suitable for reconstruction according to phylogenetic inference using maximum-likelihood methods, namely coxsackievirus A21 and wild poliovirus. Given that the wild poliovirus coordinates for 1HXS represented the most complete structures at the best resolution, comparative modelling was performed using these coordinates. The percentage amino acid identity for the template structures ranged from 59.7% to 84.0% with an overall protomer sequence identity of 68.5% (Table 5.11).

EV-C96	VP1	VP2	VP3	VP4	Protomer
PV1-Wild	59.7%	72.4%	70.9%	84.0%	68.5%

Table 5.11 EV-C96 protomer amino acid percentage identity to wild poliovirus type 1 template structure. Comparison of EV-C96 sequence 118030-PNG-05 to sequence data derived from the PDB structure file 1HXS for wild poliovirus.

Reconstruction and atomistic method dynamics simulation was performed using the methods described previously in chapter 4 and chapter 5. Simulations were performed as previously described (section 5.4.1.1) and the stability of the capsid structures were examined using RMSD calculations for the entire protein over a 10 ns simulation time (Figure 5.11). As described for the examination of CV-A21 (section 5.4.1.1), virus expansion and contraction was determined using measurements of the virus capsid across the x, y and z axes (Figure 5.12). Observations over the course of the 10ns simulation indicated subtle changes in virus diameter and plateau of RMSD values, indicating that the virus capsid had stabilised relatively quickly compared to the previous EV-A71 model. The final capsid structure is shown in figure 5.13.

Quality assessment using the ProQ website ranked the predicted model as a “very good model” using the LGscore, with the protomer-fit average structure falling just short of the “extremely good model” ranking at 3.914. MaxSub values were similar to those values for the experimentally derived and predicted models for CV-A21 and EV-A71, all of which fell within the “fairly good model” range of 0.1 to 0.5 (Table 5.12).

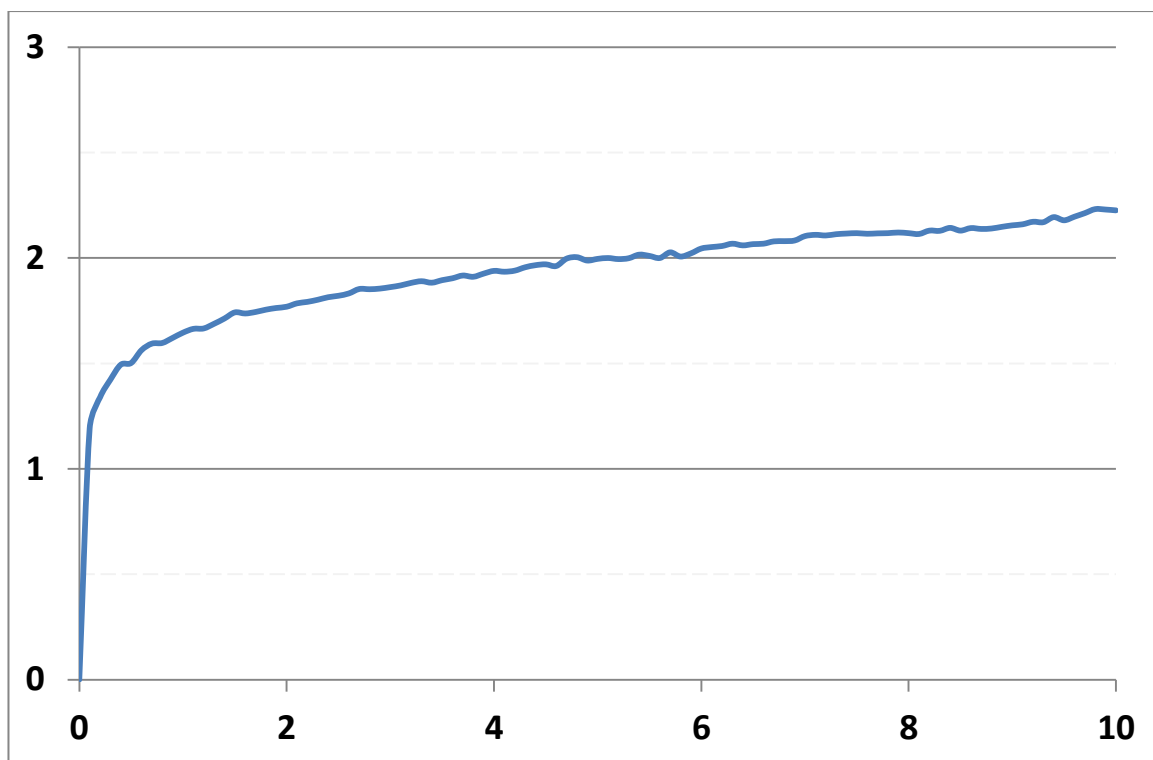


Figure 5.11 RMSD values for capsid simulation of EV-C96 homology model after atomistic MD simulation for 10 ns after 50,000 steps of minimisation. *x*- axis = time (ns), *y*- axis = RMSD (Å).

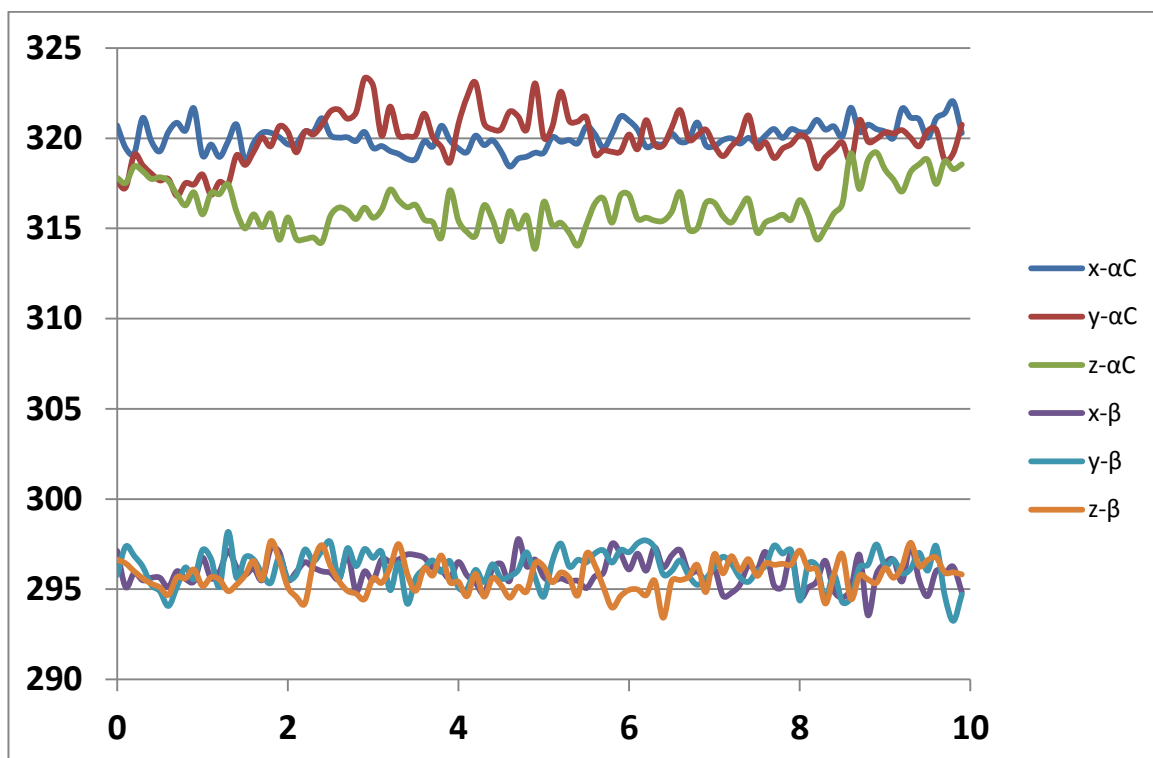


Figure 5.12 Virus diameter measurements for EV-C96 over the *x*, *y* and *z* axes in Å measured for the α - carbons ($C\alpha$) of the protein backbone and the extended beta sheets comprising the primary structural regions of the Swiss jellyrolls (β). *x*- axis = time (ns), *y*- axis = diameter (Å).

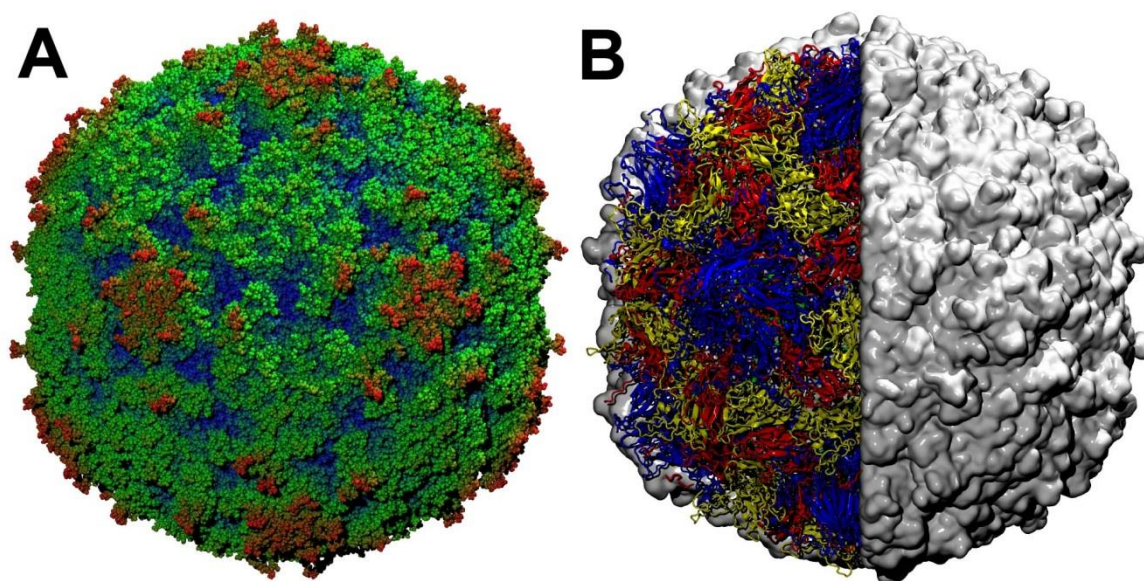


Figure 5.13 Final structure output after 10 ns of atomistic MD simulation for EV-C96. A) radial depth shaded Van der Waals representation of the capsid measured from the origin (blue = 133 Å to Red 166 Å), B) cutaway density model calculated at 3.0Å on the $+x$ and $-z$ axes. $-x$ and $+y$ axes are represented in new cartoon format and shaded by protein chain VP1 (blue), VP2 (yellow), VP3 (red) and VP4 (not visible in this representation). Refer to section 5.4.3 for topological comparison of the EV-C96 predicted model to other representatives of the family *Picornaviridae*.

LG Score		MaxSub	
Protomer No. 1	Averaged Protomer	Protomer No. 1	Averaged Protomer
3.593	3.914	0.319	0.359

Table 5.12 Quality assessment of EV-C96 predicted model using the ProQ Protein quality prediction website. LGscore and MaxSub values obtained for a single protomer derived from position 1 of the predicted capsid structure after 10ns of atomistic MD simulation and the protomer-fit averaged structure derived from all 60 protomers after 10ns of MD simulation. For LGscore and MaxSub ranking see section 5.3.

5.4.2.2 Enterovirus EV-A120

Given that the complete capsid region was now sequenced as described in chapter 3, it was possible to use the above described comparative protein modelling methods to reconstruct the capsid structure of the newly discovered enterovirus EV-A120. In order to do so, appropriate templates were selected from homologous structures determined using x-ray diffraction methods, which represented a high degree of amino acid sequence identity. The capsid region sequence was queried against available data using the SwissModel website in order to determine suitable templates.

This method indicated that X-ray crystallography data were available for EV-A71 (amino acid identity of 72.1%, Table 5.13), CV-A16 (amino acid identity of 71.6%) and cryo-EM data for CV-A7 (amino acid identity 80.0%), Figure 5.14. The x-ray diffraction data for EV-A71 was more complete in terms of chain data for VP1, VP2, VP3 and VP4. The x-ray diffraction data for CV-A7 (PDB file 4AGY, 4AGX) contained only partial chain data. The experimental error of 3.8 Å for EV-A71 and 6.09 Å for CV-A7 indicated that although CV-A7 showed a higher amino acid sequence identity with EV-A120, the completeness and high resolution of the x-ray diffraction data for EV-A71 (PDB file 4AED) was a better choice for modelling the novel virus EV-A120.

EV-A120	VP1	VP2	VP3	VP4	Protomer
EV-A71	64.6%	82.6%	71.4%	69.5%	72.1%

Table 5.13 EV-A120 protomer amino acid percentage identity to EV-A71 template structure.

Comparison of EV-A120 to sequence data derived from the PDB structure file 4AED for EV-A71 genogroup C4.

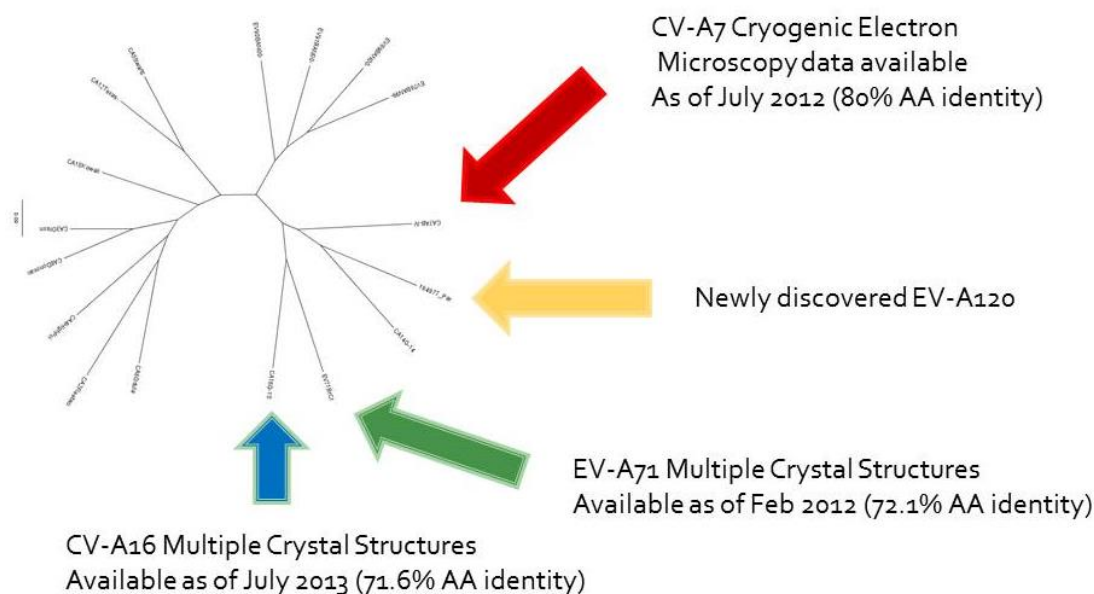


Figure 5.14 Ancestral relationships from the available crystal structures for EV-A120. Radial phylogenetic tree calculated using maximum likelihood method using the HKY+G+I model for the complete capsid encoding region. Amino acid percent identity given in parentheses.

After 10 ns of atomistic MD simulation using NAMD with a constant number of atoms, pressure and temperature (NPT), the stability of the capsid structures were examined using RMSD calculations for the entire protein over the 10 ns simulation time (Figure 5.15). Virus expansion and contraction was determined using measurements of the virus capsid across the x, y and z axes for the α - carbons and extended beta sheets located towards the centre of mass of each protomer (Figure 5.16). It was noted that the RMSD values calculated over the duration of the simulation showed evidence of stabilisation after the initial “solvent shock”, but the assessment of the capsid expansion indicated that although the extended β -sheets remained at a constant radius from the origin, the outermost α -carbons contracted by almost 5 Å. Examination of the simulation indicated that the outermost loop structures of the BC loop and the C-terminal end of VP1 settled toward the capsid surface. The final capsid structure after 10 ns is indicated in Figure 5.17.

Quality assessment using the ProQ website ranked the predicted model as a “extremely good model” using the LGscore for both the single protomer and protomer-fit averaged structures. MaxSub values were once again within the 0.1 to 0.5 range corresponding to the “fairly good model” ranking, as seen for all experimentally derived and predicted structures examined in this body of work. Table 5.14.

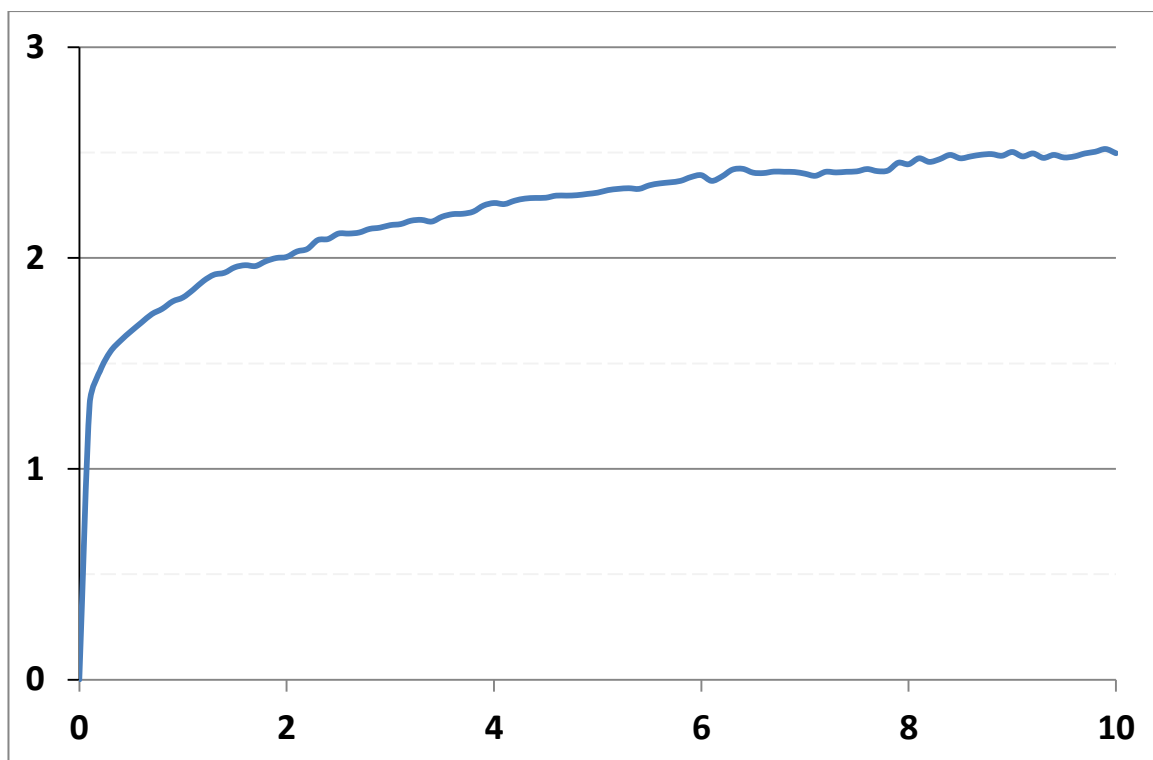


Figure 5.15 RMSD values for capsid simulation of EV-A120 homology model after atomistic MD simulation for 10 ns after 50,000 steps of minimisation. *x*- axis = time (ns), *y*- axis = RMSD (Å).

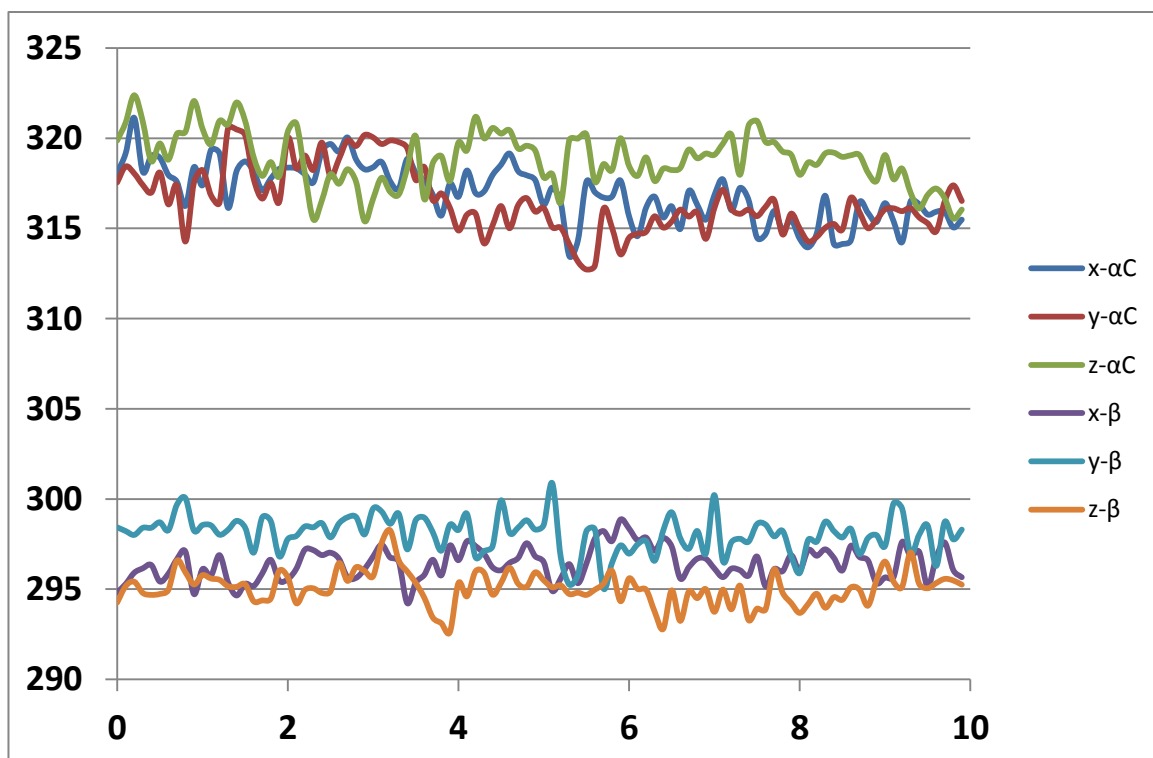


Figure 5.16 Virus diameter measurements for EV-A120 over the *x*, *y* and *z* axes in Å measured for the α - carbons ($C\alpha$) of the protein backbone and the extended beta sheets comprising the primary structural regions of the Swiss jellyrolls (β). *x*- axis = time (ns), *y*- axis = diameter (Å).

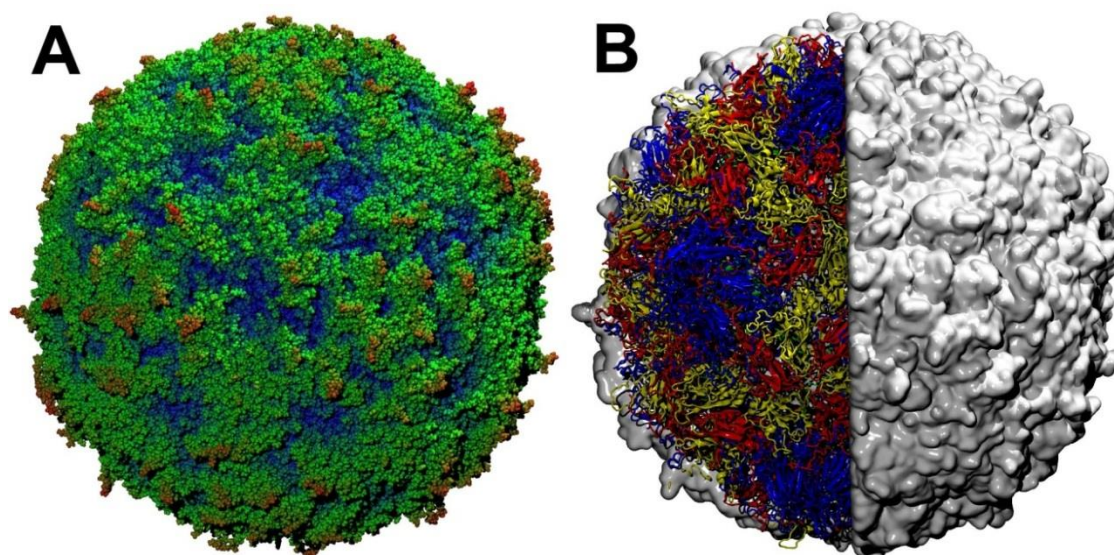


Figure 5.17 Final structure output after 10 ns of atomistic MD simulation for EV-A120. A) Radial depth shaded Van der Waals representation of the capsid measured from the origin (blue = 133 Å to Red 166 Å), B) cutaway density model calculated at 3.0Å, the +x and -z axes. -x and +y axes are represented in new cartoon format and shaded by protein chain VP1 (blue), VP2 (yellow), VP3 (red) and VP4 (not visible in this representation). Refer to section 5.4.3 for topological comparison of the EV-A120 predicted model to other representatives of the family *Picornaviridae*.

LG Score		MaxSub	
Protomer No. 1	Averaged Protomer	Protomer No. 1	Averaged Protomer
4.257	4.437	0.305	0.350

Table 5.14 Quality assessment of EV-A120 predicted model using the ProQ Protein quality prediction website. LGscore and MaxSub values obtained a single protomer derived from position 1 of the predicted capsid structure after 10ns of atomistic MD simulation and the protomer-fit averaged structure derived from all 60 protomers after 10ns of MD simulation. For LGscore and MaxSub ranking see section 5.3.

Given the issues surrounding the lack of success cultivating the virus, it was decided to concentrate on the identification of a potential receptor binding site in order to help select a more appropriate cell line for cultivation. A review of the literature indicated that members of the clade shared by EV-A120, specifically CV-A7, CV-A14, CV-A16 and EV-A71 show evidence of the utilisation of the SCARB2 receptor for binding. A binding motif has been described that indicated a series of mutations that affected SCARB2 binding affinity^{368,369}. An amino acid alignment of the clade members, CV-A7 and EV-A71, with EV-A120 was performed to indicate whether this motif was present, Figure 5.18

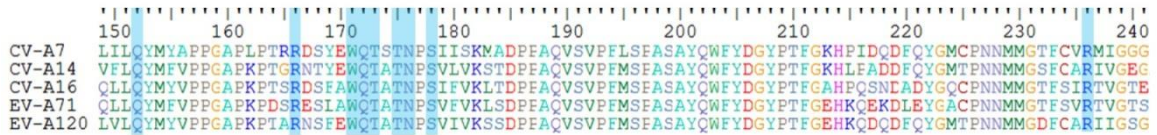


Figure 5.18 Multiple amino acid sequence alignment for CV-A7, CV-A14, CV-A16, EV-A71 and EV-A120. A fragment of the VP1 encoding region of the genome indicating the residues (blue shading) comprising an area in the “canyon” wall considered significant to binding of the cellular receptor SCARB2^{368,369}.

The alignment of CV-A7, EV-A71 and EV-A120 indicated that the hypothesised binding motif was present. CV-A7 and EV-A71 were selected for the alignment based on the availability of x-ray diffraction derived coordinate data that could be used to compare the 3-dimension presentation of the binding motif that lies on the canyon wall (Figure 5.19). Structurally, the motif present in CV-A7 and EV-A71 was present in EV-A120.

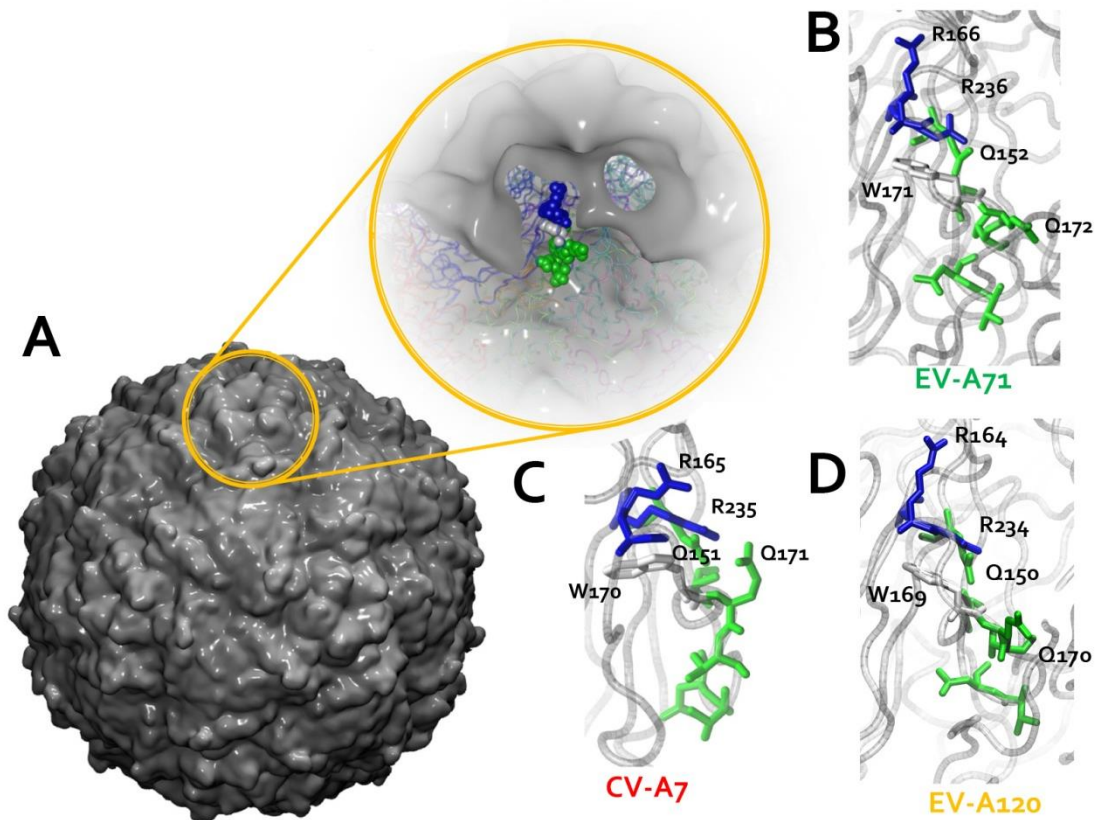


Figure 5.19 3-Dimensional reconstruction of EV-A120 after 10ns MD simulation indicating residues considered significant to SCARB2 binding. A) Surface representation of EV-A120 capsid at 5 Å resolution, break-out figure indicates residues flanking the EF loop of VP1 as described ^{368,369}, B) hypothesised SCARB2 amino acid motif indicated for EV-A71, C) CV-A7 and D) EV-A120 represented in a “liquorice” format with colouring by residue type, Basic (Blue), Hydrophobic (White) and polar uncharged (Green).

5.4.3 Comparison of Picornavirus Capsid Topography.

Comparing the topological features of the newly described enterovirus EV-C96 and the novel enterovirus EV-A120 with other members of the family *Picornaviridae*, (Figure 5.20), indicated a number of structural variations worthy of discussion. One feature of prominence common to all viruses represented in figure 5.20 was the 2-fold axis of symmetry, which was represented by a deep depression in the surface of all viruses depicted.

With regard to the models presented in this chapter, firstly, EV-C96 bore a striking resemblance to the related species *C Enterovirus* (wild poliovirus, PDB 1HXS), with prominent “mesa” representing the five-fold axis of symmetry, punctuated by BC-loop structures that extended into the solvent accessible space to a radial distance of 166 Å (Figure 5.20). The puff of the EF loop of VP2 on the tips of the pseudo three-fold axis of symmetry, extended to 160 Å. Also of note were the deep well-defined canyons, reminiscent of poliovirus. These canyons were not as wide as those found in poliovirus and were partially obscured by the C-terminal aspect of the VP1 protein, indicating a tendency to make contact with the canyon wall. These features were in stark contrast to that of the aphthoviruses- Foot and Mouth Disease virus and Equine Rhinitis A virus, which showed a relatively modest prominence of the 5-fold axis of symmetry, but a significantly well-defined pseudo 3-fold axis of symmetry. This observation was also made for the species *F Enterovirus*, *Bovine Enterovirus 2*.

Secondly, the EV-A120 virus although modelled on the EV-A71 coordinate data, showed a substantially different morphological presentation in comparison to the marked similarity of the predicted model for EV-C96 to poliovirus. The most significant feature missing from EV-A120 when compared to the EV-A71 structure was that of the prominent canyon feature (though not as prominent as poliovirus). The pseudo 3-fold axis of symmetry was similar in height and features but the puff representing the EF loop of VP2 was more prominent extending out to the solvent accessible space. As with EV-C96, the fivefold axis of symmetry was highlighted by a VP1 BC-loop that extended further into the solvent accessible space than EV-A71.

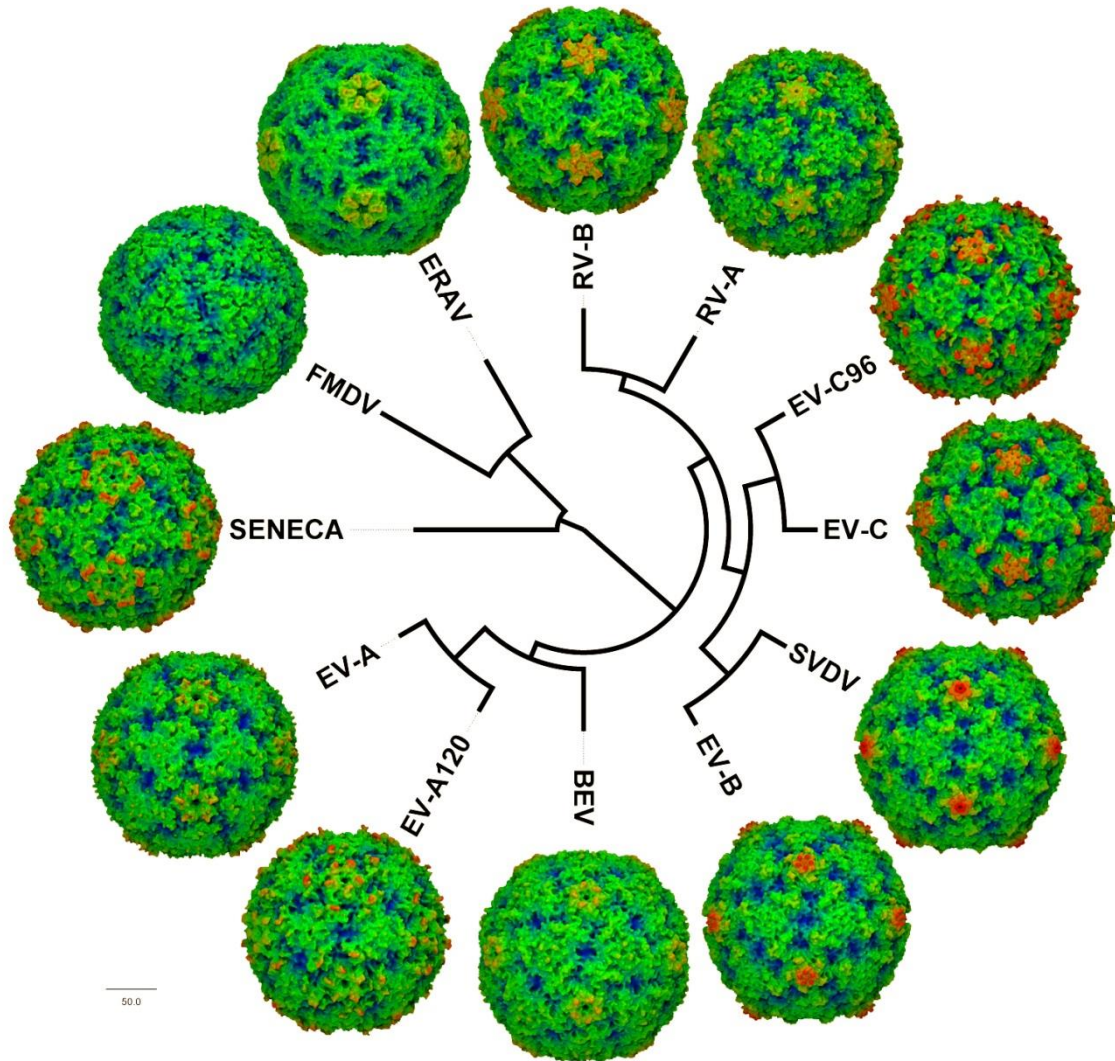


Figure 5.20 Picornavirus phylogeny compared with capsid topography. Amino acid maximum-likelihood polar phylogram, indicating genetic relationships of the capsid encoding region between representative members of the family *Picornaviridae*. Each representative displays a capsid reconstruction created using the VMD software package³³⁵ with a “Surface” representation calculated at a 1 Å radius from x-ray diffraction data obtained from the RCSB Protein Data Bank, Exceptions are EV-A120 and EV-C96 which were reconstructed using comparative protein modelling methods, equilibrated for 10ns using atomistic MD simulation (section 5.4.2). Capsids are coloured by radial distance from the point of origin, Blue- 133 Å to Red- 166 Å. Picornavirus representatives used: **SENECA** (Seneca valley virus, *GenBank Accession No. DQ641257, PDB Accession No.3CJI*), **FMDV** (Foot and mouth disease virus-A, *AY593803, IZBE*), **ERAV** (Equine rhinitis A virus, *DQ272127, 2WFF*), **RV-B** (Rhinovirus-B3, *DQ473485, 1RHI*), **RV-A** (Rhinovirus-A1, *FJ445111, 1RIA*), **EV-C96** (sections 3.4.3 and 5.4.2.1), **EV-C** (Wild Poliovirus-1, *NC002058, IHXS*), **SVDV** (Swine Vesicular Disease Virus, *EU151448, 1OOP*), **EV-B** (Echovirus-1, *AF029859, 1EVI*), **BEV** (Bovine Enterovirus-2, *NC021220*), **EV-A120** (sections 3.4.4 and 5.4.2.2), **EV-A** (Enterovirus-A71, *ETU22521, 4AED*). Scale bar indicates amino acid substitutions per site. Alignment performed using the ClustalW algorithm using the BioEdit software package version 7.1.20³⁷⁰. Branch lengths calculated using the Phylip Protein Distance model³⁷¹.

5.5 Discussion

This chapter investigated the use of comparative protein modelling methods to reconstruct enterovirus capsid structures. Examination of two manual threading methods for the VP1 of CV-A21, using prototype sequence and freely available software packages, indicated that SwissPDBviewer was easier to use than Modeller and yielded the most accurate and precise outputs when compared to experimentally derived x-ray diffraction data. As a result, SWISSPDBviewer was used to model the predicted capsid structures for EV-A71 (sub-genogroup C4), EV-C96 and EV-A120 using a manual threading method.

A quality assessment for all models was determined for the predicted structures. The capsid structures were examined for any indication of instability using an RMSD calculation and observation of capsid expansion or contraction using diameter measurements of the whole capsid over the course of the complete MD simulation. In all instances, the observations indicated that the predicted capsid structures were stable. In the case of the EV-A71 simulation, a significant increase in RMSD was observed after the initial minimisation and “solvent shock”. This can be attributed to a number of steric clashes that existed within the initial reconstruction stage. Resolution of these clashes in the first stages of simulation resulted in relatively large increase in RMSD followed by a plateau phase.

The characteristics of the RMSD increases over time could also be indicative of mechanical characteristics inherent in each viral capsid. CV-A21 showed a gradual and smooth approach to plateau, with a significant RMSD gradient still present after 10 ns simulation time. This may be indicative of greater structural elasticity present in CV-A21 capsids when compared to other virus capsids. In contrast, EV-A71 showed an abrupt RMSD jump to $\sim 4 \text{ \AA}$ in the first 100 ps, where it remained for the rest of the simulation (reaching a true plateau). This may be due to the lower sequence identity of the query sequence to template structure, resulting in an increased initial rearrangement of side chains in comparison to the CV-A21 model. The RMSD observations for EV-A71 may also indicate that this capsid is structurally less malleable, with a propensity for stepwise, rigid motions. EV-C96 and EV-A120, on the other hand, are likely to possess mechanical

properties intermediate between those of CV-A21 and A71. The functional and biological implications of the mechanical properties of different capsids deserve further study.

During the simulation of EV-A71, no significant changes in capsid diameter were observed, indicating that overall the capsid was structurally sound. Predicted structures for CV-A21, EV-C96 and EV-A120 were derived from templates with a closer match to template data, with protomer amino acid sequence percentage identity values of 68.1% for CV-A21, 68.5% for EV-C96 and 72.% for EV-A120, compared to a sequence identity of 43% for the EV-A71 predicted model, a value that lies between the 30% lower limit and the 50% cut-off described by Chothia and Lesk³⁵⁹.

Model assessment was performed using ProQ on the original x-ray diffraction data, on the individual protomer assemblies, and on the data derived using the protomer fit-averaging method. In most instances, the LGscore was above 4, indicating an excellent model. In the case of EV-C96, this value was 3.593 for the single protomer, but increased to 3.914 when using the protomer fit-average model. The MaxSub score was below 0.5 for all models examined, including those derived using x-ray diffraction methods. Values for predicted models ranged from 0.305 for EV-A120 (single protomer) to 0.375 for EV-A71 (protomer fit-averaged). These values were comparable to those of experimentally derived structures, indicating that the predicted models were of a quality comparable to those derived using x-ray diffraction methods.

For the prediction of EV-A71, significant differences were observed in the N-terminal region of VP4 and the N-terminal α -helix structure positioning of VP1. With respect to the variation in the conformation of VP4 in EV-A71, this phenomenon has been described previously and resulted in difficulty during the reconstruction of cryo-EM data for CV-A7. As a result, the authors elected to exclude the entire protein chain from the final structure²³⁴. The significant variation in structure is indicative of the problems associated with the prediction of protein chain configuration in 3-dimensional space. Given that the N-terminal helix of VP4 plays a critical role in the replication of enteroviruses, and is known to transition from the interior surface of the capsid to the exterior in the replication of poliovirus⁸⁹, then the data for the N-terminal aspect of VP4 for EV-A71 may be correct and representative of a transitional state of the protein (Figure 5.21).



Figure 5.21 Structural variation of poliovirus and EV-A71 VP4 crystal structures. Left) poliovirus type I 1HXS and Right) EV-A71 4AED. Single protomer reconstructed to form a pentamer using the matrix data contained within the respective PDB files. The initial protomer (position one) is coloured in solid green and subsequent mapping of the remaining four protomers are ghosted to allow visualisation of the differences in the N-terminal aspect of the VP4.

As for the placement of the N-terminal α -helix of the VP1 protein, the clockwise rotation of the motif is not surprising given the different conditions used in the derivation of the experimental data. X-ray diffraction data were derived from the poliovirus crystal at pH 7.0 and pH 8.0 for EV-A71. Temperature values at which the x-ray diffraction data were obtained are known for poliovirus (277°K) but unknown for EV-A71. Both temperature and pH are significant variables implicated in the uncoating process in members of the genus *Enterovirus*^{39,52,89,372} and may explain the differing spatial alignments of not only the VP1, but also possibly the VP4. Substantial mobility of the N-terminal α -helix was observed in the long time-scale simulations performed for poliovirus in chapter 4 and these observations were published in 2012².

Taking into account the observation of the spatial variation of the internal structures of the capsid, it is prudent to investigate whether these variations impact the structures associated with receptor binding, neutralising antibody binding sites or antiviral interaction sites. Specific sites of interest for receptor or neutralising antibody binding are the canyon and flanking regions defined by the BC and EF loops of VP1, and the EF loop of VP2^{75,76}. Investigation of pentamers, empty capsids and complete virion forms of

poliovirus were undertaken in chapter 4, and no obvious deviation from crystal structure was observed, with the exception of movement in the solvent accessible aspects of the BC and EF loops. This observation is expected as part of the normal movement of these structures in an aqueous environment. Importantly, the hydrophobic pocket representing the binding site of a number of antiviral drugs which inhibit virus entry into cells and the uncoating process, were observed to maintain their integrity during the MD simulation process when compared to crystal structures as described in section 5.4.1.2 (Figures 5.9 and 5.10). The examination of this site in EV-A71 indicated a correlation between amino acid side-chain placement for the predicted model and B-factors obtained during x-ray diffraction, indicating that the protein comparative modelling methods combined with MD simulation are appropriate for the examination of this structure.

Quality assessments for the predicted models for EV-C96 and EV-A120 capsids indicated that the predicted models were of good quality and that the simulations of these capsid structures were stable, supporting the use of the models for future research to examine the contribution of structural motifs to receptor binding, neutralising antibody binding, or antiviral-drug interactions.

For EV-A120, significant difficulty was experienced deriving the genetic sequence, due to the virus not growing in culture. The specimen in question was extraordinarily toxic to any cells used in cell culture and molecular detection methods proved to be the only possible method to identify this virus. It is of interest to investigate the possibility of cultivating this virus for further research, and given the relatively small amount of inoculum available, it is critical to select a cell line that best suits the tropism of the virus. Examination of a motif found on the canyon wall common to members of the species A *Enteroviruses* for which the cellular receptor SCARB2 is suspected to bind, was found to be conserved in EV-A120, indicating that cell lines expressing SCARB2 may be appropriate to the cultivation of EV-A120. Research efforts into this line of investigation are ongoing.

With respect to the topographical differences observed between predicted virus models for EV-C96 and EV-A120, it was expected that they would be virtually identical to the viruses on which they had been modelled. This was for the most part the case for EV-C96, with only modest variation in structures defining the canyon and a more punctuated BC-loop when compared to the poliovirus crystal structure used for threading.

EV-A120 exhibited a significant and unexpected morphological difference to EV-A71, even though it shared a higher amino acid percentage identity when compared to the predicted EV-C96 model (68.5% for EV-C96 versus 72.1% for EV-A120). These variations may be due to artefacts as a result of interactions of the solvent accessible loop structures encountered during the crystallisation process, or possibly the observed differences in available chain data representing the C-terminal aspect of VP1. Also of note, is that (with the exception of the EV-C96 and EV-A120 predicted models) the structures depicted in Figure 5.20 are derived exclusively from x-ray diffraction data, which represents data collected from crystals that are not representative of a biological environment. The simulation data from which the predicted EV-C96 and EV-A120 models were derived were designed to emulate an aqueous 37°C saline environment, which may have allowed a freedom of movement not necessarily experienced in the crystalline lattice used for x-ray diffraction.

In summary, this chapter has investigated the use of comparative protein modelling methods incorporating manual threading of amino acid sequences on template structures, with amino acid identities as low as 43% for protomers (EV-A71) and 32.1% for single amino acid chains (VP1 of EV-A71). These values can be considered to be approaching the minimum sequence identity acceptable for effective comparative modelling methods. The outcome of the investigation showed that even when reconstructing models with such low amino acid identities, accurate and precise models were produced for further research. Accordingly, the methods were applied to newly-described and novel viruses for which no experimentally derived structures exist. This will allow the use of these models in the investigation of receptor binding motifs, neutralising antibody binding sites and the examination of antiviral drug binding³⁷³. Importantly, these models represent a starting point for integration with experimentally derived data such as cryo-EM density data using methods such as the MD flexible fitting process^{241,374,375}.

Chapter Six

Concluding Remarks and Future Research

6.1 Manuscripts Published in Association with this Chapter

- Liu, H., **Roberts J. A.**, Moore, D., Anderson, B., Pallansch, M. A., Pevear, D. C., Collett M. S., Oberste, M. S., 2012. Characterization of Poliovirus Variants Selected for Resistance to the Antiviral Compound V-073 *Antimicrobial Agents and Chemotherapy*, 56 (11): 5568-5574. [Refer to appendix A8.2.7].

6.2 Concluding Remarks

The primary aim of the research for this thesis was to examine ways to integrate a range of statistical, bioinformatics and computational biophysical approaches, and to exploit the massive increase in computational power available to researchers and public health professionals for the detection, surveillance and study of disease causing agents, specifically viruses of the genus *Enterovirus*.

The approaches taken to incorporate computational methods included sequence database searching and retrieval, multiple sequence alignment, and phylogenetic analysis using a number of different algorithms to analyse viruses at the nucleic acid and amino acid level. Additional to the incorporation of bioinformatics approaches was the inclusion of 3-dimensional structure prediction and analysis, using comparative protein modelling and atomistic MD simulation to reconstruct viruses of interest. In essence, this body of work examines the utility of taking data retrieved from clinical samples in a 2-dimensional format (sequences) and using computational biophysics, to convert them to a 3-dimensional structure, then to incorporate a 4th dimension, time, using MD methods.

These approaches were only possible as a result of (i) the exponential increase in computational capacity both at the institutional level (supercomputer) and locally (personal computer), combined with (ii) easy access to data in the form of thousands of sequences via GenBank and (iii) the increase in freely available software packages designed specifically for bioinformatics applications, MD simulation and structure visualisation. The software packages allowed a pragmatic approach to the research questions to be undertaken, resulting in a stepwise methodology with practical applications to any disease surveillance laboratory environment (Appendices A5.9, A5.10).

In summary, the research incorporated leading-edge computational methodologies and resulted in the following outcomes:

- i.) Utilised bioinformatics methods to improve the detection of enteroviruses in clinical samples, by examining hundreds of available sequences to refine the detection of human enteroviruses, as distinct from human rhinoviruses. This resulted in a novel real-time reverse-transcription PCR method which streamlined the detection and differentiation of *Enterovirus* species C including poliovirus, from *Enterovirus* species A and B, which is applicable to the WHO polio eradication initiative.
- ii.) Used simple phylogenetic methods and more complex, computationally intensive algorithms which incorporated temporal and geographic variables, to examine the chains of transmission for specific enteroviruses within Australia and the Western Pacific region.
- iii.) Analysed enterovirus serotypes identified within Australia and the Western Pacific region that represent a significant or potential threat to public health; specifically EV-A71 and EV-C96.
- iv.) Discovered a novel enterovirus, EV-A120, from a patient suffering AFP. The virus is closely related to CV-A7 and was detected and sequenced direct from faecal extract as it was unable to be grown in culture. More than 2,500 base pairs of sequence were determined by primer walking.
- v.) Developed an atomistic MD simulation approach to the reconstruction of a complete poliovirus virion, including RNA, utilising x-ray diffraction data of the capsid. The methodology was used to reconstruct CV-A21 and EV-A71 models of the capsids from atomic coordinate data.
- vi.) Validated the use of comparative protein modelling methods combined with atomistic MD simulation methods to predict enterovirus structure, using prototype sequences and 3-dimensional threading methods. The use of these methods indicated that comparative protein modelling methods are an accurate and precise way of reproducing 3-dimensional structures from 2-dimensional sequence data.
- vii.) Applied bioinformatics and computational biophysics approaches to the detection, characterization and *in-silico* reconstruction of EV-A71, EV-C96 and novel virus EV-A120. The methodologies developed in this thesis enabled atomistic MD simulations to be performed of EV-C96 and EV-A120, by using atomic coordinate data derived from related virus structures.

6.3 Future Research

The workflows established in this thesis can enable information derived from clinical samples to be used to create a functional *in-silico* virus model for further research. Examination of anti-viral drug binding events, receptor binding sites, and major conformational shifts in virus structure can be assessed using these models.

As the technologies advance, it will become faster and easier to obtain genetic sequence data from pathogens, and rapid dissemination of these data will allow real-time observation of disease emergence. Theoretically, the combination of increased computational power and more advanced simulation methods would allow reconstruction of pathogens, using comparative protein modelling techniques to “build” an *in-silico* copy of the pathogen or possibly even a component that is a known target for existing therapeutics. This process could in the future be done locally on a personal computer, as shown by the simulation of EV-A71 on a MicroSD card running a virtual operating system on a desktop workstation (section 5.4.1.2).

These experiments could allow the examination of the drug binding site and any known mutations, by performing *in-silico* experiments using advanced MD methods, to determine the effectiveness of a range of therapeutic options for specific viruses and any mutants that may arise over time. The examination of drug resistance mechanisms in viruses, specifically influenza, have recently been examined using free energy perturbation methods, exploring the mechanism that confers resistance to Oseltamivir but not Relenza³⁷⁶. This particular method may be useful when applied to enteroviruses and the examination of the hydrophobic pocket with existing capsid binding anti-virals such as V-073, BTA-798 or other WIN compound derivatives.

An example of the application of atomistic MD simulation to the examination of this hydrophobic pocket, can be found in work examining anti-viral drug resistance in vaccine derived polioviruses³⁷³ (Appendix A8.2.7). In this work, fourteen simulations representing V-073 resistant and sensitive vaccine derived polioviruses were constructed representing pentamer configurations of the virus proteins, VP1 VP2 and VP3 and VP4. Positioning of the anti-viral compound within each protomer was performed in comparison with the known crystal structure of V-073 complexed with poliovirus type 2. After 10 ns of simulation time the pentamers underwent the protomer fit methodology to examine the positioning of point mutants that confer resistance.

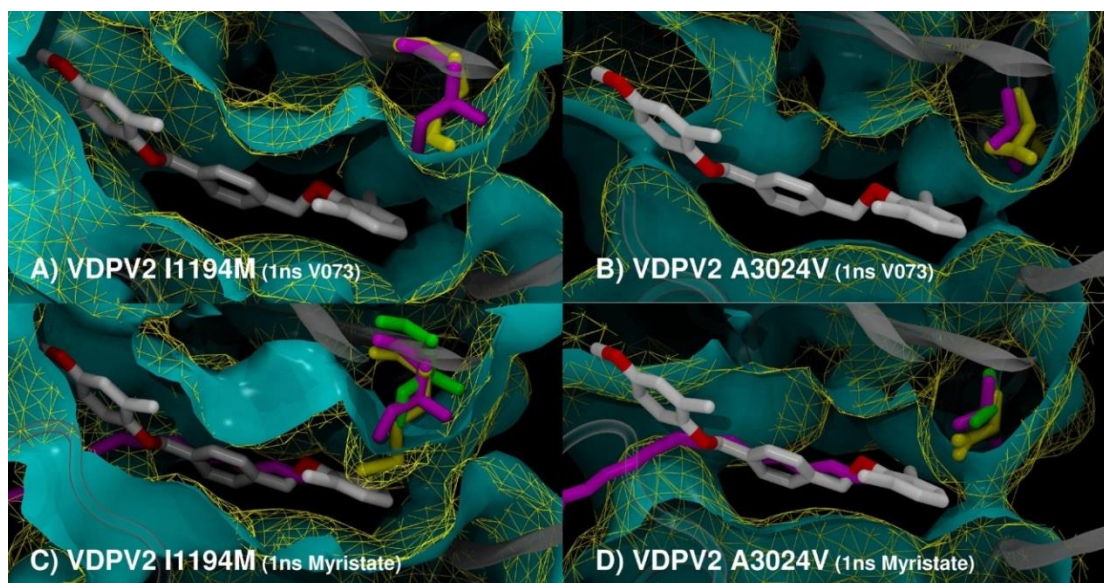


Figure 6.1 Vaccine Derived Poliovirus type 2 with Myristate as the pocket factor. Simulations performed in the absence of drug, substituted with a single myristate molecule occupying a hydrophobic pocket in order to examine conformational distortion. The figure indicates the location in the poliovirus type 2 structure of predominant amino acids substitutions that confer resistance to V-073, based on the crystallographic structure of V-073 bound to poliovirus type 2 (Lentz *et al.*, 1997)³⁷⁷, PDB file 1EAH. Superolateral cut-away view of electron density maps showing V073 bound within the hydrophobic pocket of VP1. Density mapping at 1 Å resolution and 2 Å radius representing the original parental strain (cyan) with a wireframe overlay (yellow) of the resistant virus. A) and B) original results for VDPV2 as presented by Liu *et al.*, (2012)³⁷³, amino acid differences are shown as a liquorice representation with parental (magenta) and variant (yellow), C) and D). repeat of the VDPV2 simulation experiments with myristate in the hydrophobic pocket, amino acid point mutations are shown as liquorice representations with parental (magenta) and variant (yellow), the original simulation results with V-073 in-situ are indicated in Green. Note the differences in yellow wireframe configurations between figures A and C, and between figures B and D.

In order to further examine the potential for free energy perturbation methods to predict anti-viral drug resistance, preliminary experiments involving atomistic MD simulations were performed based on the previous work described above, but instead of using the drug bound in the hydrophobic pocket, myristic acid was used in place of the drug. This lipid is a common element found within members of the genus *Enterovirus*, and allowed the examination of pocket deformation in the presence of a hydrophobic small molecule. Given that this thesis has examined the binding effect of small anti-viral (capsid binding) molecule and established that there is an observable effect on the plasticity of the capsid (section 4.4.6), it stands to reason that the placement of a hydrophobic non-binding small

molecule would elucidate any pocket deformation more effectively than any bound anti-viral which inhibits structural deformation (Figure 6.1).

The observed distortion of the hydrophobic pocket described in figure 6.1 indicates that free energy perturbation methods, specifically alchemical approaches, would be suitable to examine the potential for binding capacity of existing enterovirus capsid binding anti-virals. Free energy perturbation approaches and normal-modes analysis methods have historically been used to examine anti-viral drug binding in members of the genus *Enterovirus*, as described in section 1.5.1 of this thesis. The difference now is that experiments of this type no longer require the use of institutional supercomputers. Multiple experiment instances can now be performed on a relatively modest personal computer.

An exciting potential for the advancements revealed by the research performed in this thesis, indicates that virus models created using comparative protein modelling methods combined with atomistic MD simulation are both accurate and precise. This allows the application of molecular dynamics flexible fitting methods (MDFF) to be used in the application and refinement of experimentally derived cryo-EM density data. An MDFF approach has been used to examine the capsid structure of the HIV capsid in one of the largest atomistic MD simulations ever performed to date at 64 million atoms²⁴¹. This advance is particularly relevant in the context of examining emerging pathogens at the atomic scale, using workflows such as those established in this thesis.

Given the above, it is further theoretically possible to monitor an outbreak in real time, to generate relevant virus structures, and to determine anti-viral resistance in order to enhance the clinical approach and the use of the correct anti-viral drug. In the event that anti-viral resistance occurs during the course of the outbreak, molecular modelling can be performed in response to amino acid sequence changes and the underlying antiviral mechanisms can be determined, enabling the potential to recommend alternative anti-viral drug therapies to combat the emergence of resistant phenotypes. This would revolutionise the clinical approach to the emergence of disease in humans with a broader applicability to agriculture.

This thesis has successfully demonstrated the use of comparative modelling and all-atom MD to model full virions under simulated physiological conditions. Additionally, this

this thesis has demonstrated a range of preliminary, simple analyses on virion MD trajectories, forming a platform for the development of new analysis methods and protocols. The sheer size of whole virions means that analysis methodologies must be developed which ensure efficient use of emerging supercomputing capabilities. One obvious future direction has been described above - the study of drug-virion interactions. But more subtle, yet possibly important, behaviours of the virion could potentially be brought to light. These include (i) the examination of tectonic movement across virus capsids, (ii) the derivation of potential of mean force profiles for putative ion channels traversing virus capsids and (iii) the examination of vibrational states of virus capsids via the extraction of large-scale (slow) motions and small-scale (localised, rapid) motions for native virions, empty capsids in concert with bound antivirals or receptor even molecules. The scope for the examination of the trajectory data derived from the *in-silico* experiments performed in this research work is vast and will require a pragmatic approach and significant computational resources. Given the rate at which computer power is evolving, this is an exciting proposition.

References

- 1 Freddolino, P. L., Arkhipov, A. S., Larson, S. B., McPherson, A. & Schulten, K. Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure* **14**, 437-449 (2006).
- 2 Roberts, J. A., Kuiper, M. J., Thorley, B. R., Smooker, P. M. & Hung, A. Investigation of a predicted N-terminal amphipathic α -helix using atomistic molecular dynamics simulation of a complete prototype poliovirus virion. *Journal of Molecular Graphics and Modelling* **38**, 165-173 (2012).
- 3 Wimmer, E. The test-tube synthesis of a chemical called poliovirus. *EMBO reports* **7**, S3-S9 (2006).
- 4 Kozma, C. Skeletal dysplasia in ancient Egypt. *American Journal of Medical Genetics Part A* **146**, 3104-3112 (2008).
- 5 Faraj, A. A. Poliomyelitis: orthopaedic management. *Current Orthopaedics* **20**, 41-46 (2006).
- 6 Hughes, S. *The Virus. A History of the Concept*. (Taylor & Francis, 1977).
- 7 Iwanowski, D. Über die Mosaikkrankheit der Tabakspflanze. *Bulletin Scientifique publié par l'Académie Impériale des Sciences de Saint-Pétersbourg* **35**, 67-70 (1892).
- 8 Waksman, S. A. What is an antibiotic or an antibiotic substance? *Mycologia* **39**, 565-569 (1947).
- 9 Weller, T. H. & Enders, J. F. in *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, NY)*. 124-128 (Royal Society of Medicine).
- 10 Chonmaitree, T., Ford, C., Sanders, C. & Lucia, H. Comparison of cell cultures for rapid isolation of enteroviruses. *Journal of clinical microbiology* **26**, 2576-2580 (1988).
- 11 Dalldorf, G. The coxsackie virus group. *Annals of the New York Academy of Sciences* **56**, 583-586 (1953).
- 12 Robbins, F., Enders, J., Weller, T. & Florentino, G. Studies on the cultivation of poliomyelitis viruses in tissue culture. V. The direct isolation and serologic identification of virus strains in tissue culture from patients with nonparalytic and paralytic poliomyelitis. *American journal of hygiene* **54**, 286 (1951).
- 13 Stanley, W. M. Isolation of a crystalline protein possessing the properties of tobacco-mosaic virus. *Science* **81**, 644-645 (1935).
- 14 Bernal, J. & Fankuchen, I. X-Ray and Crystallographic Studies of Plant Virus Preparations. III. *The Journal of general physiology* **25**, 147-165 (1941).
- 15 Rossmann, M. G. *et al.* Structure of a human common cold virus and functional relationship to other picornaviruses. *nature* **317**, 145-153 (1984).
- 16 Hogle, J., Chow, M. & Filman, D. Three-dimensional structure of poliovirus at 2.9 Å resolution. *Science* **229**, 1358 (1985).
- 17 Nathanson, N. & Martin, J. R. The epidemiology of poliomyelitis: enigmas surrounding its appearance, epidemicity, and disappearance. *American Journal of Epidemiology* **110**, 672 (1979).
- 18 Dauer, C. C. The changing age distribution of paralytic poliomyelitis. *Annals of the New York Academy of Sciences* **61**, 943-955 (1955).
- 19 Bodian, D., Morgan, I. M. & A, H. H. Differentiation of Types of Poliomyelitis Viruses III. The Grouping of Fourteen Strains Into Three Basic Immunological Types. *American Journal of Epidemiology* **49**, 234-247 (1949).

- 20 Burnet, F. & Macnamara, J. Immunological differences between strains of poliomyelitic virus. *British journal of experimental pathology* **12**, 57 (1931).
- 21 Sabin, A. B. Oral, live poliovirus vaccine for elimination of poliomyelitis. *Archives of Internal Medicine* **106**, 5 (1960).
- 22 Salk, J. E. Poliomyelitis vaccine in the fall of 1955. *American Journal of Public Health and the Nations Health* **46**, 1-14 (1956).
- 23 Khetsuriani, N., LaMonte-Fowlkes, A., Oberst, S. & Pallansch, M. A. Enterovirus surveillance—United States, 1970–2005. *MMWR Surveill Summ* **55**, 1-20 (2006).
- 24 Shindarov, L. *et al.* Epidemiological, clinical, and pathomorphological characteristics of epidemic poliomyelitis-like disease caused by enterovirus 71. *Journal of hygiene, epidemiology, microbiology, and immunology* **23**, 284 (1979).
- 25 Chan, L. *et al.* Deaths of children during an outbreak of hand, foot, and mouth disease in Sarawak, Malaysia: clinical and pathological characteristics of the disease. *Clinical infectious diseases* **31**, 678-683 (2000).
- 26 Gorgievski-Hrisoho, M., Schumacher, J.-D., Vilimonovic, N., Germann, D. & Matter, L. Detection by PCR of enteroviruses in cerebrospinal fluid during a summer outbreak of aseptic meningitis in Switzerland. *Journal of clinical microbiology* **36**, 2408-2412 (1998).
- 27 Brown, B. A., Oberste, M. S., Alexander, J. P., Kennett, M. L. & Pallansch, M. A. Molecular epidemiology and evolution of enterovirus 71 strains isolated from 1970 to 1998. *Journal of virology* **73**, 9969-9975 (1999).
- 28 Strikas, R. A., Anderson, L. J. & Parker, R. A. Temporal and geographic patterns of isolates of nonpolio enterovirus in the United States, 1970–1983. *Journal of infectious diseases* **153**, 346-351 (1986).
- 29 Irvine, D., Irvine, A. & Gardner, P. Outbreak of ECHO virus type 30 in a general practice. *British Medical Journal* **4**, 774 (1967).
- 30 Louie, J. K. *et al.* Rhinovirus outbreak in a long term care facility for elderly persons associated with unusually high mortality. *Clinical infectious diseases* **41**, 262-265 (2005).
- 31 Kiang, D. *et al.* Molecular characterization of a variant rhinovirus from an outbreak associated with uncommonly high mortality. *Journal of clinical virology* **38**, 227-237 (2007).
- 32 Kohler, K. A., Banerjee, K., Gary Hlady, W., Andrus, J. K. & Sutter, R. W. Vaccine-associated paralytic poliomyelitis in India during 1999: decreased risk despite massive use of oral polio vaccine. *Bulletin of the World Health Organization* **80**, 210-216 (2002).
- 33 Kitamura, N. *et al.* Primary structure, gene organization and polypeptide expression of poliovirus RNA. (1981).
- 34 Racaniello, V. R. & Baltimore, D. Molecular cloning of poliovirus cDNA and determination of the complete nucleotide sequence of the viral genome. *Proceedings of the National Academy of Sciences* **78**, 4887-4891 (1981).
- 35 Racaniello, V. R. & Baltimore, D. Cloned poliovirus complementary DNA is infectious in mammalian cells. *Science* **214**, 916-919 (1981).
- 36 Cello, J., Paul, A. V. & Wimmer, E. Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. *science* **297**, 1016-1018 (2002).
- 37 Adams, M., King, A. & Carstens, E. Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2013). *Archives of virology* **158**, 2023-2030 (2013).

- 38 King, A. M., Adams, M. J., Lefkowitz, E. J. & Carstens, E. B. *Virus taxonomy: IXth report of the International Committee on Taxonomy of Viruses*. Vol. 9 (Access Online via Elsevier, 2011).
- 39 Hughes, J. H., Thomas, D. C. & Hamparian, V. V. in *Proceedings of the Society for Experimental Biology and Medicine*. Society for Experimental Biology and Medicine (New York, NY). 555-560 (Royal Society of Medicine).
- 40 Stott, E. & Killington, R. Rhinoviruses. *Annual Reviews in Microbiology* **26**, 503-524 (1972).
- 41 Lennette, D. & Schmidt, N. General principles for laboratory diagnosis of viral, rickettsial, and chlamydial infections. *Diagnostic procedures for viral, rickettsial, and chlamydial infections, 7th ed.* American Public Health Association, Washington, DC, 7-14 (1995).
- 42 Rotbart, H. A. Enteroviral infections of the central nervous system. *Clinical infectious diseases* **20**, 971-981 (1995).
- 43 Lau, S. K. *et al.* Detection of human rhinovirus C in fecal samples of children with gastroenteritis. *Journal of Clinical Virology* **53**, 290-296 (2012).
- 44 Oberste, M. S. *et al.* Enteroviruses 76, 89, 90 and 91 represent a novel group within the species Human enterovirus A. *Journal of general virology* **86**, 445-451 (2005).
- 45 Oberste, M. S. *et al.* Molecular identification and characterization of two proposed new enterovirus serotypes, EV74 and EV75. *Journal of general virology* **85**, 3205-3212 (2004).
- 46 Norder, H. *et al.* Sequencing of 'untypable' enteroviruses reveals two new types, EV-77 and EV-78, within human enterovirus type B and substitutions in the BC loop of the VP1 protein for known types. *Journal of general virology* **84**, 827-836 (2003).
- 47 Tracy, S. *et al.* Molecular approaches to enteroviral diagnosis in idiopathic cardiomyopathy and myocarditis. *Journal of the American College of Cardiology* **15**, 1688-1694 (1990).
- 48 Weiss, L. M., Liu, X.-F., Chang, K. L. & Billingham, M. E. Detection of enteroviral RNA in idiopathic dilated cardiomyopathy and other human cardiac tissues. *Journal of Clinical Investigation* **90**, 156 (1992).
- 49 Brown, J. *et al.* Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses. *Part II-single stranded DNA viruses*. Elsevier, 359-373 (2012).
- 50 Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution* **28**, 2731-2739 (2011).
- 51 Hasegawa, M., Kishino, H. & Yano, T.-a. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of molecular evolution* **22**, 160-174 (1985).
- 52 Racaniello, V. R. in *Fields virology, 6th Edition*. Vol. 1 (ed David M.; Howley Knipe, Peter M.) Ch. 16, 453-489 (Lippincott Williams & Wilkins, 2013).
- 53 Yeates, T. O., Thompson, M. C. & Bobik, T. A. The protein shells of bacterial microcompartment organelles. *Current opinion in structural biology* **21**, 223-231 (2011).
- 54 Lee, Y. F., Nomoto, A., Detjen, B. M. & Wimmer, E. A protein covalently linked to poliovirus genome RNA. *Proceedings of the National Academy of Sciences* **74**, 59-63 (1977).

- 55 Flanagan, J. B., Petterson, R., Ambros, V., Hewlett, N. & Baltimore, D. Covalent linkage of a protein to a defined nucleotide sequence at the 5'-terminus of virion and replicative intermediate RNAs of poliovirus. *Proceedings of the National Academy of Sciences* **74**, 961-965 (1977).
- 56 Chow, M. *et al.* Myristylation of picornavirus capsid protein VP4 and its structural significance. *Nature* **327**, 482-486 (1987).
- 57 Larsen, G. R., Anderson, C. W., Dorner, A. J., Semler, B. L. & Wimmer, E. Cleavage sites within the poliovirus capsid protein precursors. *Journal of virology* **41**, 340-344 (1982).
- 58 Palmenberg, A. C. Picornaviral processing: some new ideas. *Journal of cellular biochemistry* **33**, 191-198 (1987).
- 59 Hewlett, M. J., Rose, J. K. & Baltimore, D. 5'-terminal structure of poliovirus polyribosomal RNA is pUp. *Proceedings of the National Academy of Sciences* **73**, 327-330 (1976).
- 60 Nomoto, A., Lee, Y. F. & Wimmer, E. The 5' end of poliovirus mRNA is not capped with m7G (5') ppp (5') Np. *Proceedings of the National Academy of Sciences* **73**, 375-380 (1976).
- 61 Macejak, D. G. & Sarnow, P. Internal initiation of translation mediated by the 5' leader of a cellular mRNA. *Nature* **353**, 90-94 (1991).
- 62 Rivera, V. M., Welsh, J. D. & Maizel Jr, J. V. Comparative sequence analysis of the 5' noncoding region of the enteroviruses and rhinoviruses. *Virology* **165**, 42-50 (1988).
- 63 Skinner, M. A. *et al.* New model for the secondary structure of the 5' non-coding RNA of poliovirus is supported by biochemical and genetic data that also show that RNA secondary structure is important in neurovirulence. *Journal of molecular biology* **207**, 379-392 (1989).
- 64 Summers, D. F. & Maizel Jr, J. V. Evidence for large precursor proteins in poliovirus synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **59**, 966 (1968).
- 65 Roberts, J. A. & Thorley, B. R. in *PCR for Clinical Microbiology* (eds Margret Schuller *et al.*) Ch. 32 Enterovirus, 229-233 (Springer Netherlands, 2010).
- 66 Curry, S. *et al.* Dissecting the roles of VP0 cleavage and RNA packaging in picornavirus capsid stabilization: the structure of empty capsids of foot-and-mouth disease virus. *Journal of virology* **71**, 9743-9752 (1997).
- 67 Hindiyeh, M., Li, Q.-H., Basavappa, R., Hogle, J. M. & Chow, M. Poliovirus mutants at histidine 195 of VP2 do not cleave VP0 into VP2 and VP4. *Journal of virology* **73**, 9072-9079 (1999).
- 68 Basavappa, R. *et al.* Role and mechanism of the maturation cleavage of VP0 in poliovirus assembly: structure of the empty capsid assembly intermediate at 2.9 Å resolution. *Protein Science* **3**, 1651-1669 (1994).
- 69 Stanway, G. & Hyypiä, T. Parechoviruses. *Journal of virology* **73**, 5249-5254 (1999).
- 70 Law, R. H. *et al.* The structural basis for membrane binding and pore formation by lymphocyte perforin. *Nature* **468**, 447-451 (2010).
- 71 De Colibus, L. *et al.* Structures of lysenin reveal a shared evolutionary origin for pore-forming proteins and its mode of sphingomyelin recognition. *Structure* **20**, 1498-1507 (2012).
- 72 Bamford, D. H., Grimes, J. M. & Stuart, D. I. What does structure tell us about virus evolution? *Current opinion in structural biology* **15**, 655-663 (2005).

- 73 Benson, S. D., Bamford, J. K., Bamford, D. H. & Burnett, R. M. Does common architecture reveal a viral lineage spanning all three domains of life? *Molecular cell* **16**, 673-685 (2004).
- 74 Ren, R., Moss, E. G. & Racaniello, V. R. Identification of two determinants that attenuate vaccine-related type 2 poliovirus. *Journal of virology* **65**, 1377-1382 (1991).
- 75 Murdin, A. D. & Wimmer, E. Construction of a poliovirus type 1/type 2 antigenic hybrid by manipulation of neutralization antigenic site II. *Journal of virology* **63**, 5251-5257 (1989).
- 76 Wien, M. W. *et al.* Structure of the complex between the Fab fragment of a neutralizing antibody for type 1 poliovirus and its viral epitope. *Nature Structural & Molecular Biology* **2**, 232-243 (1995).
- 77 Arnold, E. & Rossmann, M. G. Analysis of the structure of a common cold virus, human rhinovirus 14, refined at a resolution of 3.0 Å. *Journal of molecular biology* **211**, 763-801 (1990).
- 78 Pevear, D. C., Tull, T. M., Seipel, M. E. & Groarke, J. M. Activity of pleconaril against enteroviruses. *Antimicrobial agents and chemotherapy* **43**, 2109-2115 (1999).
- 79 Zhang, A., Nanni, R., Oren, D., Rozhon, E. & Arnold, E. in *Seminars in Virology*. 453-471.
- 80 Pallansch, M., Oberste, M. S. & Whitton, J. L. in *Fields Virology, 6th Edition*. Vol. 1 (ed David M.; Howley Knipe, Peter M.) Ch. 17, 490-530 (Lippincott Williams & Wilkins, 2013).
- 81 Kono, R. Apollo 11 disease or acute hemorrhagic conjunctivitis: a pandemic of a new enterovirus infection of the eyes. *American journal of epidemiology* **101**, 383-390 (1975).
- 82 Westrop, G. *et al.* Genetic basis of attenuation of the Sabin type 3 oral poliovirus vaccine. *Journal of virology* **63**, 1338-1344 (1989).
- 83 Ren, R. & Racaniello, V. R. Poliovirus spreads from muscle to the central nervous system by neural pathways. *Journal of Infectious Diseases* **166**, 747-752 (1992).
- 84 Bodain, D. in *Viral and rickettsial infections of man* (eds TM Rivers & FL Horsfall Jr) Ch. Poliomyelitis: pathogenesis and histopathology, 479-518 (Lippincott, 1955).
- 85 Ohka, S., Nihei, C.-i., Yamazaki, M. & Nomoto, A. Poliovirus trafficking toward central nervous system via human poliovirus receptor-dependent and-independent pathway. *Frontiers in microbiology* **3** (2012).
- 86 Bubeck, D., Filman, D. J. & Hogle, J. M. Cryo-electron microscopy reconstruction of a poliovirus-receptor-membrane complex. *Nature structural & molecular biology* **12**, 615-618 (2005).
- 87 Mendelsohn, C. L., Wimmer, E. & Racaniello, V. R. Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. *Cell* **56**, 855-865 (1989).
- 88 Ahvenainen, P. *et al.* Nanoscale structural studies of cellulose using X-ray methods.
- 89 Hogle, J. M. Poliovirus cell entry: common structural themes in viral cell entry pathways. *Annual review of microbiology* **56**, 677 (2002).
- 90 Bubeck, D. *et al.* The structure of the poliovirus 135S cell entry intermediate at 10-angstrom resolution reveals the location of an externalized polypeptide that binds to membranes. *Journal of virology* **79**, 7745-7755 (2005).

- 91 Levy, H. C., Bostina, M., Filman, D. J. & Hogle, J. M. Catching a virus in the act
of RNA release: a novel poliovirus uncoating intermediate characterized by cryo-
electron microscopy. *Journal of virology* **84**, 4426 (2010).
- 92 Pelletier, J. & Sonenberg, N. Internal initiation of translation of eukaryotic mRNA
directed by a sequence derived from poliovirus RNA. *Nature* **334**, 320-325
(1988).
- 93 Chen, C.-y. & Sarnow, P. Initiation of protein synthesis by the eukaryotic
translational apparatus on circular RNAs. *Science*, 415-415 (1995).
- 94 Ambros, V., Pettersson, R. F. & Baltimore, D. An enzymatic activity in
uninfected cells that cleaves the linkage between poliovirion RNA and the 5'
terminal protein. *Cell* **15**, 1439-1446 (1978).
- 95 Hsu, N.-Y. *et al.* Viral reorganization of the secretory pathway generates distinct
organelles for RNA replication. *Cell* **141**, 799-811 (2010).
- 96 Minor, P. D. The molecular biology of poliovaccines. *Journal of General
Virology* **73**, 3065-3077 (1992).
- 97 Khetsuriani, N. *et al.* Persistence of vaccine-derived polioviruses among
immunodeficient persons with vaccine-associated paralytic poliomyelitis. *Journal
of Infectious Diseases* **188**, 1845-1852 (2003).
- 98 Trask, J. D., Vignec, A. & Paul, J. R. Poliomyelitis virus in human stools. *Journal
of the American Medical Association* **111**, 6-1 (1938).
- 99 Landsteiner, K. & Popper, E. Mikroskopische Präparate von einem menschlichen
und zwei Affenrückenmarken. *Wien Klin Wochenschr* **21**, 1830 (1908).
- 100 Enders, J. F., Weller, T. H. & Robbins, F. C. Cultivation of the Lansing strain of
poliomyelitis virus in cultures of various human embryonic tissues. *Science* **109**,
85-87 (1949).
- 101 Storch, G. A. & Wang, D. in *Fields Virology* Vol. 1 (ed David M.; Howley
Knipe, Peter M.) Ch. 15, 414-451 (Lippincott Williams & Wilkins, 2013).
- 102 Lim, K. & Benyesh-Melnick, M. Typing of viruses by combinations of antiserum
pools. Application to typing of enteroviruses (Coxsackie and ECHO). *The Journal
of Immunology* **84**, 309-317 (1960).
- 103 Van Doornum, G. & De Jong, J. Rapid shell vial culture technique for detection of
enteroviruses and adenoviruses in fecal specimens: comparison with conventional
virus isolation method. *Journal of clinical microbiology* **36**, 2865-2868 (1998).
- 104 Olive, D. M. *et al.* Detection and differentiation of picornaviruses in clinical
samples following genomic amplification. *Journal of General Virology* **71**, 2141-
2147 (1990).
- 105 Zoll, G. *et al.* General primer-mediated polymerase chain reaction for detection of
enteroviruses: application for diagnostic routine and persistent infections. *Journal
of clinical microbiology* **30**, 160 (1992).
- 106 Verstrepen, W. A., Kuhn, S., Kockx, M. M., Van De Vyvere, M. E. & Mertens, A.
H. Rapid detection of enterovirus RNA in cerebrospinal fluid specimens with a
novel single-tube real-time reverse transcription-PCR assay. *Journal of clinical
microbiology* **39**, 4093-4096 (2001).
- 107 Espy, M. *et al.* Real-time PCR in clinical microbiology: applications for routine
laboratory testing. *Clinical microbiology reviews* **19**, 165-256 (2006).
- 108 McIver, C. J. *et al.* Development of multiplex PCRs for detection of common viral
pathogens and agents of congenital infections. *Journal of Clinical Microbiology*
43, 5102-5110 (2005).
- 109 Mullis, K. B. & Faloona, F. A. Specific synthesis of DNA in vitro via a
polymerase-catalyzed chain reaction. *Methods in enzymology* **155**, 335 (1987).

- 110 Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74**, 5463-5467 (1977).
- 111 Shi, W.-f. *et al.* Positive selection analysis of VP1 genes of worldwide human enterovirus 71 viruses. *Virologica Sinica* **24**, 59-64 (2009).
- 112 Oberste, M. S. *et al.* Typing of human enteroviruses by partial sequencing of VP1. *Journal of clinical microbiology* **37**, 1288-1293 (1999).
- 113 McIntyre, C. L., Knowles, N. J. & Simmonds, P. Proposals for the classification of human rhinovirus species A, B and C into genotypically assigned types. *Journal of General Virology* (2013).
- 114 Oberste, M. S. *et al.* Species-specific RT-PCR amplification of human enteroviruses: a tool for rapid species identification of uncharacterized enteroviruses. *Journal of general virology* **87**, 119 (2006).
- 115 Oberste, M. S., Maher, K. & Pallansch, M. A. Molecular phylogeny of all human enterovirus serotypes based on comparison of sequences at the 5' end of the region encoding VP2. *Virus research* **58**, 35-43 (1998).
- 116 Cardoso, M. J. *et al.* Molecular epidemiology of human enterovirus 71 strains and recent outbreaks in the Asia-Pacific region: comparative analysis of the VP1 and VP4 genes. *Emerging infectious diseases* **9**, 462 (2003).
- 117 Nix, W. A., Oberste, M. S. & Pallansch, M. A. Sensitive, seminested PCR amplification of VP1 sequences for direct identification of all enterovirus serotypes from original clinical specimens. *Journal of clinical microbiology* **44**, 2698-2704 (2006).
- 118 Feng, D.-F. & Doolittle, R. F. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *Journal of molecular evolution* **25**, 351-360 (1987).
- 119 Hogeweg, P. & Hesper, B. The alignment of sets of sequences and the construction of phyletic trees: an integrated method. *Journal of molecular evolution* **20**, 175-186 (1984).
- 120 Needleman, S. B. & Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of molecular biology* **48**, 443-453 (1970).
- 121 Dayhoff, M. O. & Schwartz, R. M. *A model of evolutionary change in proteins.* (Citeseer, 1978).
- 122 Henikoff, S. & Henikoff, J. G. Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences* **89**, 10915-10919 (1992).
- 123 Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* **22**, 4673-4680 (1994).
- 124 Felsenstein, J. Distance methods for inferring phylogenies: a justification. *Evolution*, 16-24 (1984).
- 125 Yang, Z. Phylogenetic analysis using parsimony and likelihood methods. *Journal of Molecular Evolution* **42**, 294-307 (1996).
- 126 Baldauf, S. L. Phylogeny for the faint of heart: a tutorial. *TRENDS in Genetics* **19**, 345-351 (2003).
- 127 Lemey, P., Rambaut, A., Drummond, A. J. & Suchard, M. A. Bayesian phylogeography finds its roots. *PLoS computational biology* **5**, e1000520 (2009).
- 128 Smith, G. J. *et al.* Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* **459**, 1122-1125 (2009).

- 129 Bielejec, F., Rambaut, A., Suchard, M. A. & Lemey, P. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* **27**, 2910-2912 (2011).
- 130 Palacios, G. & Oberste, M. Enteroviruses as agents of emerging infectious diseases. *Journal of neurovirology* **11**, 424-433 (2005).
- 131 Abbasian, F., Saberbaghi, T. & Moosapour, A. Role of Non-Poliioviruses in Acute Flaccid Paralysis (AFP). *Journal of Gastroenterology and Hepatology Research* **1** (2012).
- 132 Hobday, L. K. *et al.* Potential for the Australian and New Zealand paediatric intensive care registry to enhance acute flaccid paralysis surveillance in Australia: a data-linkage study. *BMC infectious diseases* **13**, 1-6 (2013).
- 133 Habel, K. & Loomis, L. N. Cocksackie A7 Virus and the Russian "Poliovirus Type 4". *Experimental Biology and Medicine* **95**, 597-605 (1957).
- 134 Ooi, M. H., Wong, S. C., Lewthwaite, P., Cardosa, M. J. & Solomon, T. Clinical features, diagnosis, and management of enterovirus 71. *The Lancet Neurology* **9**, 1097-1105 (2010).
- 135 Abzug, M. J. The enteroviruses: Problems in need of treatments. *Journal of Infection* **68**, S108-S114 (2014).
- 136 Horstmann, D. & Yamada, N. Enterovirus infections of the central nervous system. *Research publications-Association for Research in Nervous and Mental Disease* **44**, 236 (1968).
- 137 Stewardson, A. J. *et al.* Imported case of poliomyelitis, Melbourne, Australia, 2007. *Emerging Infectious Diseases* **15**, 63-65 (2009).
- 138 Riley JR, H. D. & Batson, R. The Poliomyelitis Patient with Respiratory Failure: Diagnosis, Management, and Results or Treatment. *Southern medical journal* **50**, 1357-1368 (1957).
- 139 Wang, S.-M. *et al.* Pathogenesis of enterovirus 71 brainstem encephalitis in pediatric patients: roles of cytokines and cellular immune activation in patients with pulmonary edema. *Journal of Infectious Diseases* **188**, 564-570 (2003).
- 140 Lum, L. *et al.* Fatal enterovirus 71 encephalomyelitis. *The Journal of pediatrics* **133**, 795-798 (1998).
- 141 Rotbart, H. A. *Human enterovirus infections.* (American Society for Microbiology (ASM), 1995).
- 142 Kaplan, M. H., Klein, S. W., McPhee, J. & Harper, R. G. Group B coxsackievirus infections in infants younger than three months of age: a serious childhood illness. *Review of Infectious Diseases* **5**, 1019-1032 (1983).
- 143 Modlin, J. F. Perinatal echovirus infection: insights from a literature review of 61 cases of serious infection and 16 outbreaks in nurseries. *Review of Infectious Diseases* **8**, 918-926 (1986).
- 144 Dominguez, S. R. *et al.* Multiplex MassTag-PCR for respiratory pathogens in pediatric nasopharyngeal washes negative by conventional diagnostic testing shows a high prevalence of viruses belonging to a newly recognized rhinovirus clade. *Journal of Clinical Virology* **43**, 219-222 (2008).
- 145 Linsuwanon, P. *et al.* High prevalence of human rhinovirus C infection in Thai children with acute lower respiratory tract disease. *Journal of Infection* **59**, 115-121 (2009).
- 146 Imamura, T. *et al.* Enterovirus 68 among children with severe acute respiratory infection, the Philippines. *Emerging infectious diseases* **17** (2011).

- 147 Mirkovic, R. *et al.* Enterovirus type 70: the etiologic agent of pandemic acute haemorrhagic conjunctivitis. *Bulletin of the World Health Organization* **49**, 341 (1973).
- 148 Wadia, N., Wadia, P., Katrak, S. & Misra, V. A study of the neurological disorder associated with acute haemorrhagic conjunctivitis due to enterovirus 70. *Journal of Neurology, Neurosurgery & Psychiatry* **46**, 599-610 (1983).
- 149 Flett, K. *et al.* Hand, foot, and mouth disease caused by coxsackievirus A6. *Emerging Infectious Diseases* **18**, 1702 (2012).
- 150 Lu, Q.-B. *et al.* Circulation of Coxsackievirus A10 and A6 in Hand-Foot-Mouth Disease in China, 2009–2011. *PloS one* **7**, e52073 (2012).
- 151 Mirand, A. *et al.* Outbreak of hand, foot and mouth disease/herpangina associated with coxsackievirus A6 and A10 infections in 2010, France: a large citywide, prospective observational study. *Clinical Microbiology and Infection* **18**, E110-E118 (2012).
- 152 Cabrerizo, M. *et al.* Molecular epidemiology of enterovirus 71, coxsackievirus A16 and A6 associated with hand, foot and mouth disease in Spain. *Clinical Microbiology and Infection* (2013).
- 153 Yeung, W.-C. G., Rawlinson, W. D. & Craig, M. E. Enterovirus infection and type 1 diabetes mellitus: systematic review and meta-analysis of observational molecular studies. *Bmj* **342** (2011).
- 154 Richardson, S., Willcox, A., Bone, A., Foulis, A. & Morgan, N. The prevalence of enteroviral capsid protein vp1 immunostaining in pancreatic islets in human type 1 diabetes. *Diabetologia* **52**, 1143-1151 (2009).
- 155 Gamble, D. & Taylor, K. Seasonal incidence of diabetes mellitus. *British medical journal* **3**, 631 (1969).
- 156 Rothwell, P. *et al.* Seasonality of birth in children with diabetes in Europe: multicentre cohort study. *BMJ: British Medical Journal* **319**, 887 (1999).
- 157 Taylor, K. W. in *Diabetes and Viruses* 101-107 (Springer, 2013).
- 158 Underwood, M. *A treatise on the diseases of children, with directions for the management of infants from the birth; especially such as are brought up by hand.* (printed for J. Mathews, 1784).
- 159 Dunn, P. Michael Underwood, MD (1737–1820): physician-accoucheur of London. *Archives of Disease in Childhood-Fetal and Neonatal Edition* **91**, F150-F152 (2006).
- 160 Sabin, A. Eradication of poliomyelitis--present status and future prospects. *Rhode Island medical journal* **45**, 453 (1962).
- 161 Thorley, B., Kelly, H. & Roberts, J. Importation of wild poliovirus into Australia, July 2007. *Communicable diseases intelligence* **31**, 299 (2007).
- 162 Kogon, A. *et al.* The Virus Watch Program: A Continuing Surveillance Of Viral Infections In Metropolitan New York Families Vii. Observations On Viral Excretion, Seroimmunity, Intra-Familial Spread And Illness Association In Coxsackie And Echovirus Infections. *American journal of epidemiology* **89**, 51-61 (1969).
- 163 Wong, S., Yip, C., Lau, S. & Yuen, K. Human enterovirus 71 and hand, foot and mouth disease. *Epidemiology and infection* **138**, 1071-1089 (2010).
- 164 Wu, Y. *et al.* The largest outbreak of hand; foot and mouth disease in Singapore in 2008: the role of enterovirus 71 and coxsackievirus A strains. *International Journal of Infectious Diseases* **14**, e1076-e1081 (2010).

- 165 Khanh, T. H., Sabanathan, S., Thanh, T. T., Thoa, L. P. K. & Thuong, T. C. Enterovirus 71-associated Hand, Foot, and Mouth Disease, Southern Vietnam, 2011. *Emerging infectious diseases* **18**, 2002 (2012).
- 166 Zhu, F.-C. *et al.* Immunogenicity and safety of an enterovirus 71 vaccine in healthy Chinese children and infants: a randomised, double-blind, placebo-controlled phase 2 clinical trial. *The Lancet* **381**, 1037-1045 (2013).
- 167 Cheng, A. *et al.* A Phase I, randomized, open-label study to evaluate the safety and immunogenicity of an enterovirus 71 vaccine. *Vaccine* **31**, 2471-2476 (2013).
- 168 Li, R. *et al.* An Inactivated Enterovirus 71 Vaccine in Healthy Children. *New England Journal of Medicine* **370**, 829-837 (2014).
- 169 Li, J.-X., Mao, Q.-Y., Liang, Z.-L., Ji, H. & Zhu, F.-C. Development of enterovirus 71 vaccines: from the lab bench to Phase III clinical trials. *Expert review of vaccines* **13**, 609-618 (2014).
- 170 Huang, M.-L. *et al.* Cross-reactive neutralizing antibody responses to enterovirus 71 infections in young children: implications for vaccine development. *PLoS neglected tropical diseases* **7**, e2067 (2013).
- 171 WHO Global Polio Eradication Initiative., <<http://www.polioeradication.org/>> (Accessed 20th March 2014.).
- 172 Control, C. f. D. & Prevention. Progress toward interruption of wild poliovirus transmission--worldwide, January 2011-March 2012. *MMWR. Morbidity and mortality weekly report* **61**, 353 (2012).
- 173 Wassilak, S. & Orenstein, W. Challenges faced by the global polio eradication initiative. *Expert review of vaccines* **9**, 447-449 (2010).
- 174 Ahmed, Q. A., Nishtar, S. & Memish, Z. A. Poliomyelitis in Pakistan: time for the Muslim world to step in. *The Lancet* **381**, 1521-1523 (2013).
- 175 Groce, N. E., Banks, L. M. & Stein, M. A. Surviving Polio in a post-polio World. *Social Science & Medicine* (2014).
- 176 Thompson, K. M. & Tebbens, R. J. D. Eradication versus control for poliomyelitis: an economic analysis. *The Lancet* **369**, 1363-1371 (2007).
- 177 Kew, O. M., Sutter, R. W., de Gourville, E. M., Dowdle, W. R. & Pallansch, M. A. Vaccine-Derived Polioviruses and the Endgame Strategy for Global Polio Eradication. *Annu. Rev. Microbiol.* **59**, 587-635 (2005).
- 178 Kapp, C. Nigerian states again boycott polio-vaccination drive. *Lancet* **363**, 709 (2004).
- 179 Wassilak, S. *et al.* Outbreak of type 2 vaccine-derived poliovirus in Nigeria: emergence and widespread circulation in an underimmunized population. *Journal of Infectious Diseases* **203**, 898-909 (2011).
- 180 Jegede, A. S. What led to the Nigerian boycott of the polio vaccination campaign? *PLoS Medicine* **4**, e73 (2007).
- 181 Diop, O. M., Burns, C. C., Wassilak, S. G. & Kew, O. M. Update on Vaccine-Derived Polioviruses — Worldwide, July 2012–December 2013. *Morbidity and Mortality Weekly Report* **61**, 242-248 (2014).
- 182 Tebbens, R. J. D. *et al.* Risks of Paralytic Disease Due to Wild or Vaccine-Derived Poliovirus After Eradication. *Risk Analysis* **26**, 1471-1505 (2006).
- 183 DeVries, A. S. *et al.* Vaccine-derived poliomyelitis 12 years after infection in Minnesota. *New England Journal of Medicine* **364**, 2316-2323 (2011).
- 184 Manor, Y. *et al.* Intensified environmental surveillance supporting the response to wild poliovirus type 1 silent circulation in Israel, 2013. *Euro Surveill* **19** (2014).

- 185 Dowdle, W. R., De Gourville, E., Kew, O. M., Pallansch, M. A. & Wood, D. J. Polio eradication: the OPV paradox. *Reviews in medical virology* **13**, 277-291 (2003).
- 186 Exploring the role of antiviral drugs in the eradication of polio. (National Research Council, 2006).
- 187 De Palma, A. M., Vliegen, I., De Clercq, E. & Neyts, J. Selective inhibitors of picornavirus replication. *Medicinal research reviews* **28**, 823-884 (2008).
- 188 Thibaut, H. J. *et al.* Towards the design of combination therapy for the treatment of enterovirus infections. *Antiviral research* **90**, 213-217 (2011).
- 189 Rhoden, E., Liu, H.-M., Wang-Chern, S.-w. & Oberste, M. S. Anti-poliovirus activity of protease inhibitor AG-7404, and assessment of in vitro activity in combination with antiviral capsid inhibitor compounds. *Antiviral research* **98**, 186-191 (2013).
- 190 Pincus, S. & Wimmer, E. Production of guanidine-resistant and-dependent poliovirus mutants from cloned cDNA: mutations in polypeptide 2C are directly responsible for altered guanidine sensitivity. *Journal of virology* **60**, 793-796 (1986).
- 191 Heinz, B. A. & Vance, L. M. The antiviral compound enviroxime targets the 3A coding region of rhinovirus and poliovirus. *Journal of virology* **69**, 4189-4197 (1995).
- 192 Rodriguez, P. L. & Carrasco, L. Gliotoxin: inhibitor of poliovirus RNA synthesis that blocks the viral RNA polymerase 3Dpol. *Journal of virology* **66**, 1971-1976 (1992).
- 193 Binford, S. *et al.* Conservation of amino acids in human rhinovirus 3C protease correlates with broad-spectrum antiviral activity of rupintrivir, a novel human rhinovirus 3C protease inhibitor. *Antimicrobial agents and chemotherapy* **49**, 619-626 (2005).
- 194 Verlinden, Y., Cuconati, A., Wimmer, E. & Rombaut, B. The antiviral compound 5-(3, 4-dichlorophenyl) methylhydantoin inhibits the post-synthetic cleavages and the assembly of poliovirus in a cell-free system. *Antiviral research* **48**, 61-69 (2000).
- 195 Collett, M. S., Neyts, J. & Modlin, J. F. A case for developing antiviral drugs against polio. *Antiviral research* **79**, 179-187 (2008).
- 196 Badger, J., Minor, I., Oliveira, M. A., Smith, T. J. & Rossmann, M. G. Structural analysis of antiviral agents that interact with the capsid of human rhinoviruses. *Proteins: Structure, Function, and Bioinformatics* **6**, 1-19 (1989).
- 197 Badger, J. *et al.* Structural analysis of a series of antiviral agents complexed with human rhinovirus 14. *Proceedings of the National Academy of Sciences* **85**, 3304-3308 (1988).
- 198 Kausche, G. A., Pfankuch, E. & Ruska, H. Die Sichtbarmachung von pflanzlichem Virus im Übermikroskop. *Naturwissenschaften (translated via Google Translate 10th February 2014)* **27**, 292-299 (1939).
- 199 Lakshmanan Govindasamy, M. A.-M. a. R. M. in *Structural Virology* Vol. 21 *RSC Biomolecular Sciences* (ed Mavis Agbandje-McKenna and Robert McKenna) Ch. Ch 6. X-ray Crystallography of Virus Capsids, (Royal Society of Chemistry, 2011).
- 200 Alasdair C. Steven, G. C., Carmen Butan, Dennis C. Winkler and J. Bernard Heymann. in *Structural Virology* Vol. 21 *RSC Biomolecular Sciences* (ed Mavis Agbandje-McKenna and Robert McKenna) Ch. Ch. 4 Three-dimensional

- Structures of Pleiomorphic Viruses from Cryo-Electron Tomography, (Royal Society of Chemistry, 2011).
- 201 Liu, H. *et al.* Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. *Science* **329**, 1038-1043 (2010).
- 202 Henderson, R. & McMullan, G. Problems in obtaining perfect images by single-particle electron cryomicroscopy of biological structures in amorphous ice. *Microscopy* **62**, 43-50 (2013).
- 203 Tronrud, D. E. Introduction to macromolecular refinement. *Acta Crystallographica Section D: Biological Crystallography* **60**, 2156-2168 (2004).
- 204 Branden, C. & Tooze, J. *Introduction to protein structure*. Vol. 2 (Garland New York, 1991).
- 205 Warren, B. E. *X-ray Diffraction*. (Courier Dover Publications, 1969).
- 206 Perutz, M. Structure of hemoglobin. *Brookhaven Symp Biol* **13**, 165-183 (1960).
- 207 Kendrew, J. C. *et al.* A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662-666 (1958).
- 208 Karle, J. Some developments in anomalous dispersion for the structural investigation of macromolecular systems in biology. *International Journal of Quantum Chemistry* **18**, 357-367 (1980).
- 209 Argos, P. & Rossmann, M. G. in *Theory and practice of direct methods in crystallography* 361-417 (Springer, 1980).
- 210 Suryanarayana, C. & Norton, M. G. X-ray diffraction: a practical approach. *New York and London* (1998).
- 211 Wan, W., Sun, J., Su, J., Hovmoller, S. & Zou, X. Three-dimensional rotation electron diffraction: software RED for automated data collection and data processing. *Journal of applied crystallography* **46**, 1863-1873 (2013).
- 212 Bingel-Erlenmeyer, R. *et al.* SLS Crystallization Platform at Beamline X06DA□ A Fully Automated Pipeline Enabling in Situ X-ray Diffraction Screening. *Crystal Growth & Design* **11**, 916-923 (2011).
- 213 Winter, G. & McAuley, K. E. Automated data collection for macromolecular crystallography. *Methods* **55**, 81-93 (2011).
- 214 Lipkin, H. J. Physics of Debye-Waller Factors. *arXiv preprint cond-mat/0405023* (2004).
- 215 Wang, X. *et al.* A sensor-adaptor mechanism for enterovirus uncoating from structures of EV71. *Nature Structural & Molecular Biology* (2012).
- 216 Ruska, E. The early development of electron lenses and electron microscopy. *Microscopica acta. Supplement*, 1 (1980).
- 217 Gonen, T. *et al.* Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* **438**, 633-638 (2005).
- 218 Mitsuma, T. *et al.* Influence of the cytoplasmic domains of aquaporin-4 on water conduction and array formation. *Journal of molecular biology* **402**, 669-681 (2010).
- 219 Drummy, L. F., Yang, J. & Martin, D. C. Low-voltage electron microscopy of polymer and organic molecular thin films. *Ultramicroscopy* **99**, 247-256 (2004).
- 220 Roberts, J. A., Thorley, B. R., Bruggink, L. D. & Marshall, J. A. Electron microscope detection of an endogenous infection of retrovirus-like particles in L20B cells. *Microscopy* **62**, 485-486 (2013).
- 221 Locker, J. K. & Schmid, S. L. Integrated Electron Microscopy: Super-Duper Resolution. *PLoS biology* **11**, e1001639 (2013).

- 222 Taylor, K., Chanzy, H. & Marchessault, R. Electron diffraction for hydrated crystalline biopolymers: nigeran. *Journal of molecular biology* **92**, 165-167 (1975).
- 223 Dubochet, J., Lepault, J., Freeman, R., Berriman, J. & Homo, J. C. Electron microscopy of frozen water and aqueous solutions. *Journal of Microscopy* **128**, 219-237 (1982).
- 224 Adrian, M., Dubochet, J., Lepault, J. & McDowell, A. W. Cryo-electron microscopy of viruses. *Nature* **308**, 32-36 (1984).
- 225 Orlova, E. V. & Saibil, H. R. Structural analysis of macromolecular assemblies by electron microscopy. *Chemical reviews* **111**, 7710-7748 (2011).
- 226 Grigorieff, N. & Harrison, S. C. Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy. *Current opinion in structural biology* **21**, 265-273 (2011).
- 227 Koster, A. J. *et al.* Perspectives of molecular and cellular electron tomography. *Journal of structural biology* **120**, 276-308 (1997).
- 228 Milne, J. L. *et al.* Cryo-electron microscopy—a primer for the non-microscopist. *FEBS Journal* **280**, 28-45 (2013).
- 229 White, T. A. *et al.* Three-dimensional structures of soluble CD4-bound states of trimeric simian immunodeficiency virus envelope glycoproteins determined by using cryo-electron tomography. *Journal of virology* **85**, 12114-12123 (2011).
- 230 Zhang, X., Jin, L., Fang, Q., Hui, W. H. & Zhou, Z. H. 3.3 Å cryo-EM structure of a nonenveloped virus reveals a priming mechanism for cell entry. *Cell* **141**, 472-482 (2010).
- 231 Scheres, S. H. *et al.* Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nature Methods* **4**, 27-29 (2007).
- 232 Bartesaghi, A., Lecumberry, F., Sapiro, G. & Subramaniam, S. Protein secondary structure determination by constrained single-particle cryo-electron tomography. *Structure* **20**, 2003-2013 (2012).
- 233 Radermacher, M., Wagenknecht, T., Verschoor, A. & Frank, J. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of Escherichia coli. *Journal of microscopy* **146**, 113-136 (1987).
- 234 Seitsonen, J. J. *et al.* Structural analysis of coxsackievirus A7 reveals conformational changes associated with uncoating. *Journal of virology* **86**, 7207-7215 (2012).
- 235 Sanz-García, E., Stewart, A. B. & Belnap, D. M. The random-model method enables ab initio 3D reconstruction of asymmetric particles and determination of particle symmetry. *Journal of structural biology* **171**, 216-222 (2010).
- 236 Yu, H. Extending the size limit of protein nuclear magnetic resonance. *Proceedings of the National Academy of Sciences* **96**, 332-334 (1999).
- 237 Schein, C. H. *et al.* NMR structure of the viral peptide linked to the genome (VPg) of poliovirus. *Peptides* **27**, 1676-1684 (2006).
- 238 Schein, C. H. *et al.* NMR solution structure of poliovirus uridylyated peptide linked to the genome (VPgpU). *Peptides* **31**, 1441-1448 (2010).
- 239 Laflamme, R. *et al.* Introduction to NMR Quantum Information Processing. (2002).
- 240 Ernst, R. R., Bodenhausen, G. & Wokaun, A. *Principles of nuclear magnetic resonance in one and two dimensions*. Vol. 14 (Clarendon press Oxford, 1987).
- 241 Zhao, G. *et al.* Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **497**, 643-646 (2013).

- 242 Bochkov, Y. A. *et al.* Molecular modeling, organ culture and reverse genetics for a newly identified human rhinovirus C. *Nature medicine* **17**, 627-632 (2011).
- 243 Greer, J. Comparative modeling methods: application to the family of the mammalian serine proteases. *Proteins: Structure, Function, and Bioinformatics* **7**, 317-334 (1990).
- 244 Bernstein, F. C. *et al.* The protein data bank. *European Journal of Biochemistry* **80**, 319-324 (1977).
- 245 Dunbrack, R. L. in *Bioinformatics - From Genomes to Therapies*. Vol. 1 (ed Thomas Lengauer) Ch. 10, 297-350 (WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 2007).
- 246 Fernandez-Fuentes, N., Oliva, B. & Fiser, A. A supersecondary structure library and search algorithm for modeling loops in protein structures. *Nucleic acids research* **34**, 2085-2097 (2006).
- 247 Dunker, A. K. *et al.* Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling* **19**, 26-59 (2001).
- 248 Oberste, M. S., Maher, K., Kilpatrick, D. R. & Pallansch, M. A. Molecular evolution of the human enteroviruses: correlation of serotype with VP1 sequence and application to picornavirus classification. *Journal of virology* **73**, 1941-1948 (1999).
- 249 Weber, I. T. Evaluation of homology modeling of HIV protease. *Proteins: Structure, Function, and Bioinformatics* **7**, 172-184 (1990).
- 250 Taylor, E. & Jaakkola, J. A transposition of the reverse transcriptase gene reveals unexpected structural homology to E. coli DNA polymerase I. *Genetica* **84**, 77-86 (1991).
- 251 Mumenthaler, C. *et al.* A 3D model for the measles virus receptor CD46 based on homology modeling, Monte Carlo simulations, and hemagglutinin binding studies. *Protein science* **6**, 588-597 (1997).
- 252 Efremov, R. G. *et al.* Molecular modeling of HIV-1 coreceptor CCR5 and exploring of conformational space of its extracellular domain in molecular dynamics simulation. *Journal of Biomolecular Structure and Dynamics* **16**, 77-90 (1998).
- 253 Kolaskar, A. & Kulkarni-Kale, U. Prediction of three-dimensional structure and mapping of conformational epitopes of envelope glycoprotein of Japanese encephalitis virus. *Virology* **261**, 31-42 (1999).
- 254 Yoneda, T., Yoneda, S., Takayama, N., Kitazawa, M. & Umeyama, H. A homology modeling method of an icosahedral viral capsid: inclusion of surrounding protein structures. *Journal of Molecular Graphics and Modelling* **17**, 114-119 (1999).
- 255 Ranganathan, S., Singh, S., Poh, C. & Chow, V. The hand, foot and mouth disease virus capsid: sequence analysis and prediction of antigenic sites from homology modelling. *Applied bioinformatics* **1**, 43-52 (2001).
- 256 Schlick, T. *Molecular Modeling and Simulation: An Interdisciplinary Guide.*, Vol. 21 (Springer, 2010).
- 257 MacKerell, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry B* **102**, 3586-3616 (1998).
- 258 Lindahl, E. R. in *Molecular Modeling of Proteins Methods in Molecular Biology* (ed Andreas Kukol) Ch. Molecular Dynamics Simulations, (Humana Press, 2008).
- 259 Fox, N. in *ECE 697S: Topics in Computational Biology* (University of Massachusetts, 2006).

- 260 Feenstra, K. A., Hess, B. & Berendsen, H. J. Improving efficiency of large timescale molecular dynamics simulations of hydrogen-rich systems. *J Comput Chem* **20**, 786-798 (1999).
- 261 Hünenberger, P. H. & McCammon, J. A. Ewald artifacts in computer simulations of ionic solvation and ion-ion interaction: a continuum electrostatics study. *The Journal of chemical physics* **110**, 1856-1872 (1999).
- 262 Hünenberger, P. H. & McCammon, J. A. Effect of artificial periodicity in simulations of biomolecules under Ewald boundary conditions: a continuum electrostatics study. *Biophysical chemistry* **78**, 69-88 (1999).
- 263 Essmann, U. *et al.* A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **103**, 8577-8593 (1995).
- 264 Norberto, d. S. O. & Ornstein, R. L. Molecular dynamics simulations of a protein-protein dimer: particle-mesh Ewald electrostatic model yields far superior results to standard cutoff model. *Journal of biomolecular structure & dynamics* **16**, 1205-1218 (1999).
- 265 Westheimer, F. & Mayer, J. E. The Theory of the Racemization of Optically Active Derivatives of Diphenyl. *Journal of Chemical Physics* **14**, 733-738 (1946).
- 266 Goldstine, H. H. & Goldstine, A. The electronic numerical integrator and computer (ENIAC). *Mathematical Tables and Other Aids to Computation*, 97-110 (1946).
- 267 Felker, J. H. Patent US2670445 A Regenerative transistor amplifier. (1954).
- 268 Chaplin, G. The transistor regenerative amplifier as a computer element. *Proceedings of the IEE-Part III: Radio and Communication Engineering* **101**, 298-307 (1954).
- 269 Levitt, M. The birth of computational structural biology. *Nature Structural & Molecular Biology* **8**, 392-393 (2001).
- 270 Levitt, M. & Lifson, S. Refinement of protein conformations using a macromolecular energy minimization procedure. *Journal of molecular biology* **46**, 269-279 (1969).
- 271 Rahman, A. & Stillinger, F. H. Molecular dynamics study of liquid water. *The Journal of Chemical Physics* **55**, 3336-3359 (1971).
- 272 Stillinger, F. H. & Rahman, A. Improved simulation of liquid water by molecular dynamics. *The Journal of Chemical Physics* **60**, 1545-1557 (1974).
- 273 Brooks, B. R. *et al.* CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of computational chemistry* **4**, 187-217 (1983).
- 274 McCammon, J. A., Gelin BR , Karplus M. Dynamics of folded proteins. *Nature* **267**, 16 (1977).
- 275 Shrivastava, I. H. & Sansom, M. S. Simulations of ion permeation through a potassium channel: molecular dynamics of KcsA in a phospholipid bilayer. *Biophysical Journal* **78**, 557-570 (2000).
- 276 Seibel, G., Singh, U. & Kollman, P. A molecular dynamics simulation of double-helical B-DNA including counterions and water. *Proceedings of the National Academy of Sciences* **82**, 6537-6540 (1985).
- 277 Dongarra, J. J. The LINPACK benchmark: An explanation. *Supercomputing*, 456-474 (1988).
- 278 Lybrand, T. P. & McCammon, J. A. Computer simulation study of the binding of an antiviral agent to a sensitive and a resistant human rhinovirus. *Journal of computer-aided molecular design* **2**, 259-266 (1989).

- 279 Venable, R. M., Pastor, R. W., Brooks, B. R. & Carson, F. W. Theoretically
determined three-dimensional structures for amphipathic segments of the HIV-1
gp41 envelope protein. *AIDS research and human retroviruses* **5**, 7-22 (1989).
- 280 Weis, W. I., Brünger, A. T., Skehel, J. J. & Wiley, D. C. Refinement of the
influenza virus hemagglutinin by simulated annealing. *Journal of molecular
biology* **212**, 737-761 (1990).
- 281 von Itzstein, M. *et al.* Rational design of potent sialidase-based inhibitors of
influenza virus replication. *Nature* **363**, 418-423 (1993).
- 282 Lew, W., Chen, X. & Kim, C. U. Discovery and development of GS 4104
(oseltamivir) an orally active influenza neuraminidase inhibitor. *Current
medicinal chemistry* **7**, 663-672 (2000).
- 283 Phelps, D. K. & Post, C. B. Molecular dynamics investigation of the effect of an
antiviral compound on human rhinovirus. *Protein Science* **8**, 2281-2289 (1999).
- 284 Speelman, B., Brooks, B. R. & Post, C. B. Molecular dynamics simulations of
human rhinovirus and an antiviral compound. *Biophysical journal* **80**, 121-129
(2001).
- 285 Njoroge, F. G., Chen, K. X., Shih, N.-Y. & Piwinski, J. J. Challenges in modern
drug discovery: a case study of boceprevir, an HCV protease inhibitor for the
treatment of hepatitis C virus infection. *Accounts of chemical research* **41**, 50-59
(2008).
- 286 Xu, J. *et al.* Application of the Mole-8.5 supercomputer: Probing the whole
influenza virion at the atomic level. *Chinese Science Bulletin* **56**, 2114-2118
(2011).
- 287 Arkhipov, A., Freddolino, P. L. & Schulten, K. Stability and dynamics of virus
capsids described by coarse-grained modeling. *Structure* **14**, 1767-1777 (2006).
- 288 Jeong, H. *et al.* Performance of Kepler GTX Titan GPUs and Xeon Phi System.
arXiv preprint arXiv:1311.0590 (2013).
- 289 Carrico, J., Sabat, A., Friedrich, A., Ramirez, M. & Markers, E. S. G. f. E.
Bioinformatics in bacterial molecular epidemiology and public health: databases,
tools and the next-generation sequencing revolution. *Euro surveillance: bulletin
Européen sur les maladies transmissibles= European communicable disease
bulletin* **18**, 20382 (2013).
- 290 Knetsch, C. *et al.* Laboratory-based surveillance in the molecular era: the
TYPENED model, a joint data-sharing platform for clinical and public health
laboratories. (2013).
- 291 Roberts, J. A., Hung, A. & Thorley, B. R. Letter to the editor: Application of
bayesian methods to the inference of phylogeny for enterovirus surveillance.
Eurosurveillance **18** (2013).
- 292 Koopmans, M. *et al.* Authors' reply: Application of Bayesian methods to the
inference of phylogeny for enterovirus surveillance. *Euro surveillance: bulletin
européen sur les maladies transmissibles= European communicable disease
bulletin* **18** (2013).
- 293 Organization, W. H. Guidelines for environmental surveillance of poliovirus
circulation. (2003).
- 294 Manor, Y. *et al.* Detection of poliovirus circulation by environmental surveillance
in the absence of clinical cases in Israel and the Palestinian authority. *Journal of
clinical microbiology* **37**, 1670 (1999).
- 295 Carnie, J. A. *et al.* Public health response to imported case of poliomyelitis,
Australia, 2007. *Emerging Infectious Diseases* **15**, 1733-1737 (2009).

- 296 Stöcher, M., Leb, V. & Berg, J. A convenient approach to the generation of multiple internal control DNA for a panel of real-time PCR assays. *Journal of virological methods* **108**, 1-8 (2003).
- 297 Schuller, M., Carter, I. W., James, G. S., Halliday, C. L. & Sloots, T. P. *PCR for clinical microbiology: An Australian and International Perspective*. (Springer, 2010).
- 298 National Centre for Biotechnology Information., <www.ncbi.nlm.gov> (Accessed 12th September 2013.).
- 299 Honkanen, H. *et al.* Human rhinoviruses including group C are common in stool samples of young Finnish children. *Journal of Clinical Virology* **56**, 334-338 (2013).
- 300 Kesson, A. M., Choo, C. M., Troedson, C., Thorley, B. R. & Roberts, J. A. Echovirus 19 associated with a case of acute flaccid paralysis. *Journal of Paediatrics and Child Health* **49**, E239-E242 (2013).
- 301 National Centre for Biotechnology Information, Basic Local Alignment Search Tool . <<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>> (Accessed 12th September 2013.).
- 302 Knowles, N. J. *The Picornavirus Pages.*, <<http://www.picornaviridae.com/>> (Accessed 12th August 2013.).
- 303 SPREAD *Phylogeography.*, <<http://www.kuleuven.ac.be/aidslab/phylogeography/SPREAD.html>> (Accessed 29th October 2013).
- 304 Ray, S. C. SIMPLOT, SCSoftware., <<http://sray.med.som.jhmi.edu/SCSoftware/simplot/>> (Accessed 28th September 2013).
- 305 Lole, K. S. *et al.* Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *Journal of virology* **73**, 152-160 (1999).
- 306 WHO. World Health Organization, Polio laboratory manual. *World Health Organization* (2004).
- 307 Stambos, V., Brussen, K. A. & Thorley, B. R. Annual report of the Australian National Poliovirus Reference Laboratory, 2004. *Communicable diseases intelligence quarterly report* **29**, 263 (2005).
- 308 Brussen, K. A., Roberts, J., Ibrahim, A., Stambos, V. & Thorley, B. R. Annual report of the Australian National Poliovirus Reference Laboratory 2005. *Communicable diseases intelligence* **30**, 334-340 (2006).
- 309 Roberts, J., Brussen, K. A., Ibrahim, A. & Thorley, B. Annual report of the Australian National Poliovirus Reference Laboratory, 2006. *Communicable diseases intelligence* **31**, 263-269 (2007).
- 310 Roberts, J. A., Grant, K. A., Ibrahim, A. & Thorley, B. R. Annual report of the Australian National Poliovirus Reference Laboratory, 2007. *Communicable diseases intelligence* **32**, 308-315 (2008).
- 311 Roberts, J. A. *et al.* Annual report of the Australian National Poliovirus Reference Laboratory, 2008. *Communicable diseases intelligence* **33**, 291-297 (2009).
- 312 Roberts, J. A., Hobday, L., Polychronopoulos, S., Ibrahim, A. & Thorley, B. R. Annual report of the Australian National Poliovirus Reference Laboratory, 2009. *Communicable diseases intelligence* **34**, 277-284 (2010).
- 313 Roberts, J., Hobday, L., Ibrahim, A., Aitken, T. & Thorley, B. Annual report of the Australian National Enterovirus Reference Laboratory 2010-2011. *Communicable diseases intelligence quarterly report* **37**, E105-E114 (2012).

- 314 Roberts, J., Hobday, L., Ibrahim, A., Aitken, T. & Thorley, B. Annual report of
the Australian National Enterovirus Reference Laboratory 2012. *Communicable
diseases intelligence quarterly report* **37**, E97-E104 (2012).
- 315 Roberts, J. A. *Supercomputer Modelling of a Complete Human Viral Pathogen*,
<[http://www.youtube.com/watch?feature=player_embedded&v=WBDKmdS734
E](http://www.youtube.com/watch?feature=player_embedded&v=WBDKmdS734E)> (2013).
- 316 Sanders, S. *et al.* Molecular epidemiology of enterovirus 71 over two decades in
an Australian urban community. *Archives of virology* **151**, 1003-1013 (2006).
- 317 Roberts, J. T., B. Human Enterovirus 71 - Australia: Sub-Genogroup C4a, Acute
Flaccid Paralysis. *ProMED-mail* (2013).
- 318 McAnulty J. White, L. Enterovirus Neurological Disease. *NSW Health, General
Practitioner Alert* (2013).
- 319 McMinn, P. *et al.* Phylogenetic analysis of enterovirus 71 strains isolated during
linked epidemics in Malaysia, Singapore, and Western Australia. *Journal of
virology* **75**, 7732-7738 (2001).
- 320 Xing, W. *et al.* Hand, foot, and mouth disease in China, 2008–12: an
epidemiological study. *The Lancet Infectious Diseases* (2014).
- 321 Lu, J. *et al.* Whole Genomic Sequence and Replication Kinetics of a New
Enterovirus C96 Isolated from Guangdong, China with a Different Cell Tropism.
PLOS ONE **9**, e86877 (2014).
- 322 Knowles, N. J. *Picornavirus Study Group Enterovirus Species A*,
<http://www.picornastudygroup.com/types/enterovirus/ev_a_types.htm
>
(Accessed 12 September 2013).
- 323 Ng, T. F. F. *et al.* High variety of known and new RNA and DNA viruses of
diverse origins in untreated sewage. *Journal of virology* **86**, 12161-12175 (2012).
- 324 Al Rwahnih, M., Daubert, S., Golino, D. & Rowhani, A. Deep sequencing
analysis of RNAs from a grapevine showing Syrah decline symptoms reveals a
multiple virus infection that includes a novel virus. *Virology* **387**, 395-401 (2009).
- 325 Grard, G. *et al.* A novel rhabdovirus associated with acute hemorrhagic fever in
central Africa. *PLoS pathogens* **8**, e1002924 (2012).
- 326 Eswar, N. *et al.* Comparative protein structure modeling using Modeller. *Current
protocols in bioinformatics*, 5.6. 1-5.6. 30 (2006).
- 327 Guex, N., Peitsch, M. C. & Schwede, T. Automated comparative protein structure
modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective.
Electrophoresis **30**, S162-S173 (2009).
- 328 Wallner, B. & Elofsson, A. Can correct protein models be identified? *Protein
science* **12**, 1073-1086 (2003).
- 329 Cristobal, S., Zemla, A., Fischer, D., Rychlewski, L. & Elofsson, A. A study of
quality measures for protein threading models. *BMC bioinformatics* **2**, 5 (2001).
- 330 Fiser, A., Do, R. K. G. & Šali, A. Modeling of loops in protein structures. *Protein
science* **9**, 1753-1773 (2000).
- 331 Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace:
a web-based environment for protein structure homology modelling.
Bioinformatics **22**, 195-201 (2006).
- 332 Fricks, C. E. & Hogle, J. M. Cell-induced conformational change in poliovirus:
externalization of the amino terminus of VP1 is responsible for liposome binding.
Journal of Virology **64**, 1934-1945 (1990).
- 333 Baltimore, D. Picornaviruses are no longer black boxes. *Science* **229**, 1366-1367
(1985).

- 334 Hadfield, A. T. *et al.* The refined structure of human rhinovirus 16 at 2.15 Å
resolution: implications for the viral life cycle. *Structure* **5**, 427-441 (1997).
- 335 Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics.
Journal of molecular graphics **14**, 33-38 (1996).
- 336 Bostina, M., Levy, H., Filman, D. J. & Hogle, J. M. Poliovirus RNA is released
from the capsid near a twofold symmetry axis. *Journal of virology* **85**, 776 (2011).
- 337 Miller, S. T., Hogle, J. M. & Filman, D. J. Ab initio phasing of high-symmetry
macromolecular complexes: successful phasing of authentic poliovirus data to 3.0
Å resolution I. *Journal of Molecular Biology* **307**, 499-512 (2001).
- 338 McGuffin, L. J., Bryson, K. & Jones, D. T. The PSIPRED protein structure
prediction server. *Bioinformatics* **16**, 404-405 (2000).
- 339 Jones, D. T. Protein secondary structure prediction based on position-specific
scoring matrices. *Journal of molecular biology* **292**, 195-202 (1999).
- 340 VMD Scripts Library, University of Illinois,
<http://www.ks.uiuc.edu/Research/vmd/script_library/scripts> (Accessed 21
September 2013).
- 341 SWISSparam web service, Swiss Institute of Bioinformatics.,
<<http://swissparam.ch/>> (Accessed 12 September 2013).
- 342 Zoete, V., Cuendet, M. A., Grosdidier, A. & Michielin, O. SwissParam: a fast
force field generation tool for small organic molecules. *J Comput Chem* **32**, 2359-
2368, doi:10.1002/jcc.21816 (2011).
- 343 Make NA server, Chen Lab, University of Southern California.,
<<http://structure.usc.edu/make-na/index.html>> (Accessed 23rd September 2013).
- 344 Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *Journal of
Computational Chemistry* **26**, 1781 (2005).
- 345 Kavragi, LE. Molecular Distance Measures.,
<<http://cnx.org/content/m11608/latest/#>> (Accessed 13th November 2013.).
- 346 Torvalds, L. Linux: a portable operating system. *Master's thesis, University of
Helsinki, dept. of Computing Science* (1997).
- 347 Frishman, D. & Argos, P. Knowledge-based protein secondary structure
assignment. *Proteins: Structure, Function, and Bioinformatics* **23**, 566-579
(1995).
- 348 Smith, T. J. *et al.* The site of attachment in human rhinovirus 14 for antiviral
agents that inhibit uncoating. *Science* **233**, 1286-1293 (1986).
- 349 Tsang, S. K., Danthi, P., Chow, M. & Hogle, J. M. Stabilization of poliovirus by
capsid-binding antiviral drugs is due to entropic effects. *Journal of molecular
biology* **296**, 335-340 (2000).
- 350 Grant, R. A. *et al.* Structures of poliovirus complexes with anti-viral drugs:
implications for viral stability and drug design. *Current Biology* **4**, 784-797
(1994).
- 351 Hiremath, C. N., Grant, R. A., Filman, D. J. & Hogle, J. Binding of the antiviral
drug WIN51711 to the Sabin strain of type 3 poliovirus: structural comparison
with drug binding in rhinovirus 14. *Acta Crystallographica Section D: Biological
Crystallography* **51**, 473-489 (1995).
- 352 Lin, J. *et al.* Structure of the Fab-Labeled 'Breathing' State of Native Poliovirus.
Journal of virology (2012).
- 353 Wang, W., Lee, W.-M., Mosser, A. G. & Rueckert, R. R. WIN 52035-dependent
human rhinovirus 16: assembly deficiency caused by mutations near the canyon
surface. *Journal of virology* **72**, 1210-1218 (1998).

- 354 Allen, T. W., Andersen, O. S. & Roux, B. Molecular dynamics—potential of mean force calculations as a tool for understanding ion permeation and selectivity in narrow channels. *Biophysical chemistry* **124**, 251-267 (2006).
- 355 Patargias, G., Martay, H. & Fischer, W. B. Reconstructing potentials of mean force from short steered molecular dynamics simulations of Vpu from HIV-1. *Journal of Biomolecular Structure and Dynamics* **27**, 1-12 (2009).
- 356 Belnap, D. M. *et al.* Molecular tectonic model of virus structural transitions: the putative cell entry states of poliovirus. *Journal of Virology* **74**, 1342-1354 (2000).
- 357 El Omari, K. *et al.* Plate Tectonics of Virus Shell Assembly and Reorganization in Phage Φ 8, a Distant Relative of Mammalian Reoviruses. *Structure* **21**, 1384-1395 (2013).
- 358 Forrest, L. R., Tang, C. L. & Honig, B. On the accuracy of homology modeling and sequence alignment methods applied to membrane proteins. *Biophysical journal* **91**, 508-517 (2006).
- 359 Chothia, C. & Lesk, A. M. The relation between the divergence of sequence and structure in proteins. *The EMBO journal* **5**, 823 (1986).
- 360 Chung, S. Y. & Subbiah, S. in *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing.* 126-141.
- 361 MODELLER, Ben Webb, University of California San Francisco, <<http://salilab.org/modeller/>> (Accessed 12th September 2013).
- 362 DeepView-Swiss-PdbViewer, Swiss Institute of Bioinformatics., <<http://spdbv.vital-it.ch/>> (Accessed 28 September 2013).
- 363 Brown, B., Oberste, M. S., Maher, K. & Pallansch, M. A. Complete genomic sequencing shows that polioviruses and members of human enterovirus species C are closely related in the noncapsid coding region. *Journal of virology* **77**, 8973-8984 (2003).
- 364 ProQ-Protein quality prediction, Stockholm Bioinformatics Center., <<http://www.sbc.su.se/~bjornw/ProQ/ProQ.cgi>> (Accessed 20th February 2014).
- 365 Xiao, C. *et al.* The crystal structure of coxsackievirus A21 and its interaction with ICAM-1. *Structure* **13**, 1019-1033 (2005).
- 366 Mao, L.-X. *et al.* Epidemiology of hand, foot, and mouth disease and genotype characterization of Enterovirus 71 in Jiangsu, China. *Journal of Clinical Virology* **49**, 100-104 (2010).
- 367 Plevka, P., Perera, R., Cardoso, J., Kuhn, R. J. & Rossmann, M. G. Structure determination of enterovirus 71. *Acta Crystallographica Section D: Biological Crystallography* **68**, 1217-1222 (2012).
- 368 Chen, P. *et al.* Molecular Determinants of Enterovirus 71 Viral Entry Cleft Around Gln-172 On Vp1 Protein Interacts With Variable Region On Scavenge Receptor B2. *Journal of Biological Chemistry* **287**, 6406-6420 (2012).
- 369 Yamayoshi, S., Ohka, S., Fujii, K. & Koike, S. Functional comparison of SCARB2 and PSGL1 as receptors for enterovirus 71. *Journal of virology* **87**, 3335-3347 (2013).
- 370 Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series* **41**, 95-98 (1999).
- 371 Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783-791 (1985).
- 372 Sokhey, J., Gupta, C. K., Sharma, B. & Singh, H. Stability of oral polio vaccine at different temperatures. *Vaccine* **6**, 12-13 (1988).

- 373 Liu, H.-M. *et al.* Characterization of poliovirus variants selected for resistance to the antiviral compound v-073. *Antimicrobial agents and chemotherapy* **56**, 5568-5574 (2012).
- 374 Trabuco, L. G., Villa, E., Mitra, K., Frank, J. & Schulten, K. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* **16**, 673-683 (2008).
- 375 Chan, K.-Y. *et al.* Symmetry-restrained flexible fitting for symmetric EM maps. *Structure* **19**, 1211-1218 (2011).
- 376 Vergara-Jaque, A. *et al.* Molecular basis of drug resistance in A/H1N1 virus. *Journal of chemical information and modeling* **52**, 2650-2656 (2012).
- 377 Lentz, K. N. *et al.* Structure of poliovirus type 2 Lansing complexed with antiviral agent SCH48973: comparison of the structural and biological properties of the three poliovirus serotypes. *Structure* **5**, 961-978 (1997).

Appendices

A1. Material and Reagents listing

Item	Company	Cat #
0.22 µm filter units	Sartorius	17597K
50mL Falcon Tubes	Becton Dickinson	352070
Big Dye Terminator with 5 x buffer	University of Melbourne, Australia	BDT v3.1
Bioanalyzer DNA 7500 series II kit	Pacific Laboratory Products	5067-1506
Chloroform	Sigma Aldrich Fine Chemicals	496189
dNTP Mix 40mM (10mM each)	Promega	U1511
EMEM, w/ Hanks Salts and L-Glutamine,	Sigma Aldrich Fine Chemicals	M4780-500mL
Ethanol	Sigma Aldrich Fine Chemicals	E-7023-500ML
FastStart PCR Master mix	Roche	04710444001
HEPES Buffer	Sigma Aldrich Fine Chemicals	H0887-100mL
High pure PCR product purification kit	Roche	11 732 668 001
MagMax -96 Viral RNA isolation kit	Applied Biosystems	AM1836
MagMax 96 well plates	Applied Biosystems	4388474
MagMax combs	Applied Biosystems	4388452
MicroAmp 96-well optical FAST plates	Applied Biosystems	4346906
MicroAmp Optical 8-well Strip caps	Applied Biosystems	4323032
MicroAmp Optical 8-well Strip tubes	Applied Biosystems	4358293
Nuclease free water	Promega	P1193
Nunc plates (96 flat well microtitre)	Thermo Fisher Scientific	167008
Nunc TC flasks 80cm ²	Thermo Fisher Scientific	153732
Oligonucleotide probes	Biosearch	NA
Oligonucleotide probes, primers and calibrator dye	Geneworks	NA
Optical adhesive film	Applied Biosystems	4311971
Penicillin-Streptomycin	Sigma Aldrich Fine Chemicals	P0781-100mL
Protector RNase inhibitor	Roche	03 335 402 001
Sarstedt plastic Tissue Culture tubes	Sarstedt	83.9923.953
Superscript III Reverse Transcriptase	Life Technologies	18080-044
Superscript III Reverse Transcriptase and Platinum Taq	Life Technologies	12574-026
Taq DNA Polymerase	Roche	11418432001
TaqMan Fast Universal PCR master mix	Applied Biosystems	4352042
Transcriptor	Roche	03 531 287 001
Trypan Blue 0.4% (250mL)	Fluka	93595-250ML
Trypsin-EDTA (1x solution)	Sigma Aldrich Fine Chemicals	59430C-500ML

A2 Equipment.

Item	Manufacturer	Serial Number	Asset Number
7500FAST Real-time machine	Applied Biosystems	275011364	MH11617
Veriti thermalcycler	Applied Biosystems	299020589	MH14731
Bioanalyser	Agilent Technologies	DE54704689	MH11618
Precision T3500, Tesla C2070, Quadro 5000	Dell	4BQT22S	MH15107
Freezer, -30°C	Sanyo	60814358	MH12067
Freezer -70°C	Thermo Scientific	814876-2110	MH14656
Blue Gene/P supercomputer,	IBM	Grant number VR0069	Grant number VR0069
Blue Gene/Q supercomputer	IBM	Gant numbers VR0069, VR0262	Grant numbers VR0069, VR0262
Incubator with roller drums	LEEC	P33	MH12070
MagMax extraction robot	Applied Biosystems	700-450 rs-232c	MH18095
Microscope	Olympus	CK2	MH45028
Shaker	Ratek	80597342	MH13265
Refrigerator, 4°C	Westinghouse	63252923	MH11413
Centrifuge, refrigerated	Beckman Coulter	ALP07H02	MH11600
Altix XE supercomputer	SGI	Grant numbers VR0069, VR0262	Grant numbers VR0069, VR0262
Nitro G25, 4x Tesla M2090, compute node	Xenon	N/A	MH23069

A3 Specific Software Package Details

- ABI 7500 software version 2.0.5
- Agilent Bioanalyser 2100 expert version B.02.04.SI356
- Bioedit version 7.1.3.0
- BEAST 1.7.4
- BEAGLE 1.0
- Blender version 2.63
- ChromasPro version 1.5
- Chromas 2.24
- FigTree version 1.4
- GIMP
- Libre Office 3
- Linux Mint 14
- MakeNA
- MEGA 5.1
- Microsoft Office 2003, 2010
- Modeller version 9.10
- Nanoscale Molecular Dynamics (NAMD) version 2.7, 2.8, 2.9
- NCBI (GenBank portal)
- PSIPred
- RCSB (PDB portal)
- SwissPDBviewer (aka DeepView) version 4.04
- SWISSmodel
- TreeView
- Ubuntu 10.04 LTS
- Visual Molecular Dynamics (VMD) 1.9.2

A4 Sequencing Primers

	Primer	Sequence5' to 3'	Position (PV-1 Sabin V01150)
1	PanEV1_OS	CAAGCACTTCTGTTTCCCCGG	164-184
2	PanEV2_IS	TCCTCCGGCCCTGAATGCG	446-465
3	PanEV3_laS	ATTGTCACCATAAGCAGCCA	581-600

Table A4.1. Generic. Primers¹ used in sequencing of both EV-C96 and EV-A120. Primer locations based on the prototype sequence for Poliovirus type 1, GenBank accession number V01150.

	Primer	Sequence5' to 3'	Position (EV-C96 EF015886)
1	12134022_VP1_for	TACACTGATACTGTCCAACCTCAG	2841-2863
2	12134022_VP1_rev	CTGAGTTGGACAGTATCAGTGTA	2863-2841
3	12173491_VP1_for	TATACAGACACAGTACAACCTCAG	2841-2863
4	12173491_VP1_rev	CTGAGTTCTACTGTGTCTGTATA	2863-2841
5	04107711_VP1_for	TACACAGACACAGTTCAACCTCAG	2841-2863
6	04107711_VP1_rev	CTGAGTTGAACTGTGTCTGTGTA	2863-2841
7	05118030_VP1_for	TACACAGACACAGTGCAACTTAG	2841-2863
8	05118030_VP1_rev	CTAAGTTGCACTGTGTCTGTGTA	2863-2841
9	EV-C_2C_rev	TCTTGAAGTGTATGTARTTGTTG	4478-4492
10	EV96_2A_3657_rev	TCRTTRGSYTCATSAYATG	3657-3676
11	EV96_VP1_3288_for	ATGAARCCYAARCAATYAGAG	3288-3309
12	EV96_VP1_2877_rev	AARCGYGARTAKGTRAASAT	2877-2896
13	EV96_VP1_2508_for	GARWCMGTRGCYCADAAATGC	2508-2527
14	EV96_VP3_2037_rev	TGYARYCTYTGRTCACTWGC	2037-2056
15	EV96_VP3_1950_rev	TACATRTCCTAKGTRTHGC	1950-1969
16	EV96_VP3_1818_for	AAITTYCARTCRCCRTGTGC	1818-1837
17	EV96_VP2_1370_rev	ACATARCTRGTGTARTTKGTC	1370-1390
18	EV96_VP2_1090_for	ARGCHAACCCMGTTGATGC	1090-1108
19	EV96_VP4_846_rev	GCATTRCTRGCRCRTATCYTT	846-866

Table A4.2. EV-C96. Primers¹ used in sequencing of EV-C96. Primer locations based on the prototype sequence Enterovirus C96 BAN-00, GenBank accession number EF015886.

	Primer	Sequence5' to 3'	Position (CV-A16 U05876)
1	Novel_EVA_VP1rev	TCCCCATTTGGAGTAGTAGC	2863-2882
2	Novel_EVA_VP1for	GGGCTTGTGGGAGTTATAGA	2707-2726
3	Novel_EVA_2Crev	GTGTTTGCTCTTGAAGTGCATGTA	4414-4437
4	EV-A120_4235for	AAATTATACCAGCAGCCAAAGA	4235-4256
5	EV-A120_3540rev	GTGCTTCCTTTTAGAATTGCAG	3519-3540
6	EV-A120_3450for	AGTTTCCTCTACCACTGCTCAA	3450-3471
7	EV-A120_2040rev	GTAATAACGGCACAATTGACCT	2019-2040
8	EV-A120_1930for	ATACCGGTTAAAGTGCAAACAT	1930-1951
9	EV-A120_1360rev	ATACAGTGCCAGCACATATTC	1339-1360
10	EV-A120_1265for	ATTTTGCATCCACGTTTCACT	1265-1284

Table A4.3. EV-A120. Primers¹ used in sequencing of the novel virus EV-A120. Primer locations based on the prototype sequence for Coxsackievirus A16 GenBank accession number U05876.

¹Abbreviations Used:

For=Sense, *Rev*=Anti-sense, *A*=Adenine, *C*=Cytosine, *G*=Guanine, *T*=Thymine, *I*=Inosine, *B*=C,G,T, *D*=A,G,T, *H*=A,C,T, *K*=G,T, *M*=A,C, *N*=A,C,G,T, *R*=A,G, *S*=C,G, *V*=A,C,G, *W*=A,T, *Y*=C,T.

A5 Computational Biology Scripts

A5.1 TCL scripts written for atom selection for application of constraints and virus substructure positioning

A5.1.1 File Name: Beta_select.tcl

```
#select and changes beta values for alpha carbon for constraints
set sel0 [atomselect 0 "protein and type C"]
$sel0 set beta 1
```

A5.1.2 File Name: RNA_supercoil_betaselect_step1

```
# selects and changes beta values for Mahoney RNA coordinates
# 10 strands moves the centre 10 residues 1000A y axis
# step 1 of 2 step in triangulation of RNA strands
# in preparation for joining and supercoiling

set N1 [atomselect top "segname N1 and resid 370 to 380"]

    $N1 set beta 1
    $N1 set x 0
    $N1 set y 1000
    $N1 set z 1260

set N2 [atomselect top "segname N2 and resid 370 to 380"]

    $N2 set beta 1
    $N2 set x 40
    $N2 set y 1000
    $N2 set z 1260

set N3 [atomselect top "segname N3 and resid 370 to 380"]

    $N3 set beta 1
    $N3 set x 80
    $N3 set y 1000
    $N3 set z 1260

set N4 [atomselect top "segname N4 and resid 370 to 380"]

    $N4 set beta 1
    $N4 set x 120
    $N4 set y 1000
    $N4 set z 1260

set N5 [atomselect top "segname N5 and resid 370 to 380"]

    $N5 set beta 1
    $N5 set x 160
```

```

    $N5 set y 1000
    $N5 set z 1260

set N6 [atomselect top "segname N6 and resid 370 to 380"]

    $N6 set beta 1
    $N6 set x 200
    $N6 set y 1000
    $N6 set z 1260

set N7 [atomselect top "segname N7 and resid 370 to 380"]

    $N7 set beta 1
    $N7 set x 240
    $N7 set y 1000
    $N7 set z 1260

set N8 [atomselect top "segname N8 and resid 370 to 380"]

    $N8 set beta 1
    $N8 set x 280
    $N8 set y 1000
    $N8 set z 1260

set N9 [atomselect top "segname N9 and resid 370 to 380"]

    $N9 set beta 1
    $N9 set x 320
    $N9 set y 1000
    $N9 set z 1260

set N10 [atomselect top "segname N10 and resid 370 to 380"]

    $N10 set beta 1
    $N10 set x 360
    $N10 set y 1000
    $N10 set z 1260

```

A5.1.3 File Name: Origin_constraints.tcl

```

#select and changes beta values for alpha carbon for constraints

set sel0 [atomselect top all]

$sel0 set beta 1
$sel0 set x 1
$sel0 set y 1
$sel0 set z 1

```


A5.2 TCL scripts written for the calculation of virus diameter, RMSD and RMSE.

A5.2.1 File Name: diameterCA

```
set outfile [open viriondiameterCA.dat w]
set nf [molinfo top get numframes]
set sel [atomselect top "protein and type CA"]
for { set i 1 } { $i < $nf } { incr i } {
  $sel frame $i
  puts $outfile "[measure minmax $sel]"
}
close $outfile
```

A5.2.2 File Name: RMSD_segments

```
# Script to calculate RMSD's for last frame of trajectory
# Calculates each RMSD for a given number of segments

# Create output file
set output [open "RMSD_segments_PV1W.dat" w]

#calculate VP1 RMSD's

for {set VP1 1} {$VP1 <= 240} {incr VP1} {
  puts $VP1

  set frame_ref [atomselect top " segname P$VP1 and backbone " frame 0]
  set frame_calc [atomselect top " segname P$VP1 and backbone " frame last]

  set fit [measure fit $frame_calc $frame_ref]
  set frame_move [atomselect top " segname P$VP1 and backbone " frame last]
  $frame_move move $fit

  set RMSD [measure rmsd $frame_move $frame_ref]
  puts $output $VP1,$RMSD

}

close $output
```

A5.2.3 File Name: VirusCapsidRMSD

```
#####  
#  
# Calculates the RMSD for a given number of segments in a multimeric assembly #  
# May 8th 2011 by Jason A. Roberts, VIDRL/RMIT University Australia #  
#  
#####  
  
# Create prefix for output file  
# Enter the virus name after the "set virus" variable  
  
set virus PV1WBG  
  
# Calculate VP1 RMSD's  
  
# Create VP1 output file  
  
set VP1chains [open "$virus.VP1_RMSD.dat" w]  
  
# Start looping for x number of times with start point 1 "sega"  
for {set sega 1} {$sega <= 240} {incr sega} {  
  
# Print out a screen count for visual feedback  
puts $sega  
  
# Select the first and last frames for RMSD calculation  
set frame_ref [atomselect top " segname P$sega and backbone " frame first]  
set frame_calc [atomselect top " segname P$sega and backbone " frame last]  
  
# Fit last frame to reference frame and move molecule  
set fit [measure fit $frame_calc $frame_ref]  
set frame_move [atomselect top " segname P$sega and backbone " frame last]  
$frame_move move $fit  
  
# Calculate the RMSD  
set RMSD [measure rmsd $frame_move $frame_ref]  
puts $VP1chains $sega,$RMSD  
  
#Skip 3 molecules ahead according to previous chain IDs in usage  
set sega [expr 3 + $sega]  
  
}  
  
# Close the output file  
close $VP1chains  
  
# End of script for VP1, the remaining VPs are the same  
# but have a new start point "segb" starts at 2, "segc" starts at 3 etc.  
  
#calculate VP2 RMSD's  
  
set VP2chains [open "$virus.VP2_RMSD.dat" w]  
  
for {set segb 2} {$segb <= 240} {incr segb} {  
puts $segb  
  
set frame_ref [atomselect top " segname P$segb and backbone " frame first]  
set frame_calc [atomselect top " segname P$segb and backbone " frame last]  
  
set fit [measure fit $frame_calc $frame_ref]  
set frame_move [atomselect top " segname P$segb and backbone " frame last]  
$frame_move move $fit  
  
set RMSD [measure rmsd $frame_move $frame_ref]  
puts $VP2chains $segb,$RMSD  
  
set segb [expr 3 + $segb]
```

```

    }
close $VP2chains

#calculate VP3 RMSD's
set VP3chains [open "$virus.VP3_RMSD.dat" w]

for {set segc 3} {$segc <= 240} {incr segc} {
puts $segc

    set frame_ref [atomselect top " segname P$segc and backbone " frame first]
    set frame_calc [atomselect top " segname P$segc and backbone " frame last]

    set fit [measure fit $frame_calc $frame_ref]
    set frame_move [atomselect top " segname P$segc and backbone " frame last]
    $frame_move move $fit

    set RMSD [measure rmsd $frame_move $frame_ref]
    puts $VP3chains $segc,$RMSD

    set segc [expr 3 + $segc]
}

close $VP3chains

#calculate VP4 RMSD's
set VP4chains [open "$virus.VP4_RMSD.dat" w]

for {set segd 4} {$segd <= 240} {incr segd} {
puts $segd

    set frame_ref [atomselect top " segname P$segd and backbone " frame first]
    set frame_calc [atomselect top " segname P$segd and backbone " frame last]

    set fit [measure fit $frame_calc $frame_ref]
    set frame_move [atomselect top " segname P$segd and backbone " frame last]
    $frame_move move $fit

    set RMSD [measure rmsd $frame_move $frame_ref]
    puts $VP4chains $segd,$RMSD

    set segd [expr 3 + $segd]
}

close $VP4chains

```

A5.2.4 File Name: RMSF_Capsidv2

```
#####  
#  
#   RMSF script for cuboidal PV1W and PV1SL protomers   #  
#  
#####  
  
# create output file  
  
set output [open "RMSF_protomers.dat" w]  
  
#calculate RMSF  
  
for {set segA 1} {$segA <=240} {incr segA} {  
  
puts $segA  
  
    set segB [expr 1 + $segA]  
    set segC [expr 2 + $segA]  
    set segD [expr 3 + $segA]  
  
    set segment [atomselect top "segname P$segA P$segB P$segC P$segD and name CA"]  
    set result [measure rmsf $segment first 0 last 9 step 1]  
  
    puts $output "Protomer_Starting$segA"  
    puts $output $result  
  
    set segA [expr 3 +$segA]  
  
    }  
  
close $output
```

A5.3 TCL script written for the renaming of virus protein chains.

File Name: VirusCapsidChainRename.tcl

```
#####  
#  
# Renames the chains for a given number of segments in a multimeric assembly #  
# May 8th 2011 by Jason A. Roberts, VIDRL/RMIT University Australia #  
######  
  
# Start looping for x number of times with start point 1 "sega"  
for {set sega 1} {$sega <= 240} {incr sega} {  
  
# Print out a screen count for visual feedback  
puts $sega  
  
# Select the first and last frames for RMSD calculation  
set chain1 [atomelect top " segname P$sega"]  
$chain1 set chain 1  
  
#Skip 3 molecules ahead according to previous chain IDs in usage  
set sega [expr 3 + $sega]  
  
}  
  
# End of script for VP1, the remaining VPs are the same  
# but have a new start point "segb" starts at 2, "segc" starts at 3 etc.  
  
# Start looping for x number of times with start point 2 "segb"  
for {set segb 2} {$segb <= 240} {incr segb} {  
  
# Print out a screen count for visual feedback  
puts $segb  
  
# Select the first and last frames for RMSD calculation  
set chain2 [atomelect top " segname P$segb"]  
$chain2 set chain 2  
  
#Skip 3 molecules ahead according to previous chain IDs in usage  
set segb [expr 3 + $segb]  
  
}  
  
# End of script for VP2  
  
# Start looping for x number of times with start point 3 "segc"  
for {set segc 3} {$segc <= 240} {incr segc} {  
  
# Print out a screen count for visual feedback  
puts $segc  
  
# Select the first and last frames for RMSD calculation  
set chain3 [atomelect top " segname P$segc"]  
$chain3 set chain 3  
  
#Skip 3 molecules ahead according to previous chain IDs in usage  
set segc [expr 3 + $segc]  
  
}  
  
# End of script for VP3  
  
# Start looping for x number of times with start point 4 "segd"  
for {set segd 4} {$segd <= 240} {incr segd} {  
  
# Print out a screen count for visual feedback  
puts $segd  
  
# Select the first and last frames for RMSD calculation  
set chain4 [atomelect top " segname P$segd"]  
$chain4 set chain 4  
  
#Skip 3 molecules ahead according to previous chain IDs in usage  
set segd [expr 3 + $segd]  
  
}  
  
# End of script for VP4
```

A5.4 TCL script written postion all 60 virus protomers in the same space as protomer

No. 1.

File Name: ProtomerFit_capsidalternate.tcl

```
#####
#
#   Script for translating protomer positions relative
#   to protomer reference 1 coordinates for analysis
#   Jason A. Roberts 2011 VIDRL RMIT
#
#####

# Enter the number of protomers after "set protomer"
# Enter the name of the output file after "set name"
# The rest just resets all variables to 0

set protomer      60
set name          PV1W5ns
set counter       0
set loop          0
set segA          0
set segB          0
set segC          0
set segD          0
set segE          0
set segF          0
set reference [atomselect top "segname P1 P2 P3 P4 and resid 30 to 60"]

# This procedure lifted from John Stones "animatepdb"
proc animatepdb {start end fileformat} {
    set filename [format $fileformat [expr $start]]
    incr start
    puts "Reading initial frame in PDB sequence $filename"
    mol load pdb $filename

    puts "Reading PDB files as an animation..."
    for {set i $start} {$i <= $end} {incr i 1} {
        set filename [format $fileformat [expr $i]]
        animate read pdb $filename
    }
}

# This part loops the selections for each protein and lipid/drug
# Then remaps them to the new coordinates and updates the new pdb
for {set counter 1} {$counter <= $protomer} {incr counter} {

    set segA [expr 1 + $loop]
    set segB [expr 2 + $loop]
    set segC [expr 3 + $loop]
    set segD [expr 4 + $loop]

    set segF [expr $segD / 2]
    set segE [expr $segF - 1]

    set segment [atomselect top "segname P$segA P$segB P$segC P$segD O$segE O$segF"]
    set query [atomselect top "segname P$segA P$segB P$segC P$segD and resid"]
    set fit [measure fit $query $reference]
    $segment move $fit

    $segment writepdb "$counter.pdb"

    puts "protomer $counter"
    puts "loop number $loop"
    puts "P$segA"
    puts "P$segB"
    puts "P$segC"
    puts "P$segD"
    puts "O$segE"
    puts "O$segF"

    set loop [expr 1 + $segC]
}

animatepdb 1 $loop %d.pdb
```

A5.5 Selection parameters for the selection and comparison of multiple poliovirus capsid proteins with standard naming convention.

File Name: 1HXS_crystal_selection.txt

1HXS.pdb specific selection data

(chain 1 and resid 20 to 302) or (chain 2 and resid 6 to 272) or (chain 3 and resid 1 to 235)
or (chain 4 and resid 2 to 69)

Matches across the range of available type 1 crystal structures

(chain 1 and resid 70 to 302) or (chain 2 and resid 6 to 272) or (chain 3 and resid 1 to 235)
or (chain 4 and resid 2 to 13) or (chain 4 and resid 25 to 69)

A5.6. Small molecule parameterization, customised topology data file for virus construction in VMD.

File Name: Picornavirus_small_molecules_rev2.top

```

*>Small molecule topology for picornavirus simulations<
*>>>>>>> JA Roberts, MJ Kuiper and A Hung 2012 <<<<<<<<

!The following residue PLM (PALM) derived from original
CHARMM file.
!partial charges derived from (PALM) in original CHARMM
!file and renamed to suit virus reconstructions
!Jason Roberts 2012 VIDRL/RMIT University 2012

MASS 201 CO2M 12.011000
MASS 202 O2CM 15.999400
MASS 203 CR 12.011000
MASS 204 HCMM 1.007940

AUTOGENERATE ANGLES DIHE
DEFA FIRS NONE LAST NONE

RESI PLM -1.000
GROUP
ATOM C1 CO2M 0.62
ATOM O1 O2CM -0.76
ATOM O2 O2CM -0.76
ATOM C2 CR -0.28
ATOM C3 CR -0.18
ATOM C4 CR -0.18
ATOM C5 CR -0.18
ATOM C6 CR -0.18
ATOM C7 CR -0.18
ATOM C8 CR -0.18
ATOM C9 CR -0.18
ATOM CA CR -0.18
ATOM CB CR -0.18
ATOM CC CR -0.18
ATOM CD CR -0.18
ATOM CE CR -0.18
ATOM CF CR -0.18
ATOM CG CR -0.27
ATOM H22 HCMM 0.09
ATOM H23 HCMM 0.09
ATOM H32 HCMM 0.09
ATOM H33 HCMM 0.09
ATOM H42 HCMM 0.09
ATOM H43 HCMM 0.09
ATOM H52 HCMM 0.09
ATOM H53 HCMM 0.09
ATOM H62 HCMM 0.09
ATOM H63 HCMM 0.09
ATOM H72 HCMM 0.09
ATOM H73 HCMM 0.09
ATOM H82 HCMM 0.09
ATOM H83 HCMM 0.09
ATOM H92 HCMM 0.09
ATOM H93 HCMM 0.09
ATOM HA2 HCMM 0.09
ATOM HA3 HCMM 0.09
ATOM HB2 HCMM 0.09
ATOM HB3 HCMM 0.09
ATOM HC2 HCMM 0.09
ATOM HC3 HCMM 0.09
ATOM HD2 HCMM 0.09
ATOM HD3 HCMM 0.09
ATOM HE2 HCMM 0.09
ATOM HE3 HCMM 0.09
ATOM HF2 HCMM 0.09
ATOM HF3 HCMM 0.09
ATOM HG1 HCMM 0.09
ATOM HG2 HCMM 0.09
ATOM HG3 HCMM 0.09
BOND C1 O1
BOND C1 O2
BOND C1 C2
BOND C2 C3
BOND C3 C4
BOND C4 C5
BOND C5 C6
BOND C6 C7
BOND C7 C8
BOND C8 C9
BOND C9 CA
BOND CA CB
BOND CB CC
BOND CC CD
BOND CD CE
BOND CE CF
BOND CF CG
BOND CG HG2
BOND HG2 CG
BOND HG1 CG
BOND HG3 CG
BOND HF2 CF
BOND HF3 CF
BOND HE2 CE
BOND HE3 CE
BOND HD2 CD
BOND HD3 CD
BOND HC2 CC
BOND HC3 CC
BOND HB2 CB
BOND HB3 CB
BOND HA2 CA
BOND HA3 CA
BOND H93 C9
BOND H92 C9
BOND H83 C8
BOND H82 C8
BOND H73 C7
BOND H72 C7
BOND H63 C6
BOND H62 C6
BOND H53 C5
BOND H52 C5
BOND H43 C4
BOND H42 C4
BOND H33 C3
BOND H32 C3
BOND H23 C2
BOND H22 C2
IMPH C1 O1 C2 O2
IMPH C2 C3 C1 H23
IMPH C2 C3 C1 H22
IMPH C3 C4 C2 H33
IMPH C3 C4 C2 H32
IMPH C4 C5 C3 H43
IMPH C4 C5 C3 H42
IMPH C5 C6 C4 H53
IMPH C5 C6 C4 H52
IMPH C6 C7 C5 H63
IMPH C6 C7 C5 H62
IMPH C7 C8 C6 H73
IMPH C7 C8 C6 H72
IMPH C8 C9 C7 H83
IMPH C8 C9 C7 H82
IMPH C9 CA C8 H93
IMPH C9 CA C8 H92
IMPH CA CB C9 HA2
IMPH CA CB C9 HA3
IMPH CB CC CA HB3
IMPH CB CC CA HB2
IMPH CC CD CB HC2
IMPH CC CD CB HC3
IMPH CD CE CC HD3
IMPH CD CE CC HD2
IMPH CE CF CD HE3
IMPH CE CF CD HE2
IMPH CF CG CE HF2
IMPH CF CG CE HF3
IMPH CG HG2 CF HG1
IMPH CG HG2 CF HG3
IC C1 C2 C3 C4 1.55 119.16 -175.79 110.92
1.57
IC C1 C2 C3 H32 1.55 119.16 -55.61 109.08
1.09
IC C1 C2 C3 H33 1.55 119.16 63.78 109.25
1.09
IC O1 C1 C2 C3 1.26 120.56 -103.84 119.16
1.55
IC O1 C1 C2 H22 1.26 120.56 17.52 107.00
1.09
IC O1 C1 C2 H23 1.26 120.56 133.11 107.73
1.09
IC O2 C1 C2 C3 1.25 118.42 79.75 119.16
1.55
IC O2 C1 C2 H22 1.25 118.42 -158.89 107.00
1.09
IC O2 C1 C2 H23 1.25 118.42 -43.30 107.73
1.09
IC C2 C3 C4 C5 1.55 110.92 62.15 109.64
1.53
IC C2 C3 C4 H42 1.55 110.92 -57.86 109.41
1.09
IC C2 C3 C4 H43 1.55 110.92 -177.79 109.43
1.09
IC C3 C4 C5 C6 1.57 109.64 -176.79 114.94
1.54
IC C3 C4 C5 H52 1.57 109.64 62.41 108.09
1.09

```


IC C3	C4	C5	H53	1.57	109.64	-55.11	108.52	1.09
IC C4	C3	C2	H22	1.57	110.92	62.83	106.99	1.09
IC C4	C3	C2	H23	1.57	110.92	-52.76	107.77	1.09
IC C4	C5	C6	C7	1.53	114.94	-167.63	107.67	1.56
IC C4	C5	C6	H62	1.53	114.94	-47.91	109.90	1.09
IC C4	C5	C6	H63	1.53	114.94	72.91	109.78	1.09
IC C5	C4	C3	H32	1.53	109.64	-58.03	109.10	1.09
IC C5	C4	C3	H33	1.53	109.64	-177.40	109.21	1.09
IC C5	C6	C7	C8	1.54	107.67	-161.51	109.67	1.54
IC C5	C6	C7	H72	1.54	107.67	-41.48	109.41	1.09
IC C5	C6	C7	H73	1.54	107.67	78.40	109.47	1.09
IC C6	C5	C4	H42	1.54	114.94	-56.80	109.43	1.09
IC C6	C5	C4	H43	1.54	114.94	63.16	109.46	1.09
IC C6	C7	C8	C9	1.56	109.67	-161.40	113.63	1.53
IC C6	C7	C8	H82	1.56	109.67	77.99	108.41	1.09
IC C6	C7	C8	H83	1.56	109.67	-40.09	108.75	1.09
IC C7	C6	C5	H52	1.56	107.67	-46.83	108.08	1.09
IC C7	C6	C5	H53	1.56	107.67	70.69	108.52	1.09
IC C7	C8	C9	CA	1.54	113.63	177.94	113.11	1.55
IC C7	C8	C9	H92	1.54	113.63	-61.54	108.53	1.09
IC C7	C8	C9	H93	1.54	113.63	56.79	108.87	1.09
IC C8	C7	C6	H62	1.54	109.67	78.79	109.93	1.09
IC C8	C7	C6	H63	1.54	109.67	-42.05	109.77	1.09
IC C8	C9	CA	CB	1.53	113.11	-67.56	119.48	1.54
IC C8	C9	CA	HA2	1.53	113.11	171.07	106.90	1.09
IC C8	C9	CA	HA3	1.53	113.11	55.52	107.68	1.09
IC C9	C8	C7	H72	1.53	113.63	78.58	109.42	1.09
IC C9	C8	C7	H73	1.53	113.63	-41.30	109.44	1.09
IC C9	CA	CB	CC	1.55	119.48	-167.44	109.12	1.53
IC C9	CA	CB	HB2	1.55	119.48	-47.50	109.56	1.09
IC C9	CA	CB	HB3	1.55	119.48	72.66	109.56	1.09
IC CA	C9	C8	H82	1.55	113.11	-61.46	108.43	1.09
IC CA	C9	C8	H83	1.55	113.11	56.65	108.78	1.09
IC CA	CB	CC	CD	1.54	109.12	-168.50	116.66	1.53
IC CA	CB	CC	HC2	1.54	109.12	-47.49	107.62	1.09
IC CA	CB	CC	HC3	1.54	109.12	69.28	108.22	1.09
IC CB	CA	C9	H92	1.54	119.48	171.93	108.55	1.09
IC CB	CA	C9	H93	1.54	119.48	53.59	108.85	1.09
IC CB	CC	CD	CE	1.53	116.66	-168.77	110.30	1.54
IC CB	CC	CD	HD2	1.53	116.66	-48.63	109.26	1.09
IC CB	CC	CD	HD3	1.53	116.66	70.99	109.32	1.09
IC CC	CB	CA	HA2	1.53	109.12	-46.06	106.89	1.09
IC CC	CB	CA	HA3	1.53	109.12	69.46	107.66	1.09
IC CC	CD	CE	CF	1.53	110.30	-167.93	114.49	1.53
IC CC	CD	CE	HE2	1.53	110.30	71.35	108.18	1.09
IC CC	CD	CE	HE3	1.53	110.30	-46.37	108.61	1.09
IC CD	CC	CB	HB2	1.53	116.66	71.56	109.55	1.09
IC CD	CC	CB	HB3	1.53	116.66	-48.58	109.53	1.09
IC CD	CE	CF	CG	1.54	114.49	178.82	114.17	1.57
IC CD	CE	CF	HF2	1.54	114.49	58.13	108.25	1.09
IC CD	CE	CF	HF3	1.54	114.49	-59.73	108.67	1.09
IC CE	CD	CC	HC2	1.54	110.30	70.23	107.64	1.09
IC CE	CD	CC	HC3	1.54	110.30	-46.53	108.20	1.09
IC CE	CF	CG	HG1	1.53	114.17	-179.99	109.51	1.09
IC CE	CF	CG	HG2	1.53	114.17	60.01	109.44	1.09
IC CE	CF	CG	HG3	1.53	114.17	-59.97	109.50	1.09
IC CF	CE	CD	HD2	1.53	114.49	71.94	109.27	1.09
IC CF	CE	CD	HD3	1.53	114.49	-47.70	109.33	1.09
IC CG	CF	CE	HE2	1.57	114.17	-60.47	108.20	1.09
IC CG	CF	CE	HE3	1.57	114.17	57.26	108.61	1.09
IC H22	C2	C3	H32	1.09	106.99	-176.98	109.08	1.09
IC H22	C2	C3	H33	1.09	106.99	-57.60	109.25	1.09
IC H23	C2	C3	H32	1.09	107.77	67.43	109.08	1.09
IC H23	C2	C3	H33	1.09	107.77	-173.19	109.25	1.09
IC H32	C3	C4	H42	1.09	109.10	-178.04	109.41	1.09
IC H32	C3	C4	H43	1.09	109.10	62.04	109.43	1.09
IC H33	C3	C4	H42	1.09	109.21	62.59	109.41	1.09
IC H33	C3	C4	H43	1.09	109.21	-57.33	109.43	1.09
IC H42	C4	C5	H52	1.09	109.43	-177.59	108.09	1.09
IC H42	C4	C5	H53	1.09	109.43	64.89	108.52	1.09
IC H43	C4	C5	H52	1.09	109.46	-57.64	108.09	1.09
IC H43	C4	C5	H53	1.09	109.46	-175.16	108.52	1.09
IC H52	C5	C6	H62	1.09	108.08	72.88	109.90	1.09
IC H52	C5	C6	H63	1.09	108.08	-166.29	109.78	1.09
IC H53	C5	C6	H62	1.09	108.52	-169.60	109.90	1.09
IC H53	C5	C6	H63	1.09	108.52	-48.77	109.78	1.09
IC H62	C6	C7	H72	1.09	109.93	-161.18	109.41	1.09
IC H62	C6	C7	H73	1.09	109.93	-41.29	109.47	1.09
IC H63	C6	C7	H72	1.09	109.77	77.98	109.41	1.09
IC H63	C6	C7	H73	1.09	109.77	-162.13	109.47	1.09
IC H72	C7	C8	H82	1.09	109.42	-42.04	108.41	1.09
IC H72	C7	C8	H83	1.09	109.42	-160.11	108.75	1.09
IC H73	C7	C8	H82	1.09	109.44	-161.92	108.41	1.09
IC H73	C7	C8	H83	1.09	109.44	80.01	108.75	1.09
IC H82	C8	C9	H92	1.09	108.43	59.07	108.53	1.09
IC H82	C8	C9	H93	1.09	108.43	177.40	108.87	1.09
IC H83	C8	C9	H92	1.09	108.78	177.17	108.53	1.09
IC H83	C8	C9	H93	1.09	108.78	-64.50	108.87	1.09
IC H92	C9	CA	HA2	1.09	108.55	50.56	106.90	1.09
IC H92	C9	CA	HA3	1.09	108.55	-64.99	107.68	1.09
IC H93	C9	CA	HA2	1.09	108.85	-67.78	106.90	1.09
IC H93	C9	CA	HA3	1.09	108.85	176.68	107.68	1.09
IC HA2	CA	CB	HB2	1.09	106.89	73.87	109.56	1.09
IC HA2	CA	CB	HB3	1.09	106.89	-165.97	109.56	1.09
IC HA3	CA	CB	HB2	1.09	107.66	-170.60	109.56	1.09
IC HA3	CA	CB	HB3	1.09	107.66	-50.44	109.56	1.09
IC HB2	CB	CC	HC2	1.09	109.55	-167.43	107.62	1.09
IC HB2	CB	CC	HC3	1.09	109.55	-50.66	108.22	1.09
IC HB3	CB	CC	HC2	1.09	109.53	72.43	107.62	1.09
IC HB3	CB	CC	HC3	1.09	109.53	-170.80	108.22	1.09
IC HC2	CC	CD	HD2	1.09	107.64	-169.63	109.26	1.09
IC HC2	CC	CD	HD3	1.09	107.64	-50.01	109.32	1.09

IC HC3	CC	CD	HD2	1.09	108.20	73.61	109.26
1.09							
IC HC3	CC	CD	HD3	1.09	108.20	-166.77	109.32
1.09							
IC HD2	CD	CE	HE2	1.09	109.27	-48.79	108.18
1.09							
IC HD2	CD	CE	HE3	1.09	109.27	-166.50	108.61
1.09							
IC HD3	CD	CE	HE2	1.09	109.33	-168.42	108.18
1.09							
IC HD3	CD	CE	HE3	1.09	109.33	73.86	108.61
1.09							
IC HE2	CE	CF	HF2	1.09	108.20	178.84	108.25
1.09							
IC HE2	CE	CF	HF3	1.09	108.20	60.98	108.67
1.09							
IC HE3	CE	CF	HF2	1.09	108.61	-63.43	108.25
1.09							
IC HE3	CE	CF	HF3	1.09	108.61	178.71	108.67
1.09							
IC HF2	CF	CG	HG1	1.09	108.30	-59.32	109.51
1.09							
IC HF2	CF	CG	HG2	1.09	108.30	-179.33	109.44
1.09							
IC HF2	CF	CG	HG3	1.09	108.30	60.69	109.50
1.09							
IC HF3	CF	CG	HG1	1.09	108.66	58.56	109.51
1.09							
IC HF3	CF	CG	HG2	1.09	108.66	-61.45	109.44
1.09							
IC HF3	CF	CG	HG3	1.09	108.66	178.57	109.50
1.09							
IC O1	C2	*C1	O2	0.00	0.00	180.00	0.00
0.00							
IC C3	C1	*C2	H23	0.00	0.00	120.00	0.00
0.00							
IC C3	C1	*C2	H22	0.00	0.00	-120.00	0.00
0.00							
IC C4	C2	*C3	H33	0.00	0.00	120.00	0.00
0.00							
IC C4	C2	*C3	H32	0.00	0.00	-120.00	0.00
0.00							
IC C5	C3	*C4	H43	0.00	0.00	120.00	0.00
0.00							
IC C5	C3	*C4	H42	0.00	0.00	-120.00	0.00
0.00							
IC C6	C4	*C5	H53	0.00	0.00	120.00	0.00
0.00							
IC C6	C4	*C5	H52	0.00	0.00	-120.00	0.00
0.00							
IC C7	C5	*C6	H63	0.00	0.00	120.00	0.00
0.00							
IC C7	C5	*C6	H62	0.00	0.00	-120.00	0.00
0.00							
IC C8	C6	*C7	H73	0.00	0.00	120.00	0.00
0.00							
IC C8	C6	*C7	H72	0.00	0.00	-120.00	0.00
0.00							
IC C9	C7	*C8	H83	0.00	0.00	120.00	0.00
0.00							
IC C9	C7	*C8	H82	0.00	0.00	-120.00	0.00
0.00							
IC CA	C8	*C9	H93	0.00	0.00	120.00	0.00
0.00							
IC CA	C8	*C9	H92	0.00	0.00	-120.00	0.00
0.00							
IC CB	C9	*CA	HA2	0.00	0.00	120.00	0.00
0.00							
IC CB	C9	*CA	HA3	0.00	0.00	-120.00	0.00
0.00							
IC CC	CA	*CB	HB3	0.00	0.00	120.00	0.00
0.00							
IC CC	CA	*CB	HB2	0.00	0.00	-120.00	0.00
0.00							
IC CD	CB	*CC	HC2	0.00	0.00	120.00	0.00
0.00							
IC CD	CB	*CC	HC3	0.00	0.00	-120.00	0.00
0.00							
IC CE	CC	*CD	HD3	0.00	0.00	120.00	0.00
0.00							
IC CE	CC	*CD	HD2	0.00	0.00	-120.00	0.00
0.00							
IC CF	CD	*CE	HE3	0.00	0.00	120.00	0.00
0.00							
IC CF	CD	*CE	HE2	0.00	0.00	-120.00	0.00
0.00							
IC CG	CE	*CF	HF2	0.00	0.00	120.00	0.00
0.00							
IC CG	CE	*CF	HF3	0.00	0.00	-120.00	0.00
0.00							
IC HG2	CF	*CG	HG1	0.00	0.00	120.00	0.00
0.00							
IC HG2	CF	*CG	HG3	0.00	0.00	-120.00	0.00
0.00							

!The following residue MYR derived from SWISSparam output,
!partial charges derived from (PALM) in original CHARMM file and renamed to suit virus reconstructions
!Jason Roberts 2012 VIDRL/RMIT University 2012

RESI MYR	-1.000		
GROUP			
ATOM C1	CO2M	0.62	
ATOM O1	O2CM	-0.76	
ATOM O2	O2CM	-0.76	
ATOM C2	CR	-0.28	
ATOM C3	CR	-0.18	
ATOM C4	CR	-0.18	
ATOM C5	CR	-0.18	
ATOM C6	CR	-0.18	
ATOM C7	CR	-0.18	
ATOM C8	CR	-0.18	
ATOM C9	CR	-0.18	
ATOM C10	CR	-0.18	
ATOM C11	CR	-0.18	
ATOM C12	CR	-0.18	
ATOM C13	CR	-0.18	
ATOM C14	CR	-0.27	
ATOM H22	HCMM	0.09	
ATOM H23	HCMM	0.09	
ATOM H32	HCMM	0.09	
ATOM H33	HCMM	0.09	
ATOM H42	HCMM	0.09	
ATOM H43	HCMM	0.09	
ATOM H52	HCMM	0.09	
ATOM H53	HCMM	0.09	
ATOM H62	HCMM	0.09	
ATOM H63	HCMM	0.09	
ATOM H72	HCMM	0.09	
ATOM H73	HCMM	0.09	
ATOM H82	HCMM	0.09	
ATOM H83	HCMM	0.09	
ATOM H92	HCMM	0.09	
ATOM H93	HCMM	0.09	
ATOM H102	HCMM	0.09	
ATOM H103	HCMM	0.09	
ATOM H112	HCMM	0.09	
ATOM H113	HCMM	0.09	
ATOM H122	HCMM	0.09	
ATOM H123	HCMM	0.09	
ATOM H132	HCMM	0.09	
ATOM H133	HCMM	0.09	
ATOM H141	HCMM	0.09	
ATOM H142	HCMM	0.09	
ATOM H143	HCMM	0.09	
BOND C1	O1		
BOND C1	O2		
BOND C1	C2		
BOND H23	C2		
BOND H22	C2		
BOND C2	C3		
BOND H33	C3		
BOND H32	C3		
BOND C3	C4		
BOND H43	C4		
BOND H42	C4		
BOND C4	C5		
BOND H53	C5		
BOND H52	C5		
BOND C5	C6		
BOND H63	C6		
BOND H62	C6		
BOND C6	C7		
BOND H73	C7		
BOND H72	C7		
BOND C7	C8		
BOND H83	C8		
BOND H82	C8		
BOND C8	C9		
BOND H93	C9		
BOND H92	C9		
BOND C10	C9		
BOND H103	C10		
BOND H102	C10		
BOND C10	C11		
BOND H113	C11		
BOND H112	C11		
BOND C11	C12		
BOND H123	C12		
BOND H122	C12		
BOND C12	C13		
BOND H133	C13		
BOND H132	C13		
BOND C13	C14		
BOND H142	C14		
BOND H143	C14		
BOND H141	C14		
IMPH C1	O1	C2	O2
IMPH C2	C3	C1	H23
IMPH C2	H23	C1	H22
IMPH C3	C4	C2	H33
IMPH C3	H33	C2	H32
IMPH C4	C5	C3	H43
IMPH C4	H43	C3	H42
IMPH C5	C6	C4	H53
IMPH C5	H53	C4	H52
IMPH C6	C7	C5	H63
IMPH C6	H63	C5	H62
IMPH C7	C8	C6	H73
IMPH C7	H73	C6	H72
IMPH C8	C9	C7	H83
IMPH C8	H83	C7	H82

A5.7.2 File Name: Disoxaril.par

* ----
 * Built parameters for disoxaril.mol2
 * Thu Mar 24 01:54:06 CET 2011
 * Jason A. Roberts 2011 RMIT University/VIDRL
 * ----
 *

BONDS

C5A	OFUR	416.469	1.3600
CR	C5A	322.481	1.4710
HCMM	CR	342.991	1.0930
CR	CR	306.432	1.5080
CR	OR	363.214	1.4180
CB	OR	404.019	1.3760
CB	CB	401.068	1.3740
C5A	CB	372.641	1.4360
C5B	C5A	512.256	1.3770
C5B	N5B	320.682	1.3690
C5A	N5B	599.191	1.3130
HCMM	C5B	396.246	1.0800
HCMM	C5A	398.045	1.0800
HCMM	CB	381.853	1.0840
C5B	C5B	310.390	1.4180
C5B	CR	325.144	1.4690
C5B	N5A	594.297	1.3350
N5A	OFUR	342.271	1.3880

ANGLES

C5B	N5A	OFUR	128.675	103.4520
N5A	C5B	C5B	65.921	113.5700
N5A	C5B	CR	69.303	120.6400
C5B	C5B	CR	55.126	128.0610
C5B	C5B	C5A	62.322	108.2390
C5B	C5B	HCMM	39.293	127.4050
C5A	C5B	HCMM	36.055	126.1700
C5B	C5A	OFUR	74.484	110.1080
C5B	C5A	CR	53.039	131.3780
OFUR	C5A	CR	84.560	115.2530
CB	CB	CB	48.145	119.9770
CB	CB	OR	69.663	116.4950
CB	CB	HCMM	40.517	120.5710
CB	CB	C5A	64.337	120.1900
CB	C5A	N5B	62.682	128.1300
CB	C5A	OFUR	87.366	114.2110
N5B	C5A	OFUR	84.991	115.5920
C5A	N5B	C5B	86.791	103.7790
N5B	C5B	C5A	74.700	111.6210
N5B	C5B	HCMM	50.304	120.4780
C5B	C5A	HCMM	41.524	131.7210
OFUR	C5A	HCMM	56.421	114.0760
N5A	OFUR	C5A	125.940	107.7550
C5B	CR	HCMM	44.763	110.4570
HCMM	CR	HCMM	37.134	108.8360
C5A	CR	CR	72.397	110.0580
C5A	CR	HCMM	44.691	110.4670
CR	CR	HCMM	45.770	110.5490
CR	CR	CR	61.243	109.6080
CR	CR	OR	71.390	108.1330
OR	CR	HCMM	56.205	108.5770
CB	OR	CR	77.363	102.8460
C5A	OFUR	C5A	91.612	106.3130

DIHEDRALS

N5A	C5B	C5B	C5A	3.500	2	180.00
N5A	C5B	C5B	HCMM	3.500	2	180.00
N5A	C5B	CR	HCMM	0.000	1	0.00
N5A	OFUR	C5A	C5B	3.500	2	180.00
N5A	OFUR	C5A	CR	3.500	2	180.00
C5B	N5A	OFUR	C5A	3.500	2	180.00
C5B	C5B	C5A	OFUR	3.500	2	180.00
C5B	C5B	C5A	CR	3.500	2	180.00
C5B	C5B	N5A	OFUR	3.500	2	180.00
C5B	C5B	CR	HCMM	0.000	1	0.00
C5B	C5A	CR	CR	0.000	1	0.00
C5B	C5A	CR	HCMM	0.000	1	0.00
C5A	C5B	C5B	CR	3.500	2	180.00
C5A	CR	CR	CR	0.150	3	0.00
C5A	CR	CR	HCMM	0.150	3	0.00
CB	CB	CB	CB	3.500	2	180.00
CB	CB	CB	HCMM	3.500	2	180.00
CB	OR	CR	CR	0.100	3	0.00
CB	OR	CR	HCMM	0.053	3	0.00
CB	CB	OR	CR	2.191	2	180.00
CB	CB	CB	C5A	3.500	2	180.00
CB	CB	CB	OR	3.500	2	180.00
CB	CB	C5A	N5B	3.500	2	180.00
CB	CB	C5A	OFUR	3.500	2	180.00
CB	C5A	N5B	C5B	0.000	1	0.00
CB	C5A	OFUR	C5A	0.000	1	0.00
C5A	CB	CB	HCMM	1.000	2	180.00
C5A	N5B	C5B	C5A	3.500	2	180.00
C5A	N5B	C5B	HCMM	3.500	2	180.00
C5A	OFUR	C5A	C5B	3.500	2	180.00
C5A	OFUR	C5A	HCMM	3.500	2	180.00
N5B	C5A	OFUR	C5A	3.500	2	180.00
N5B	C5B	C5A	OFUR	3.500	2	180.00
N5B	C5B	C5A	HCMM	3.500	2	180.00
C5B	N5B	C5A	OFUR	3.500	2	180.00

OFUR	N5A	C5B	CR	3.500	2	180.00	OFUR	C5A	C5B
HCMM		3.500	2	180.00					
OFUR	C5A	CR	CR	0.000	1	0.00			
OFUR	C5A	CR	HCMM	0.000	1	0.00			
CR	C5B	C5B	HCMM	3.500	2	180.00			
CR	C5A	C5B	HCMM	3.500	2	180.00			
CR	CR	CR	CR	0.051	1	0.00			
CR	CR	CR	CR	0.341	2	180.00			
CR	CR	CR	CR	0.166	3	0.00			
CR	CR	CR	HCMM	0.320	1	0.00			
CR	CR	CR	HCMM	-0.315	2	180.00			
CR	CR	CR	HCMM	0.132	3	0.00			
CR	CR	CR	OR	-0.344	1	0.00			
CR	CR	CR	OR	0.878	2	180.00			
CR	CR	CR	OR	0.238	3	0.00			
OR	CB	CB	HCMM	3.500	2	180.00			
OR	CR	CR	HCMM	-0.327	1	0.00			
OR	CR	CR	HCMM	0.536	2	180.00			
OR	CR	CR	HCMM	0.140	3	0.00			
HCMM	CR	CR	HCMM	0.142	1	0.00			
HCMM	CR	CR	HCMM	-0.693	2	180.00			
HCMM	CR	CR	HCMM	0.157	3	0.00			
HCMM	CB	CB	HCMM	3.500	2	180.00			
HCMM	C5A	C5B	HCMM	3.500	2	180.00			

IMPROPER

C5B	C5B	N5A	CR	2.879	0	0.00
C5A	C5B	OFUR	CR	3.598	0	0.00
CR	CR	C5A	HCMM	0.000	0	0.00
CR	HCMM	C5A	HCMM	0.000	0	0.00
CR	CR	CR	HCMM	0.000	0	0.00
CR	HCMM	CR	HCMM	0.000	0	0.00
CR	OR	CR	HCMM	0.000	0	0.00
CB	CB	OR	CB	3.454	0	0.00
CB	CB	CB	HCMM	1.079	0	0.00
CB	CB	CB	C5A	2.519	0	0.00
C5A	OFUR	CB	N5B	3.598	0	0.00
C5A	C5B	OFUR	HCMM	2.375	0	0.00
C5B	C5A	N5B	HCMM	3.095	0	0.00
C5B	C5A	C5B	HCMM	0.432	0	0.00
CR	HCMM	C5B	HCMM	0.000	0	0.00

NONBONDED nbxmod 5 atom cdiel shift vatom vdistance
 vswitch -
 cutnb 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0
 wmin 1.5

N5A	0.000000	-0.200000	1.850000		
C5B	0.000000	-0.050000	2.040000		
C5A	0.000000	-0.050000	2.040000		
CB	0.000000	-0.070000	1.992400		
N5B	0.000000	-0.200000	1.850000		
OFUR	0.000000	-0.200000	1.850000		
CR	0.000000	-0.055000	2.175000	0.000000	-
	0.010000	1.900000			
OR	0.000000	-0.152100	1.770000		
HCMM	0.000000	-0.022000	1.320000		

A5.7.3 File Name: 1EAH_SCH4.par

```
* -----
* Built parameters for 1EAH_SCH4+H.mol2
* Fri Oct 7 00:21:01 CEST 2011
* Jason A. Roberts 2011 RMIT University/VIDRL
* -----
*
BONDS
CB CB 401.068 1.3740
CB CR 356.737 1.4860
CR OR 363.214 1.4180
CB OR 404.019 1.3760
CB CLMM 243.102 1.7210
HCMM CB 381.853 1.0840
HCMM CR 342.991 1.0930

ANGLES
CB CB CB 48.145 119.9770
CB CB HCMM 40.517 120.5710
CB CB OR 69.663 116.4950
CB CB CLMM 68.367 118.4950
CB CB CR 57.788 120.4190
CB OR CR 77.363 102.8460
CB CR OR 63.186 107.9780
CB CR HCMM 45.122 109.4910
OR CR HCMM 56.205 108.5770
HCMM CR HCMM 37.134 108.8360

DIHEDRALS
CB CB CB CB 3.500 2 180.00
CB CB CB OR 3.500 2 180.00
CB CB CB HCMM 3.500 2 180.00
CB CB OR CR 2.191 2 180.00
CB CB CB CLMM 3.500 2 180.00
CB OR CR CB 0.100 3 0.00
CB OR CR HCMM 0.053 3 0.00
CB CB CR OR 0.075 3 0.00
CB CB CR HCMM -0.210 2 180.00
CB CB CR HCMM 0.196 3 0.00
CB CB CB CR 3.500 2 180.00
CLMM CB CB OR 3.500 2 180.00
CLMM CB CB HCMM 3.500 2 180.00
OR CB CB HCMM 3.500 2 180.00
CR CB CB HCMM 3.500 2 180.00
HCMM CB CB HCMM 3.500 2 180.00

IMPROPER
CB CB CB HCMM 1.079 0 0.00
CB OR CB CB 3.454 0 0.00
CB CLMM CB CB 2.519 0 0.00
CR CB OR HCMM 0.000 0 0.00
CB CB CR CB 2.879 0 0.00
CB CB CB OR 3.454 0 0.00
CR OR CB HCMM 0.000 0 0.00
CB CB CB CLMM 2.519 0 0.00
CR HCMM OR HCMM 0.000 0 0.00

NONBONDED nbxmod 5 atom cdiel shift vatom vdistance vswitch -
cutnb 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5
CB 0.000000 -0.070000 1.992400
CLMM 0.000000 -0.260000 2.040000
OR 0.000000 -0.152100 1.770000
CR 0.000000 -0.055000 2.175000 0.000000 -0.010000 1.900000
HCMM 0.000000 -0.022000 1.320000
```

A5.8 NAMD NPT simulation parameters for BlueGene/P and /Q supercomputers with file compression activated and custom load balancing intervals.

File Name: ExampleBlueGeneSimulationRun.conf

```
#####
## JOB DESCRIPTION ##
#####
# Generic optimization/equilibration script
#####
## ADJUSTABLE PARAMETERS ##
#####

# structure          Examplexxxx.psf
# coordinates        Examplexxxx.pdb

usecompressedpsf    on

structure           Examplexxxx.psf.inter
binCoordinates      Examplexxxx.coor

outputName          Examplexxxx_EM+ins

## Memory optimization parameters (useful for BlueGene node 0 issues)

# noPatchesOnZero yes
# ldbUnloadZero yes
# maxSelfPart 1
# maxPairPart 1
# pairlistMinProcs 128
# PMEPencils 1

## Load balancing parameters

ldbPeriod 50000
firstLdbStep 5000

#####
## SIMULATION PARAMETERS ##
#####

# Input
paraTypeCharmm      on
parameters          par_all27_prot_lipid_na.inp
parametersPicornavirus_small_molecules.par

set Temp            310
temperature          $Temp

# constraints        on
# consexp            2
# consref            xxxx.pdb
# conskfile          xxxx.pdb
# conskcol           B
# constraintScaling  0.001

# Force-Field Parameters
exclude             scaled1-4
l-4scaling          1.0
cutoff              12
switching           on
switchdist          10
pairlistdist        14

# Integrator Parameters
timestep            2
rigidBonds          all
nonbondedFreq       1
fullElectFrequency  2
stepspercycle       10

# Constant Temperature Control
langevin            on
langevinDamping     5
langevinTemp        $Temp
langevinHydrogen    off

# Periodic Boundary Conditions

cellBasisVector1    364.425 0.0 0.0
cellBasisVector2    0.0 364.425 0.0
cellBasisVector3    182.212 182.212 257.688

cellOrigin           0 0 0

wrapAll              on
wrapNearest          on
```



```
# PME (for full-system periodic electrostatics)
PME                yes
PMEGridSpacing     1.0

# Constant Pressure Control (variable volume)
useGroupPressure   yes
useFlexibleCell    no
useConstantArea    no

langevinPiston     on
langevinPistonTarget 1.01325
langevinPistonPeriod 100.
langevinPistonDecay 50.
langevinPistonTemp $Temp

# Output

restartfreq        50000
dcdfreq            50000
xstFreq            5000
outputEnergies     5000
outputPressure     5000
outputTiming       5000

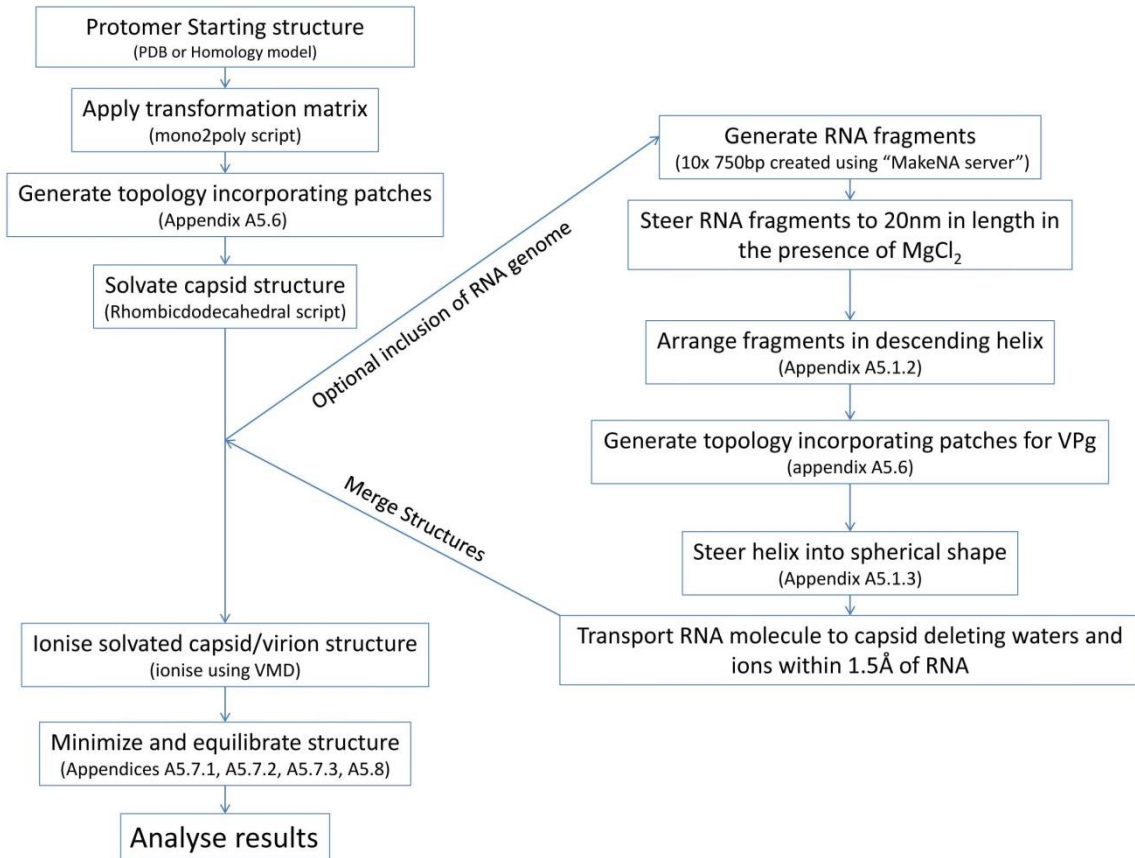
#####
## EXECUTION SCRIPT                               ##
#####

# Minimization

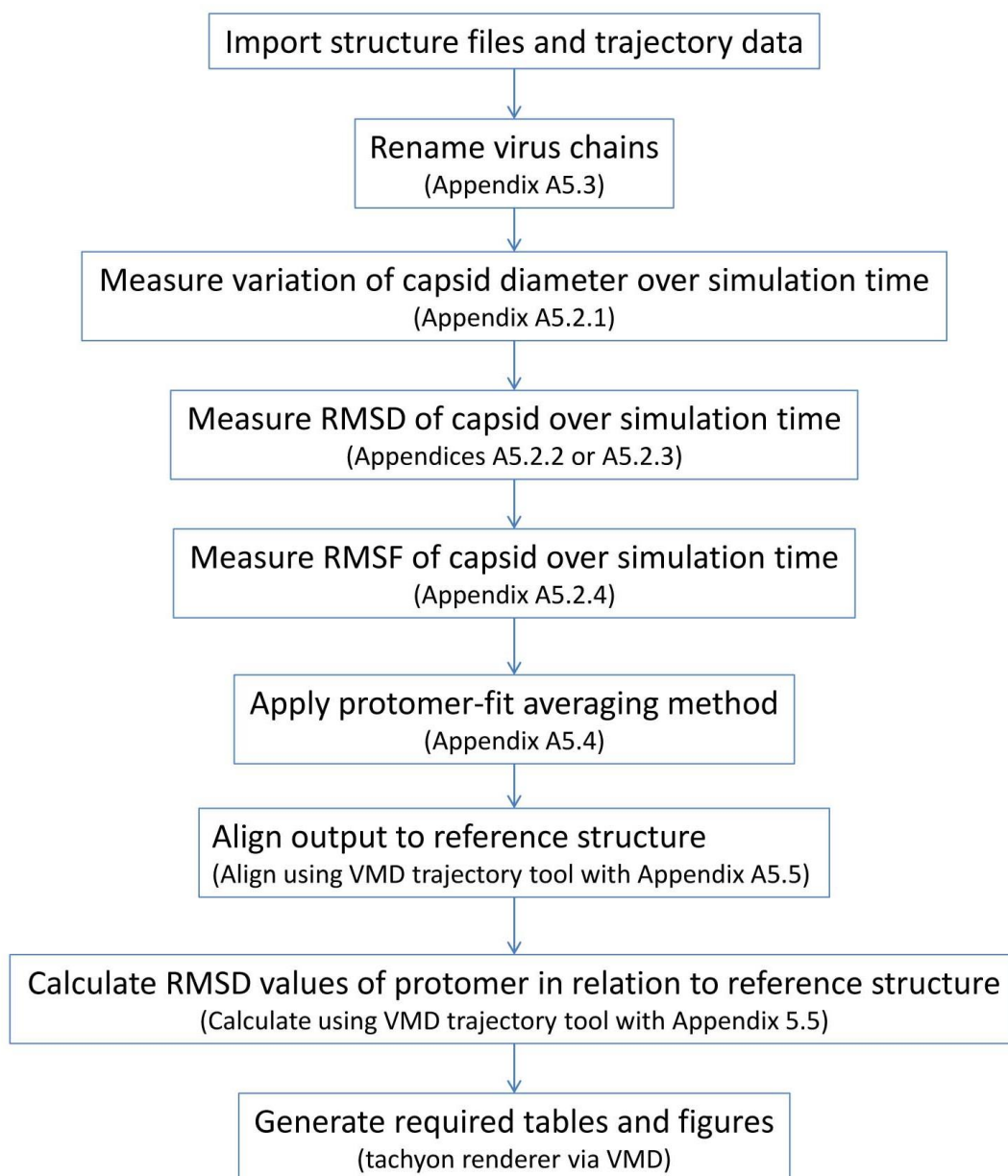
minimize           50000
run                 500000

#
```

A5.9 Simulation Construction Flowchart



A5.10 Simulation Analysis Flowchart



A6 Supporting Grants

2012-2015 *Title: National Poliovirus Laboratory and AFP surveillance System*

(incorporating environmental surveillance and enterovirus surveillance).

Department of Health and Ageing. Total grant value \$691,764.

Principle Investigator - Thorley, B. R., Investigator - **Roberts, J. A.**,

2011-2014 *Title: Molecular Modelling of Novel Enterovirus Proteins Associated With*

Acute Flaccid Paralysis. Victorian Life Sciences Computation Initiative,

Resource Allocation Scheme. 8.2 million service units awarded in total.

Principle Investigators - **Roberts, J. A.**, Thorley, B. R., Investigators -

Hung, A., Smooker, P., Kuiper, M.

2012-2013 *Title: Examination of the Specific Mechanisms of Vaccine-Derived Poliovirus*

Antiviral Drug Interactions. Victorian Life Sciences Computation

Initiative, Resource Allocation Scheme. 2.4 million service units awarded.

Principle Investigators - **Roberts, J. A.**, Thorley, B. R., Investigators -

Hung, A., Oberste, M. S.

2011-2012 *Title: Fighting Superbugs with Supercomputers*. nVidia Academic Partnership

Program. Equipment grant in kind, to the value of \$24,722.

Principle Investigators - **Roberts, J. A.**, Kuiper, M. J, Investigators - Hurt,

A., Hung, A., Wilson, D., Thorley, B.

A7 Awards

2013 – *Melbourne Health*, “Celebrating Excellence Awards” Healthcare Innovation Award, Highly Commended.

2013 – *Australian Institute of Medical and Biological Illustrators*, “Illuminate” conference awards exhibition, two bronze, one silver and one gold award.

2013 - *CSIRO Computational and Simulation Sciences Annual Conference*, Winner poster session.

2012 - *IDC HPC Innovation Excellence Award*, International Award For the Outstanding Application of HPC Computing for Business and Scientific Achievements. November 2012.

2012 – *RMIT University*, *Student Graduate Research Travel Award*, \$2000.

2012 – *Victorian Life Sciences Computation Initiative*, *Travel Award*, \$4000.

2011 - *Association of Molecular Modellers Australasia*, 2011 Meeting of Molecular Modellers, Best video featuring molecular modelling simulation.

A8 Dissemination Details

A8.1 Conference Proceedings, Seminars and Public Media

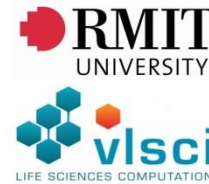
- 2014** - (Invited Speaker) Detection, Characterisation and 3D Molecular Modelling of Enteroviruses from Cases of Acute Flaccid Paralysis, *Public Health Night and ASM General Meeting, Australian Society of Microbiology (Victorian Branch)*, Melbourne, Australia.
- 2014** - (Invited Speaker) Supercomputer Simulation of Viruses, The Merging of Art and Science, *Institute of Photographic Technology*, Melbourne, Australia.
- 2014** - (Invited Speaker). Discovery and In-Silico Reconstruction of a Novel Virus Identified From a Patient Suffering Paralysis, *Melbourne Meeting of Molecular Modellers*, LaTrobe University, Melbourne, Australia.
- 2013** - (Invited Speaker). Supercomputer Simulation of Newly Discovered and Novel Enteroviruses, *University of Queensland, Australian Institute for Bioengineering and Nanotechnology Special Seminar*, Queensland, Australia.
- 2013** - (Invited Speaker). Supercomputer Simulation of Viruses, The Merging of Art and Science 2013, *Australian Institute of Medical and Biological Illustrators, "Illuminate" biennial conference*, Melbourne Australia.
- 2013** – (Television) ABC Television, "Catalyst", Series 14, Episode 17, aired 15th August, *Featured story - 3D Virus*.
- 2013** - (Invited Speaker). Molecular Dynamics Simulation of Viruses 2013, *RMIT University, 4th Workshop on Computational Modelling of Proteins and Membranes*, Melbourne Australia.
- 2013** – (Newspaper article) The Age, Melbourne Magazine: (also Sydney Morning Herald.) Feature Story: *Melbourne research that could change your life*. Including cover art.
- 2013** - (Poster) Refinement of the wild poliovirus capsid structure by atomistic molecular dynamics simulation of a complete virion. Roberts, J. A., et. al., *CSIRO Computational and Simulation Sciences and eResearch Annual Conference*, Melbourne, Australia..
- 2013** – (Newspaper article) The Australian, Page 3, 22nd January, *Scientists zoom in on a killer virus*.

- 2012** - (Invited Speaker). Computational Biophysics. Case study: In-Silico Reconstruction of Poliovirus. *University of Adelaide, AMSI Summer Symposium in Bioinformatics*. Adelaide Australia.
- 2012** - (Invited Speaker) Supercomputer Modelling of a Complete Human Viral Pathogen: Poliovirus. *Centers for Disease Control and Prevention, Bioinformatics Seminar Series*. Atlanta GA, United States.
- 2012** – (Presentation) Supercomputer Modelling of a Complete Human Viral Pathogen: Poliovirus. *Theoretical and Computational Biology Group, Beckman Institute, University of Illinois*. Champaign-Urbana IL, United States.
- 2012** - (Presentation and poster) Supercomputer Modelling of the Complete Human Viral Pathogen: Poliovirus. *Higher Degrees Research Student Conference 'From Inception to Excellence'*, RMIT University, Melbourne, Australia.
- 2012** - (Poster) Refinement of the wild poliovirus capsid structure by atomistic molecular dynamics simulation of a complete virion. Roberts, J. A., et. al., *17th European Study Group on the Molecular Biology of Picornaviruses conference*, St Raphael, France.
- 2012** - (Poster) Supercomputer Modelling of Infectious Poliovirus., Roberts, J. A., Thorley, B. R., *Melbourne Health Research Week*, Melbourne, Australia..
- 2012** – (Newspaper article) The Age, Page 6, 15th January, Polio research reveals devil in Blue Gene.
- 2011** - (Poster) Atomistic Molecular Dynamics Simulation of Complete Human Pathogenic Virions. Roberts, J., et al. *6th Australasian Virology Group meeting*, Kingscliff, Australia.
- 2011** - (Presentation) Application of Graphics Processing Unit Technology to Large-Scale Atomistic Molecular Dynamics Simulation of Pathogenic Virus Structures. *Melbourne Meeting of Molecular Modellers*, Melbourne, Australia..
- 2011** - (Invited Speaker). In-silico Reconstruction of Poliovirus Virions, *4th Sino-Australian Meeting on Infectious Diseases*, VIDRL Melbourne, Australia.

A8.1.1 Poster: Refinement of the wild poliovirus capsid structure by atomistic molecular dynamics simulation of a complete virion.



Refinement of the wild poliovirus capsid structure by atomistic molecular dynamics simulation of a complete virion.



Jason A. Roberts^{1,3}, Bruce R. Thorley^{1,3}, Mike J. Kuiper², Peter M. Smooker³, Andrew Hung³

¹Enterovirus Reference Laboratory, Victorian Infectious Diseases Reference Laboratory, North Melbourne, Australia
²Victorian Life Sciences Computation Initiative, Parkville, Australia
³School of Applied Sciences and Health Innovations Research Institute, RMIT University, Melbourne, Australia

Introduction

The poliovirus virion is composed of a single stranded positive sense RNA genome enclosed by a non-enveloped capsid composed of 60 copies of four viral proteins: VP1, VP2, VP3 and VP4. One copy of each VP protein forms a protomer subunit and five individual protomers form a pentamer structure. The complete capsid has an icosahedral formation made from 12 pentamers. Integral to the capsid structure are myristate molecules covalently bound to the N-terminus of VP4 and an endogenous lipid occupying a hydrophobic pocket accessible from the exterior surfaces of VP1 that is common to enteroviruses. The RNA genome is polyadenylated and VPg, a virus encoded protein, is covalently attached to the 5' terminus.

The computational power of supercomputers now enables molecular dynamics simulations of multi-million atom systems, facilitating the *in silico* investigation of macromolecular structures. We used atomistic molecular dynamics simulation methodology to build a complete poliovirus, including the native RNA sequence. We believe that all-atom simulations of enteroviruses will provide insights to significant molecular interactions and biological properties of the virion that may not be discerned from modeling of individual structural units of the use of static models.

Aim

To validate the atomistic molecular dynamics simulation of the complete wild poliovirus type 1 (Mahoney strain) and compare various structural elements with observed phenomena and published data.

Methods

Wild poliovirus type 1 (Mahoney strain) was reconstructed from the 1HXS pdb x-ray crystallography file using the Visual Molecular Dynamics (VMD) software package incorporating Compute Unified Device Architecture extensions. Topology files and CHARMM parameter data for the associated myristate and palmitate molecules were determined using the SWISSParam server.

The amino terminus of the VP1 protein was not resolved by x-ray crystallography. The nineteen amino acid residues constituting an amphipathic α -helix were reconstructed using the molefacture plug-in of VMD with spatial alignment based on the positioning of partial coordinate data in the original 1HXS pdb file.

The icosahedral capsid structure was constructed from metadata within the original crystal structure. The "mono2poly" Tcl script was used to position a single protomer according to the biological parameters and repeated to assemble the 60 virus protomers.

Initially, 50,000 steps of energy minimisation were performed followed by equilibration for 0.5 ns to 310 degrees Kelvin at one atmosphere using 256 cores of an SGI Altix XE supercomputer. Once equilibrated the virus simulation was moved to an IBM BlueGene/P supercomputer running Nanoscale Molecular Dynamics 2.8 beta 1 with memory optimisation extensions on 2048 cores. Total simulation times ranged from 10 nanoseconds to 0.1 microseconds. Trajectory data for each simulation (table 1) were then analysed on local computers. (1)

Simulation	Detached helix	Pentamer	Empty capsid	Complete virion
Total atoms	8,598	388,821	5,858,950	2,754,142
PBC	Cubic	Cubic	Cubic	Triclinic/orthorhombic
Amino Acids	20	4,390	52,680	52,680
Nucleotides (RNA)	0	0	0	7,640
Fatty Acids	0	10	120	120
Water molecules	2,760	106,620	1,038,857	563,929
Sodium ions	10	370	3,728	2,581
Magnesium ions	0	0	0	3,150
Chloride ions	8	309	3,011	1,927

Table 1: Table outlining specific details of each simulation including periodic boundary conditions (PBC) and total atoms used.

Results

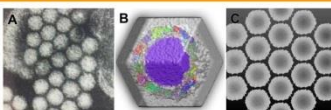


Figure 1. (A) Transmission electron micrograph of picornaviruses showing honeycomb-like aggregation of virus particles. (B) MD simulation of the full poliovirus virion suspended in saline solution with transparent density representation of water and a rhombic dodecahedral periodic cell boundary. (C) Transverse section of the MD simulation periodic cell reflections illustrating a similar honeycomb-like appearance as in (A).

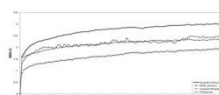


Figure 2. Root Mean Square Deviation (RMSD) of the poliovirus empty capsid, full virion, spherical RNA and an individual pentamer over the 10 ns simulation time.

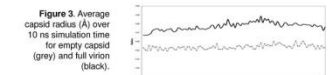


Figure 3. Average capsid radius (R) over 10 ns simulation time for empty capsid (grey) and full virion (black).

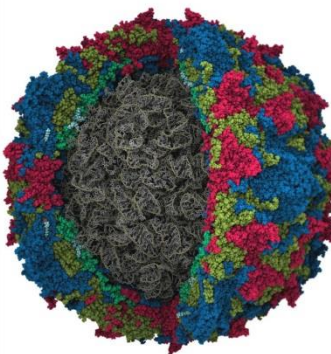


Figure 4. Poliovirus type 1 (Mahoney strain) virion. Total simulation size was approximately 2.8 million atoms. Water and surrounding ions were removed from this representation. Capsid components are represented as Van der Waals models and coloured by subunit: VP1 = blue, VP2 = yellow, VP3 = red, VP4 = green. Lipids are coloured according to the constitutive elements. RNA genome is shown using a "ligonice" representation (grey).

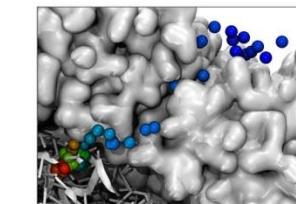


Figure 5. Migration path of a single sodium ion through a pore-like structure located at the base of the canyon (5-fold axis of symmetry): 0 ns = blue to 10 ns = red. A sodium ion transversed the pore in approximately 500 ps. Capsid protein is depicted in silver, with RNA represented in cartoon (white).

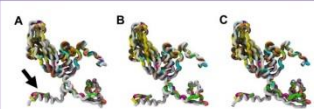


Figure 6. View of fit-averaged VP1 structural protein (white) overlaid on ten poliovirus type 1 crystal structure coordinates (coloured) and the original template crystal structure 1HXS (blue) from the RCSB PDB. Arrow indicates the position of the reconstructed amphipathic α -helix after 10 ns simulation of (A) pentamer (n=5), (B) empty capsid (n=60) and (C) full virion (n=60).

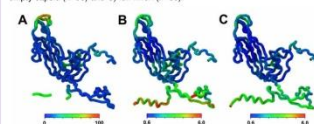


Figure 7. (A) VP1 protein of the 1HXS template crystal structure coloured by beta value. (B) and (C) are coloured "RMSD by residue" of fit-averaged VP1 structural proteins (n=60) at 10 ns time point for (B) empty capsid and (C) full virion.

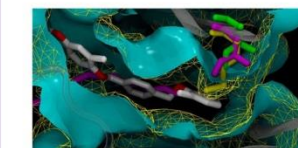


Figure 8. Simulation output showing conformational change in the hydrophobic pocket of an antiviral resistant Vaccine Derived Poliovirus pentamer. Endogenous lipid (purple), antiviral (white), Point mutations isoleucine to methionine. Variations in pocket architecture shown as details in wild type (cyan) and mutant (yellow).

Discussion and Conclusions

The use of a rhombic dodecahedral periodic cell boundary reduced the total atom count and mirrored native virus aggregation (Fig. 1). The MD simulations were stable as observed by the RMSD of the individual components (Fig. 2) and the radius of the empty capsid and full virion (Fig. 3). The MD simulation represented a complete poliovirus virion including the endogenous RNA genome (Fig. 4) that also enabled ionic movement to be studied (Fig. 5). The VP1 amphipathic α -helix was reconstructed from the available crystal data (Fig. 6) and was observed to have a high degree of flexibility (Fig. 7). Construction of the virion and biologically relevant precursors permits the examination of antiviral drug resistant mutants of Poliovirus (Fig. 8). (2)

Future work will focus on re-positioning the VP1 amphipathic α -helix and refinement of the RNA stem-loop structure. Furthermore, we will undertake homology modeling of other enteroviruses of public health significance such as vaccine derived polioviruses, enterovirus 71 and newly described and novel enteroviruses.


Acknowledgements

We thank Dallas Wilson, Thomas Alken and Mike Catton of the Victorian Infectious Diseases Reference Laboratory for supportive contributions at various stages of this work. We thank Cyril Fiebold, Gisela Meyer and Matthew Dowton for their kind assistance with the generation of the rhombic dodecahedral modelling conditions. We thank the VIDRL Electron Microscopy Unit for their kind provision of the picornavirus transmission electron microscopy image. This research was supported by a Victorian Life Sciences Computation Initiative (VLSI) grant number V10069 on its Peak Computing Facility at the University of Melbourne, an initiative of the Victorian Government, Australia. The National Enterovirus Reference Laboratory and WHO Regional Poliovirus Reference Laboratory are supported by funding from the Victorian State Government, the Department of Health and Ageing, Australia and the World Health Organization.

References


- Roberts, J.A., Kuiper, M.J., Thorley, B.R., Smooker P.M., Hung, A., 2012. Investigation of a predicted N-terminal amphipathic α -helix using atomistic molecular dynamics simulation of a complete prototype poliovirus virion. *Journal of Molecular Graphics and Modelling*, 38, September: 165-173.
- Liu, H., Roberts, J.A., Moore, D., Anderson, B., Pallansch, M.A., Peave, D.C., Collett M.S., Oberste, M.S., 2012. Characterization of Poliovirus Variants Selected for Resistance to the Antiviral Compound V-075. *Antimicrobial Agents and Chemotherapy*, 56 (11): 5568-5574.

A8.1.2 Poster: Atomistic molecular dynamics simulation of complete human pathogenic virions



VIDRL

Atomistic Molecular Dynamics Simulation of Complete Human Pathogenic Virions



Jason A. Roberts^{1, 2}, Mike J. Kuiper³, Bruce R. Thorley^{1, 2}, Peter M. Smooker², Andrew Hung².

1. World Health Organization Poliovirus Regional Reference Laboratory, Victorian Infectious Diseases Reference Laboratory, Melbourne, Victoria, Australia.
 2. Health Innovations Research Institute and School of Applied Sciences, RMIT University, Melbourne, Victoria, Australia.
 3. Victorian Partnership of Advanced Computing, Melbourne, Victoria, Australia.

Background

In 1989, the World Health Assembly initiated the Global Poliovirus Eradication Initiative (GPEI) when there were more than 350,000 polio cases annually. Since then there has been a reduction of more than 99% of paralytic cases globally, but as global eradication of wild poliovirus nears a number of issues have emerged. In particular, the phenomenon of vaccine derived polioviruses (VDPVs), neurovirulent forms of poliovirus originating from vaccine strains with the potential for sustained transmission. Molecular dynamics (MD) simulations of complete virions may enable us to elucidate the atomic-level mechanisms of neurovirulent or antiviral-resistant mutations. Studying the tertiary structural evolution of virions will thereby form a vital complement to sequence analyses of these viral strains. Until recently, computer modelling of virus capsid dynamics to investigate amino acid mutations that result in neurovirulent or antiviral-resistant phenotypes, was impractical. The first report of a full virus simulation was of Satellite Tobacco Mosaic Virus in 2006, with the full simulation including water and ions totalling 1 million atoms¹. In comparison, a single poliovirus contains over one million atoms with an additional 5 million atoms required to solvate the virus in a small cube measuring 35 nanometres on each edge. Computer hardware and software is now available that allows researchers to construct atomic models greater than 100 million atoms using a suitably equipped desktop computer. We used these methods to build a complete 3-dimensional model of a prototype poliovirus complete with proteins, lipids and RNA suspended in a virtual saline solution. The *in silico* construction and simulation of prototype poliovirus, which is analogous to a number of other icosahedral RNA viruses, will allow researchers to gain insight into the molecular mechanisms of virus interaction with the host cellular environment and also antivirals.

Results

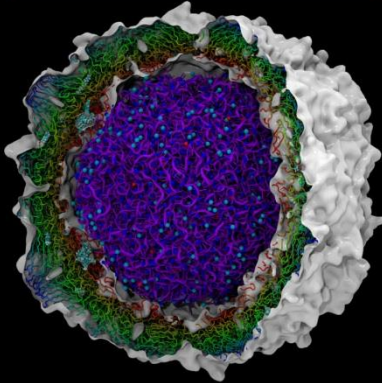


Figure 1 Poliovirus type 1 (Mahoney strain) after 10 ns of molecular dynamics simulation. Water and surrounding ions have been removed from this representation. Atomic density calculated at 2 Å resolution with a 6 Å radius is represented in white, with proteins represented as tubes with RGB colour shading from 110 to 150 Å. Lipids are represented as VDW models with colouring by name of element. RNA is shown in 'heat colour' representation (purple) with associated ions (Dark blue - Magnesium, Light blue - Sodium, Red - Chloride).

Discussion

This work represents one of the first reports describing the reconstruction and atomistic molecular dynamics simulation of a complete human viral pathogen containing the encoding genomic RNA (figure 1). Numerous issues with the generation of such large models for simulation were encountered when constructing this virus, with problems ranging from computer hardware requirements to software compatibility. Given the unknown nature of the specific packing mechanisms and final conformation of the RNA within the virion, it was decided to restructure a supercoiled version of the RNA genome, although further work involving the folding of RNA according to minimal energy states will be investigated. We find that after an initial conformational shift during the energy minimisation stage, the virus begins to settle after approximately 7 ns of simulation, with final capsid protein RMSD values below 3.1 Å at the 10 ns time point (figure 2). The final fit-averaged protomer (figure 3) showed a RMSD of 1.65 Å when compared to the original template crystal structure (1HXS). The release of RNA under heat stress conditions has been described using cryo-electron microscopy, with expulsion of RNA detected around the region of the 2-fold axis of symmetry². Molecular dynamics simulation of the complete virion with varying temperature conditions revealed a series of breaches in capsid integrity around the 2-fold axis and an associated movement of water and ions across the capsid. Subjecting the virion core to severe ionic disruption resulted in expansion of the RNA and capsid with failure points accumulating around the positions in the 2-fold axis of symmetry and partial expulsion of RNA material (figure 4). Neurovirulent forms of poliovirus are known to exhibit reduced heat lability and as such the replication *in-silico* of heat stress should assist in the elucidation of the underlying mechanics of heat induced virion degradation and the mechanisms that delay this process. Such insight may aid researchers in understanding the specifics of antiviral resistance and assist in the development of alternate antiviral compounds. Such compounds are of great interest to the GPEI as adequate therapeutic options for the chronic excretion of poliovirus are not currently available. The development of novel antiviral compounds may also prove useful for related viruses such as enterovirus 71 or the human rhinoviruses. The computational methods described can be applied to virtually any virus that displays icosahedral symmetry as long as structural data are available and the sequence identity of the virus in question is >30% compared to the reference template structure.

Methods




Poliovirus capsid reconstruction was based on data from the public domain, specifically, GeneBank sequence ID NC002058. PSIPred predictions³ derived from the amino acid sequence data and comparative information determined from enterovirus crystal structures available in the Research Collaboratory for Structural Bioinformatics protein data bank (RCSB PDB) was used to select an appropriate template for the reconstruction of the capsid and associated lipids using the SWISS-model server⁴. A template in the form of X-ray crystallographic data of poliovirus type 1 Mahoney strain (PDB file 1HXS) at 2.2 Å resolution⁵ was selected. Reconstruction of the virus was achieved using the 64-bit Linux variant of Visual Molecular Dynamics (VMD) ⁶ software incorporating Compute Unified Device Architecture extensions⁷. A significant element lacking from the original crystal structure was an α -helix located at the amino terminus of the VP1 protein common to members of the genus enterovirus. This helix was manually created and placed in the appropriate position determined from partial x-ray data. Topology files and CHARMM parameter data for the associated myristate and palmitate molecules were determined using the SWISSParam server (http://swissparam.ch⁸). Icosahedral symmetry data in the form of metadata contained within the original crystal structure was then used to assemble the 240 proteins and 120 lipids resulting in full capsid molecules of approximately 800,000 atoms. The genomic RNA was split into 10 fragments of approximately 750 bp each and a 20 bp poly(A) tail was attached to the 3' NTR of the sequence. The VPc protein was replicated using the PDB file ZBBP and covalently attached to the 5' NTR of the genome. Each RNA fragment was generated as a helical structure approximately 200nm in length. Using steered molecular dynamics the RNA fragments were compressed to 30nm and then arranged in a super-coil formation as a sphere of 18.7 nm in diameter. The RNA was then solvated and neutralized using magnesium ions. Once equilibrated the RNA was transcribed into the core of the poliovirus capsid creating a final structure totalling 1.95 million atoms. An additional 3 million atoms were required to replicate a 0.154mM sodium chloride solution with cell edge padding at 12 Å on each axis with cubic cell dimensions of 344.8 angstrom. Due to the inefficient nature of a cubic system for the simulation of an icosahedral structure, a rhombic dodecahedron structure was chosen to reduce the amount of water required to solvate the virus, thus reducing the atom count from 4 million atoms to 2.75 million atoms. The adoption of rhombic-dodecahedral periodic boundary conditions resulted in an increase in simulation efficiency of 50%. Molecular dynamics simulation was performed using supercomputers at the Victorian Life Sciences Computation Initiative (VLSICI) and the 64-bit Nanoscale Molecular Dynamics (NAMD) software package⁹ (http://www.ks.uiuc.edu/Research/namd/). Initially, 50,000 steps of energy minimisation was performed followed by equilibration for 0.5 ns to 310 degrees Kelvin at 1013.25 mbar (one atmosphere) using 256 cores of an SGI Altix XE supercomputer. Once equilibrated the virus simulation was moved to a BlueGene/P supercomputer running NAMD 2.8 beta 1 with memory optimisation extensions. The simulation was run on 1024 cores resulting in 1.4 days per ns and was stopped at the 10 ns time point. Additionally, empty capsid structures (no RNA) were simulated for a total of 10 ns. Observation of capsid structural integrity and deformation was achieved by the use of two strategies. The first involved superheating the virion to 200°C for 1.7 ns at atmospheric pressure. In parallel, a virion was cooled to 4°C for 1.7 ns using the same conditions. To replicate severe internal ionic disruption of the virion, magnesium ions were removed from the core of the virus to allow the RNA to rapidly displace and expand outward.

References

1. Freddolino PL, Arkhipov AS, Larson SB, McPherson A, Schulten K (2006) Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure* 14: 437-448.
2. McGuffee LJ, Bryson K, Jonas DT (2000) The PSIPRED protein structure prediction server. *Bioinformatics* 16: 404.
3. Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, et al. (2008) Protein structure homology modeling using SWISS-MODEL workspace. *Nature Protocols* 4: 1-13.
4. Miles ST, Hogle JM, Finlan SA (2001) A2 into phasing of high-symmetry macromolecular complexes: successful phasing of authentic poliovirus data to 3.0 Å resolution. *Journal of Molecular Biology* 307: 499-512.
5. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *Journal of Molecular Graphics* 14: 33-38.
6. Stone JE, Phillips JC, Freddolino PL, Hardy DJ, Trabuco LG, et al. (2007) Accelerating molecular modeling applications with graphics processors. *Journal of Computational Chemistry* 28: 2616-2640.
7. Zoete V, Cuendet MA, Grosdidier A, Michielin O (2011) SwissParam: A fast force field generation tool for small organic molecules. *Journal of Computational Chemistry*.
8. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, et al. (2005) Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* 26: 1781.
9. Levay, H. C., M. Boslino, D. J. Filman, and J. M. Hogle. 2010. Catching a virus in the act of RNA release: a novel poliovirus uncoating intermediate characterized by cryo-electron microscopy. *Journal of Virology* 84:4425.

Acknowledgements

We would like to thank Dallas Wilson, Thomas Aiken and Mike Catton of the Victorian Infectious Diseases Reference Laboratory. We thank Cyril Reboul, Grisha Meyer and Matthew Dowton for their kind assistance with the generation of the rhombic dodecahedral modelling conditions. We would like to acknowledge and thank nVidia™ for providing access to a Tesla C2070 GPU for the purposes of this study. This research was supported by a Victorian Life Sciences Computation Initiative (VLSICI) grant number V00069 on the Peak Computing Facility at the University of Melbourne, an initiative of the Victorian Government, Australia. The National Enterovirus and the WHO Regional Poliovirus Reference Laboratories are funded by the Australian Federal and State Government Departments of Health and by the World Health Organization.

A8.2 Publications

A8.2.1

Annual reports

Annual reports

ANNUAL REPORT OF THE AUSTRALIAN NATIONAL ENTEROVIRUS REFERENCE LABORATORY 2012

Jason Roberts, Linda Hobday, Aishah Ibrahim, Thomas Aitken and Bruce Thorley

Abstract

In 2012 no cases of poliomyelitis were reported through clinical surveillance in Australia, and poliovirus was not detected through virological surveillance. Australia conducts surveillance for cases of acute flaccid paralysis (AFP) in children less than 15 years as the main mechanism to monitor its polio-free status in accordance with World Health Organization (WHO) recommendations. Cases of AFP in children are notified to the Australian Paediatric Surveillance Unit or the Paediatric Active Enhanced Disease Surveillance System. In 2012 Australia reported 1.2 non-polio AFP cases per 100,000 children, meeting the WHO performance criterion for a sensitive system for the fifth year in a row. However the faecal specimen collection rate from AFP cases was 29%, which was well below the WHO target of 80%. Virological surveillance for poliovirus consists of two components. Firstly, the Enterovirus Reference Laboratory Network of Australia (ERLNA) reports on the typing of enteroviruses detected in or isolated from clinical specimens. Secondly, environmental surveillance is conducted at sentinel sites. These surveillance systems are co-ordinated by the National Enterovirus Reference Laboratory (NERL).

Keywords: poliovirus, acute flaccid paralysis, surveillance, enterovirus, poliomyelitis, eradication, vaccination

Introduction

Australia, along with the World Health Organization (WHO) Western Pacific Region, was declared polio-free in 2000 and has established clinical and virological surveillance schemes to monitor its polio-free status. Clinical surveillance follows the WHO recommendation of investigating cases of acute flaccid paralysis (AFP) in children less than 15 years of age. AFP cases are ascertained either by clinicians notifying the Australian Paediatric Surveillance Unit (APSU) via a monthly report card or through the Paediatric Active Enhanced Disease Surveillance System (PAEDS). PAEDS involves ward based nurses reviewing hospital records and enrolling AFP patients with the consent of a parent or guardian at four sentinel tertiary paediatric hospitals.^{1,2} The WHO recommends that two faecal specimens be collected for virological investigation at least 24

hours apart and within 14 days of the onset of paralysis from cases of AFP in order to exclude poliovirus as the causative agent. It is a requirement of the WHO polio eradication program that specimens are tested in a WHO accredited laboratory, which for Australia is the National Enterovirus Reference Laboratory (NERL), formerly the National Polio Reference Laboratory (NPRL) at the Victorian Infectious Diseases Reference Laboratory (VIDRL). The clinical and laboratory data from AFP cases in children is reviewed by the Polio Expert Panel (PEP) and reported to the WHO as evidence of Australia's continued polio-free status.

Enterovirus and environmental surveillance programs were established as virological surveillance for poliovirus to complement the clinical surveillance program focussed on AFP cases in children. Enteroviruses other than poliovirus have been associated with AFP and poliovirus infection may manifest clinically without paralysis. The Enterovirus Reference Laboratory Network of Australia (ERLNA) was established in 2009. Public diagnostic virology laboratories report their enterovirus typing results from clinical specimens to exclude poliovirus and establish the epidemiology of non-polio enteroviruses in Australia. WHO supports environmental surveillance as a sensitive means of detecting poliovirus through the testing of sewage samples. In December 2012, Egypt reported the detection of wild poliovirus type 1 in 2 sewage samples collected in Cairo.³ Genetic sequencing identified Pakistan as the source of the viruses.

The certification of India as being polio-free in January 2012 was a significant achievement for the global polio eradication program, reducing the number of endemic countries to three; Afghanistan, Nigeria and Pakistan.⁴ Furthermore, the reporting of 223 polio cases globally in 2012 represented the lowest number since the eradication program started in 1988.⁵ Nevertheless, it is important to maintain high polio vaccine coverage and sensitive surveillance systems for AFP cases in children until global eradication is achieved. As an example, China along with other countries of the Western Pacific Region, was declared polio-free in 2000. However, an outbreak in Xinjiang province in China due to a wild poliovirus type 1 importation from Pakistan caused 21 cases of polio before the country was declared polio-free once again in October 2012.⁴ A weekly situ-

ation report of polio cases worldwide is available at the WHO website <http://www.polioeradication.org/Dataandmonitoring/Poliothisweek.aspx>

This report summarises the polio surveillance program in Australia for 2012 encompassing clinical surveillance for AFP cases in children and virological surveillance for poliovirus.

Methods

AFP Surveillance

Paediatricians reviewing a patient less than 15 years of age presenting with AFP, or clinicians reviewing a patient of any age suspected of poliomyelitis, are requested to notify the NERL (telephone 03-9342 2607, email polio@mh.org.au). Paediatricians also notify the AFP case to the APSU (<http://www.apsu.org.au/>) via a monthly report card. Upon receipt of the notification, the AFP National Surveillance Co-ordinator based at VIDRL forwards a clinical questionnaire for the clinician to complete. Alternatively, AFP cases are ascertained by PAEDS nursing staff from medical records and are enrolled in the surveillance program with parental or guardian consent.

WHO classifies specimens as being adequate for virological investigation when two faecal specimens are collected more than 24 hours apart (due to intermittent virus shedding), and the specimens are collected within 14 days of the onset of paralysis (while the virus titre remains high). The faecal specimens are tested free of charge by the NERL.

The PEP, convened by the Department of Health (DoH), reviews the clinical and laboratory data for all notified cases of AFP, irrespective of whether they are an eligible or ineligible case. An eligible case is an Australian child under 15 years of age with AFP (including Guillain-Barré syndrome and transverse myelitis) or an Australian of any age with suspected polio. Ineligible cases include patients aged 15 years or older, overseas residents and cases notified in error or later determined not to be AFP.

The PEP classifies cases of AFP as:

- Poliomyelitis due to wild poliovirus, vaccine-derived poliovirus (VDPV) or vaccine associated paralytic poliomyelitis (VAPP);
- Polio compatible if there is insufficient evidence to exclude poliomyelitis;
- Non-polio AFP; or
- Non-AFP.

A follow-up questionnaire is sent to notifying clinicians if the PEP requires more information regard-

ing the AFP case before a final classification can be made. After each PEP meeting the Australian AFP case classifications are forwarded to WHO for inclusion in the global AFP surveillance data published in the Weekly Epidemiological Record (available at <http://www.who.int/wer/en/>). Ineligible cases are not reported to WHO.

The WHO AFP surveillance performance indicator for a polio non-endemic country is one case of non-polio AFP per 100,000 children aged less than 15 years. For Australia in 2012, this equated to 43 cases per year, based on the Australian Bureau of Statistics data released in December 2011. An AFP surveillance scheme that satisfies the surveillance performance indicator is deemed sufficiently sensitive to detect a wild poliovirus importation in children of that country. The WHO surveillance performance indicator for laboratory testing is that at least 80% of notified AFP cases have adequate faecal specimens collected and tested in a WHO accredited laboratory.

At the end of each calendar year a number of AFP notifications remain pending where insufficient clinical and laboratory data were made available to the PEP. The PEP classifies the remaining AFP notifications as "polio compatible-zero evidence" if a final review reveals no evidence of clustering amongst the cases.

Virus Culture

Upon receipt at the NERL, faecal specimens are treated with Minimum Essential Medium containing Hank's salts, chloroform (9.1% v/v) and foetal bovine serum (2%). The suspension is clarified and the supernatant inoculated onto a series of mammalian cell lines. Two WHO recommended cell lines are used for the isolation of poliovirus, L20B (a transgenic mouse epithelial cell line expressing the human poliovirus receptor, CD155) and RD-A (human rhabdomyosarcoma).^{6,7} Diagnostic laboratories in Australia are encouraged to refer enteroviruses of unknown serotype to the NERL for further characterisation as poliovirus infection can lead to clinical presentations without paralysis such as aseptic meningitis.

Two WHO real time reverse transcription polymerase chain reaction (RT-PCR) tests are used to determine whether a poliovirus is a wild strain, oral poliomyelitis vaccine (OPV) strain (Sabin-like) or a vaccine-derived poliovirus (VDPV), in a process known as intratypic differentiation.⁸ The NERL sequences the complete poliovirus viral protein 1 (VP1) genomic region, which contains a major neutralizing antibody binding site. The VP1 genomic sequence provides valuable biological information, including the number of mutations within a signifi-

cant region of the OPV virus strain and it enables phylogenetic analysis of wild poliovirus to rapidly determine the likely source of the virus, as utilised in the 2007 wild poliovirus importation.⁹

Enterovirus Surveillance

The ERLNA was established primarily as a means of detecting imported poliovirus amongst untyped enteroviruses from clinical specimens. The network consists of 10 public sector diagnostic virology laboratories in the Australian Capital Territory (Canberra Hospital), New South Wales (Royal Prince Alfred Hospital), Queensland (Queensland Health and Scientific Services), South Australia (Flinders Medical Centre, Institute of Medical and Veterinary Science), Tasmania (Royal Hobart Hospital), Victoria (Royal Children's Hospital, VIDRL) and Western Australia (Queen Elizabeth II Medical Centre, Princess Margaret Hospital for Children).

The NERL encourages members of the ERLNA to perform their own enterovirus typing. It has advised members of the ERLNA on enterovirus detection, supplied laboratory and computer analysis protocols and performed tests in parallel with other laboratories for quality assurance purposes. The NERL receives untyped enteroviruses from three laboratories for typing on a regular basis. The other laboratories perform their own enterovirus typing and report the results to the NERL for inclusion in the national enterovirus database.

The NERL screens clinical specimens for enteroviruses using a semi-nested RT-PCR directed to highly conserved sequence in the 5' non-translated region (NTR).¹⁰ Enterovirus typing is primarily performed by amplifying a fragment of the VP1 genomic region according to a published method,¹¹ but the complete nucleotide sequence of VP1 is required to type some enteroviruses. The enterovirus typing RT-PCR is directed to a region of sequence divergence that allows differentiation between enterovirus genomes. As a consequence, the enterovirus sequence based typing assay is not as sensitive as the pan-enterovirus detection assay. This can result in an enterovirus being detected by pan-enterovirus RT-PCR in a clinical specimen without subsequent identification by the VP1 enterovirus typing assay.

Environmental surveillance

The laboratory cell culture protocol implemented by the NERL for environmental surveillance is based on a two-phase separation procedure published by WHO and further advice was obtained from the Enterovirus Laboratory at the National Public Health Institute,¹² Finland, a Global Specialised Laboratory in the WHO Polio Laboratory Network. In brief, 800 mL of sewage is collected prior to any

biological or chemical treatment and referred to the NERL within 24 hours. At the laboratory 500 mL of the sample is centrifuged and the supernatant vigorously shaken at 4°C with dextran, polyethylene glycol and sodium chloride. The mixture is incubated overnight at 4°C in a separating funnel and the lower organic phase is collected the next day and used to re-suspend any pellet stored after the initial centrifugation. The final solution is clarified as for a faecal specimen and inoculated onto the L20B and RD-A cell lines and observed microscopically for cytopathic effect. The sewage extracts are tested in parallel by cell culture and a pan-enterovirus RT-PCR. The pan-enterovirus RT-PCR is a validated in-house test and is utilised to confirm the cell culture results as not all human enteroviruses can infect the RD-A cell line. All enterovirus isolates from cell culture and positive detections by RT-PCR were investigated to determine the virus type by nucleic acid sequencing.

Results

Classification of AFP cases

A total of 77 notifications of AFP in children less than 15 years of age were received in 2012 (Table 1). The PEP classified 51 cases as non-polio AFP with onset of paralysis in 2012. This equated to a non-polio AFP rate of 1.2 cases per 100,000 children less than 15 years of age, exceeding the WHO AFP surveillance performance criterion for a polio-free country of one case of non-polio AFP per 100,000 children (Table 2, Figure 1).

In 2012, one AFP case reviewed by the PEP had a differential diagnosis of Guillain-Barré syndrome and anterior horn cell disease due to an enteroviral infection. One faecal specimen, collected eight days after the onset of paralysis, was reported as no enterovirus isolated by cell culture by the NERL. The patient had not travelled overseas in the three months preceding the onset of symptoms. The PEP was not able to exclude polio based on the available clinical evidence and classified the case as "polio compatible" (Table 1). No further clinical information or laboratory specimens were received from one other AFP notification and the PEP classified the case as "polio compatible – zero evidence" to indicate the fact that it was a notification only with no further evidence to support a clinical diagnosis of polio.

Thirteen AFP cases were notified by more than one clinician and were regarded as duplicate notifications (Table 1). Eight AFP notifications did not meet the criteria for an eligible case. These involved either patients greater than 14 years of age, cases with symptom onset prior to 2012, or cases that were later reported as non-AFP. Three cases involving patients older than 14 years of age were all classified

by the PEP as non-polio AFP. However, they were not reported to the WHO as the global polio surveillance program focuses on AFP in children less than 15 years of age as the age group being at high risk of poliovirus infection.

Notification of AFP cases by state and territory

In 2012, eligible AFP cases were notified from all jurisdictions in Australia except the Australian Capital Territory and Tasmania (Table 1). The non-polio AFP rates for eligible cases per jurisdiction exceeded the WHO AFP surveillance performance indicator of one case per 100,000 children in New South Wales, Northern Territory, South Australia,

Victoria and Western Australia. Queensland was the only more populous state not to have achieved the WHO surveillance performance indicator.

Faecal collection from AFP cases

In 2012, a total of 64 faecal specimens from 37 of the 51 eligible cases were tested at the NERL (Table 3). No poliovirus was isolated from any of the specimens. The non-polio enteroviruses, coxsackievirus A7, coxsackievirus A16, coxsackievirus B5 and enterovirus 71 subgenogroup C2 were reported from four AFP cases in 2012. Diagnoses were transverse myelitis for the first two cases, Guillain-Barré syndrome for the third and acute disseminated encephalomyelitis for the last. Fifteen (29%) of the eligible cases had adequate specimens collected in 2012, while another 12 (24%) cases had only one specimen collected within the optimal period. This compares to the further WHO AFP surveillance criterion that 80% of the eligible AFP cases should have adequate specimens collected, a result that Australia has never achieved nationally (Figure 2). At the jurisdictional level, Queensland was the only state to reach the WHO target in 2012, with adequate specimens collected from all five cases classified (100%).

Enterovirus and environmental surveillance

No poliovirus was detected by enterovirus or environmental surveillance in 2012. The ERLNA typed 277 non-polio enteroviruses with coxsackievirus A6 and echovirus 18 amongst the most frequent detections in Australia during the year (Table 4).

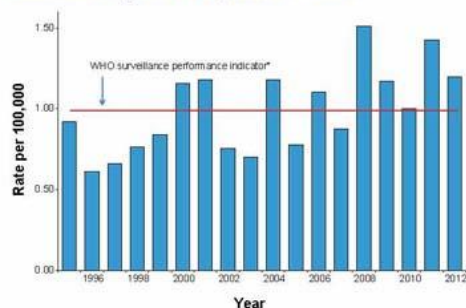
Four collections from each of the three sentinel sites (Armidale, Byron Bay and Newcastle) were tested by RT-PCR, and virus isolation. Twelve collections (four from each site) were tested by cell culture and RT-PCR. All 12 samples were positive by pan-enterovirus RT-PCR and non-polio enterovirus was isolated in cell culture from eight samples. Four samples positive by pan-enterovirus RT-PCR could not be typed due to low virus titres.

Regional reference laboratory activities

Several activities were performed as a Polio Regional Reference Laboratory in 2012. Specimens from AFP cases were referred from Brunei Darussalam (2 cases), Pacific Island countries (6 cases) and Papua New Guinea (9 cases). No poliovirus was isolated from any of the specimens but non-polio enteroviruses were reported from one case from the Pacific Islands and 5 cases from Papua New Guinea.

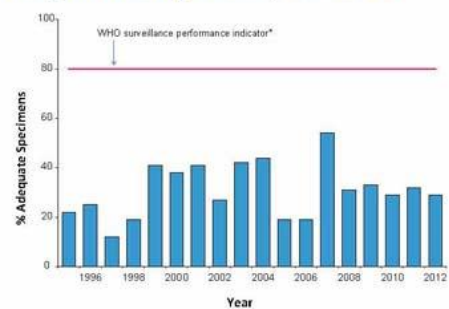
Six poliovirus type 2 and eight poliovirus type 3 isolates were referred from AFP cases in the Philippines for intratypic differentiation and all were character-

Figure 1: Non-polio AFP rate classified by the PEP, 1995 to 2012



* The WHO AFP surveillance performance indicator for a polio non-endemic country is one case per 100,000 children <15 years of age.

Figure 2: Percentage of AFP cases with adequate faecal specimens, 1995 to 2012



* The criterion for the WHO surveillance performance indicator is the collection of two faecal specimens more than 24 hours apart and within 14 days of the onset of paralysis from 80% of classified non-polio AFP cases.

Table 1: Notification of AFP cases in Australia, 2012, by state or territory

State/Territory	Estimated population aged <15 years	Expected number of AFP cases [†]	Total number of AFP cases	Ineligible notifications	Duplicate notifications	Polio compatible- zero evidence	Polio compatible	Pending	Eligible cases with final classification by PEP	Non-polio AFP rate per 100,000 children
ACT	67,397	0.5	0	0	0	0	0	0	0	0.0
NSW	1,358,279	14	21	1	2	0	1	1	16	1.1
NT	52,749	0.5	2	0	0	0	0	0	2	4.0
QLD	909,482	9	14	2	4	1	0	2	5	0.6
SA	293,392	3	5	0	0	0	0	0	5	1.7
TAS	97,694	1	1	1	0	0	0	0	0	0.0
VIC	1,027,417	10	28	4	6	0	0	0	18	1.8
WA	453,747	5	6	0	1	0	0	0	5	1.0
Australia	4,260,157	43	77	8	13	1	1	3	51	1.2

* Australian Bureau of Statistics, estimated population at 30 June 2011. Available at <http://www.abs.gov.au/>

† Based on a non-polio AFP rate of 1 case per 100,000 children less than 15 years of age

Table 2: Surveillance for AFP cases in children less than 15 years, Australia, 2012, compared with the WHO performance indicators

WHO surveillance performance indicator for AFP cases in children <15 years	Performance of Australia's AFP surveillance	
	Number of cases/specimens 2012	Comparison with WHO Indicator 2012
AFP cases		
≥1.0 non-polio AFP case / 100,000 children (43 cases for Australia in 2012).	51 cases classified as non-polio AFP	1.2 (51 / 43) non-polio AFP cases / 100,000 children <15 years
Adequate specimen collection		
≥80% of classified AFP cases with adequate specimens* (41 cases for Australia in 2012).	15 AFP cases with adequate specimens collected	29% (15 / 51) classified non-polio AFP cases with adequate specimens

* Adequate specimen collection is defined as 2 faecal specimens collected at least 24 hours apart and within 14 days of onset of paralysis

ised as being Sabin-like. Seven poliovirus type 1 and two poliovirus type 2 isolates were characterised as Sabin-like from sources other than AFP.

Quality Assurance Programs

In 2012, the NERL passed the annual WHO quality assurance panels for poliovirus isolation by cell culture and poliovirus RT-PCR for intratypic differentiation and vaccine derived poliovirus. The WHO distributed the first official poliovirus

sequencing proficiency panel and the laboratory scored full marks for sequencing RNA templates consisting of wild, Sabin and Sabin prototype mixtures of poliovirus. The laboratory also participated in the Royal College of Pathologists of Australasia quality assurance panel for enterovirus detection by RT-PCR.

Table 3: Specimens referred to the NERL Australia, 2012

Result	Specimens from AFP cases in patients < 15 years of age	Specimens from AFP cases in patients ≥15 years of age	Specimens from sources other than AFP	TOTAL
Non-polio enterovirus	5	0	150	155
Rhinovirus	1	0	2	3
No enterovirus identified	58	4	35	97
Total	64	4	187	255

Table 4: Enterovirus test results from the NERL Australia, 1995 to 2012

Year	Poliovirus		Non-polio enterovirus	No enterovirus detected	EVID results referred†	Total samples reviewed
	Sabin-like	Non-Sabin-like*				
1995	190	0	200	13	0	403
1996	224	0	198	9	0	431
1997	124	0	76	0	0	200
1998	52	0	15	4	0	71
1999	60	1	9	9	0	79
2000	45	0	44	47	0	136
2001	46	5	33	75	0	159
2002	36	0	21	49	0	106
2003	9	0	15	47	0	71
2004	6	0	26	61	0	93
2005	18	0	10	39	0	67
2006	2	0	6	71	29	108
2007	0	2	32	115	107	256
2008	0	0	20	92	77	189
2009	1	0	63	78	113	255
2010	0	0	170	39	108	317
2011	0	0	174	61	205	440
2012	0	0	155	97	123	375

* Untyped enterovirus or uncharacterised poliovirus isolates were referred for further testing after completion of a laboratory inventory in 1999 and 2001; the six non-Sabin-like isolates were identified as wild type poliovirus prototype strains and destroyed. Wild poliovirus type 1 was imported from Pakistan in 2007. A Sabin-like poliovirus type 1 was identified from an unimmunised infant in 2009.

† Enterovirus Identification (EVID) results include retrospective data made available via the ERNLA.

Discussion

In 2012, Australia reached the WHO surveillance target of ≥ 1 non-polio AFP case per 100,000 children, for the fifth year in a row. The continued participation of clinicians and health care workers in notifying cases of AFP to the APSU and VIDRL along with the involvement of the ward based nurses in the PAEDS is essential in reaching this target, indicative of a sensitive surveillance system. Collection of adequate faecal specimens from AFP cases in Australia has never met the WHO surveillance target and represents a gap in clinical surveillance for imported cases of polio. The establishment of supplementary virological surveillance for poliovirus by the typing of enteroviruses and testing of environmental samples at sentinel sites provides an additional means of monitoring Australia's polio-free status.

Since the initial target for global polio eradication by the year 2000 set by the World Health Assembly (WHA) in 1988, a number of subsequent target years set for achieving polio eradication have not been met, the most recent being the end of 2012.¹³ However significant achievements have been attained since 1988, including the last reported case of wild poliovirus type 2 in 1999 and a reduction in the number of polio endemic countries worldwide from 125 to 3 by 2012. However, cases of wild poliovirus have reviewed around 1,200 annually between 2002 and 2010 (Figure 3).¹⁴ After Egypt was certified polio-free in 2006, it has proved difficult to eradicate the virus from the remaining areas of wild poliovirus transmission (Afghanistan, India, Nigeria and Pakistan).

A new strategy was initiated from 2005 with the introduction of monovalent oral polio vaccine for poliovirus type 1 and poliovirus type 3 followed by bivalent oral polio vaccine for poliovirus types 1 and 3 in 2009.¹⁵ The judicious use of trivalent, bivalent

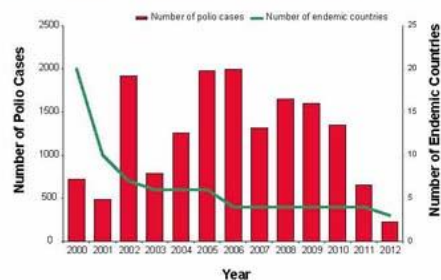
and monovalent oral polio vaccines reduced the number of polio cases in the polio endemic countries and those with re-established transmission to 217 and 6, respectively, in 2012.⁵ Changes to the WHO laboratory testing protocols were also introduced from 2006 to shorten the time taken to confirm cases of polio.¹⁶ The WHO Global Polio Laboratory Network introduced a new cell culture algorithm that halved the reporting time from 28 days to 14 days and implemented real time RT-PCR protocols that reduced the timeframe for poliovirus intratypic differentiation from 14 days to 7 days.

A further response in 2010 was the establishment of an Independent Monitoring Board (IMB) by the WHA to monitor and guide the progress of the Global Polio Eradication Initiative's 2010-2012 Strategic Plan.¹⁷ The IMB was convened quarterly to review developments in the polio program and provided independent advice regarding the requirements for the plan to succeed. The 2010 to 2012 strategic plan aimed to stop wild poliovirus transmission in two of the four endemic countries by the end of 2011, with only one country, India, achieving that goal.⁴ Considering that India reported 741 cases of polio as recently as 2009, representing 46% of the cases worldwide, its certification as polio-free in January 2012 was a significant milestone for the eradication program. Despite this, the January 2012 IMB report concluded that global polio eradication would not be achieved if it continued on its current path.¹⁷ One of the board's conclusions was to place a greater emphasis on people management, including rating the importance of having well-trained vaccinators who are valued and inspired as the most important group in the programme. The cessation of polio vaccination in response to the deliberate killing of polio vaccinators in Pakistan in late 2012 reinforced the board's opinion of their key role.¹³

In January 2012, in response to a recommendation by the IMB, the WHO Executive Board called "the completion of poliovirus eradication a programmatic emergency for global public health". This was adopted as a resolution by the WHA the following May.¹⁸ By the end of 2012, 223 cases of polio were reported for the year, the lowest annual total ever (Figure 3). Transmission of wild poliovirus was restricted to four countries, the lowest since the program began; the three endemic countries of Afghanistan, Nigeria and Pakistan, and Chad with re-established transmission.⁵ Furthermore, wild poliovirus type 3 was last reported in Afghanistan and Pakistan in April 2012, potentially leaving Nigeria as the last country to be endemic for this serotype.

The Polio Eradication Initiative appears to be at a critical juncture. A concerted international effort is required to support the final stages of wild

Figure 3: Cases of wild polio virus infection in endemic countries, 2000 to 2012, by year



Data from: <http://www.polioeradication.org/Dataandmonitoring/Poliothisweek/Wildpolioviruslist.aspx>, accessed 27 March 2013.

poliovirus eradication to avoid a repeat of 2001. This was the last time that relatively few cases were reported, and was followed in subsequent years by a rapid rise in the number of cases in the endemic countries and frequent occurrences of re-established transmission in others. To restrict the international spread of wild poliovirus, the IMB recommended in its seventh report that by May 2013 the International Health Regulations Expert Review Committee issue a standing recommendation that travellers from the remaining endemic countries receive pre-travel vaccination or a check of their vaccination status until polio transmission in the country ceases.¹⁷

Until certification of global wild poliovirus eradication, Australia remains at risk of an importation as occurred from Pakistan in 2007.⁹ The continued performance of the clinical and virological surveillance systems for poliovirus to a high international standard is essential to monitor Australia's polio-free status in order to detect and rapidly respond to any future polio importation event.

Acknowledgements

The authors thank the clinicians who participated in the AFP surveillance program in 2012 as well as the APSU and the PAEDS team. The active involvement of the laboratory members of the ERLNA is gratefully acknowledged, as are the staff of the water authorities involved with environmental surveillance. The polio surveillance program co-ordinated by the NERL is funded by the Australian Government Department of Health, the Victorian government Department of Health and VIDRL.

Author details

Mr Jason Roberts, Senior Medical Scientist¹
Ms Linda Hobday, Medical Scientist¹
Mrs Aishah Ibrahim, Medical Scientist¹
Mr Thomas Aitken
Dr Bruce Thorley, Senior Medical Scientist, Laboratory Head¹

1. National Enterovirus Reference Laboratory, Victorian Infectious Diseases Reference Laboratory, North Melbourne, Victoria, Australia.

Corresponding author: Dr Bruce Thorley, Senior Medical Scientist, Laboratory Head, National Enterovirus Reference Laboratory, WHO Polio Regional Reference Laboratory, Victorian Infectious Diseases Reference Laboratory (VIDRL), Telephone: +61 3 9342 2607; Facsimile: +61 3 9342 2665; E-mail: bruce.thorley@mh.org.au

References

1. Study Protocol, Acute Flaccid Paralysis. Australian Paediatric Surveillance Unit [Internet]. Accessed on 27 March 2013. Available from: <http://www.apsu.org.au/assets/current-studies/AFP-Study-Protocol-APSU-Final-110810.pdf>.
2. Paediatric Active Enhanced Disease Surveillance [Internet]. Accessed on 27 March 2013. Available from: <http://www.apsu.org.au/surveillance-systems/paedcs/>.

3. Outbreak news – Poliovirus isolation, Egypt. *Weekly Epidemiological Record*. 2013;88:74-75.
4. World Health Organization. Progress towards global interruption of wild poliovirus transmission, January 2011–March 2012. *Weekly Epidemiological Record*. 2012;87:189-200.
5. Global Polio Eradication Initiative [Internet]. Wild poliovirus 2008-2013. Accessed on 27 March 2013. Available from: http://www.polioeradication.org/Portals/0/Document/Data&Monitoring/Wild_poliovirus_list_2008_2013_26Mar.pdf
6. Wood DJ, Hull B. L20B Cells Simplify Culture of Polioviruses from Clinical Samples. *J Med Virol* 1999;58:188-192.
7. World Health Organization. *Polio Laboratory Manual, 4th edition*, WHO/IVB/04.10. Geneva: Department of Immunization, Vaccines and Biologicals; 2004.
8. Kilpatrick DR, Yang CF, Ching K, Vincent A, Iber J, Campagnoli R, Mandelbaum M, De L, Yang SJ, Nix A. Rapid group-, serotype-, and vaccine strain-specific identification of poliovirus isolates by real-time reverse transcription PCR using degenerate primers and probes containing deoxycytosine residues. *J Clin Microbiol* 2009;47:1939-1941.
9. Stewardson AJ, Roberts JA, Beckett CL, Prime HT, Loh PS, Thorley BR, Daffy JR. An Imported Case of Poliomyelitis in Melbourne, Australia. *Emerging Infect. Dis.* 2009;15:63-65.
10. Roberts JA, Thorley BR. Enterovirus. In: Schuller M, Sloots TP, James GS, Halliday CL, Carter IWJ, editors. *PCR for Clinical Microbiology: An Australian and International Perspective*. 1st edn. Dordrecht: Springer; 2010. p. 229-233.
11. Nix WA, Oberste MS, Pallansch MA. Sensitive, seminested PCR amplification of VP1 sequences for direct identification of all enterovirus serotypes from original clinical specimens. *J Clin Microbiol* 2006;44:2698-2704.
12. World Health Organization. Guidelines for environmental surveillance of poliovirus circulation. Geneva: Department of Vaccines and Biologicals WHO/V&B/03.03; 2003.
13. Global polio eradication: not there yet [editorial]. *Lancet* 2013;381:1.
14. Global Polio Eradication Initiative [Internet]. Wild poliovirus list. Accessed on 27 March 2013. Available from: <http://www.polioeradication.org/Dataandmonitoring/Poliothisweek/Wildpolioviruslist.aspx>.
15. Crawford NW, Buttery JP. Poliomyelitis eradication: another step forward. *The Lancet* 2010;376:1624-1625.
16. World Health Organization. Informal consultation of the Global Polio Laboratory Network-June 2008. *Weekly Epidemiological Record* 2008;83:261-268.
17. Independent Monitoring Board [Internet]. IMB meeting reports. Accessed on 27 March 2013. Available from: <http://www.polioeradication.org/Aboutus/Governance/IndependentMonitoringBoard/Reports.aspx>
18. Global Polio Eradication Initiative [Internet]. Agenda item 13.10. Poliomyelitis: intensification of the global eradication initiative. WHA65.5. May 2012. Accessed 27 March 2013. Available from <http://www.polioeradication.org/Resourcelibrary/Declarationsandresolutions.asp>

ANNUAL REPORT OF THE AUSTRALIAN NATIONAL ENTEROVIRUS REFERENCE LABORATORY 2010-2011

Jason Roberts, Linda Hobday, Aishah Ibrahim, Thomas Aitken and Bruce Thorley

Abstract

Australia conducts clinical surveillance for cases of polio-like illness in children in accordance with the World Health Organization (WHO) recommended surveillance criteria for acute flaccid paralysis (AFP). AFP cases are ascertained either by clinicians notifying the Australian Paediatric Surveillance Unit or designated nurses enrolling cases as part of the Paediatric Active Enhanced Disease Surveillance system at four sentinel tertiary paediatric hospitals. The National Enterovirus Reference Laboratory (NERL), formerly the National Poliovirus Reference Laboratory, is accredited by the World Health Organization (WHO) for the testing of faecal specimens from cases of AFP and operates as a Poliovirus Regional Reference Laboratory for the Western Pacific Region. In 2010 and 2011, for the 3rd and 4th consecutive years, Australia met the WHO AFP surveillance performance indicator. This is indicative of a sensitive surveillance system capable of detecting an imported case of polio in children. However, the faecal collection rate for the virological investigation of AFP cases was below the WHO surveillance performance indicator in both years and represented a gap in Australia's polio surveillance. Enterovirus and environmental surveillance were established in Australia as virological surveillance to complement the clinical surveillance schemes. No poliovirus was detected by the clinical or virological surveillance schemes in 2010 or 2011 and Australia maintained its polio-free status. India was declared polio-free in January 2012, a significant step towards global polio eradication, leaving Afghanistan, Nigeria and Pakistan as the remaining countries endemic for wild poliovirus.

Keywords: poliovirus, acute flaccid paralysis, surveillance, enterovirus, poliomyelitis, eradication, vaccination

Introduction

The National Enterovirus Reference Laboratory (NERL), formerly the National Polio Reference Laboratory, is responsible for the virological testing of faecal specimens from cases with a clinical suspicion of poliomyelitis. This includes cases of acute flaccid paralysis (AFP), a major clinical presentation of poliomyelitis, in children less than 15 years of age and cases of suspected poliomyelitis in patients of any age. The World Health Organization (WHO) recommends that two faecal specimens be collected

for virological investigation at least 24 hours apart and within 14 days of the onset of paralysis from cases of AFP to exclude poliovirus as the causative agent. It is a requirement of the WHO polio eradication program that the specimens are tested in a WHO accredited laboratory, which for Australia is the NERL at the Victorian Infectious Diseases Reference Laboratory (VIDRL).

Enterovirus and environmental surveillance programs were established to provide virological surveillance for poliovirus to complement the clinical surveillance program focussed on AFP cases in children. Enteroviruses other than poliovirus have been associated with AFP and poliovirus infection may manifest clinically without paralysis. The Enterovirus Reference Laboratory Network of Australia (ERLNA) was established in 2009, bringing together public diagnostic virology laboratories. ERLNA aims to identify the types of enteroviruses detected in clinical specimens to exclude poliovirus and establish the epidemiology of non-polio enteroviruses in Australia. WHO supports environmental surveillance as another aspect of polio surveillance through the testing of sewage samples. Clinical and virological surveillance schemes for poliovirus serve to monitor Australia's polio-free status.

From November 2005, inactivated poliomyelitis vaccine (IPV) replaced oral poliomyelitis vaccine (OPV) in the National Immunisation Program.¹ IPV is administered to children at 2, 4 and 6 months of age, with a booster dose at 4 years of age. With the removal of OPV (containing live attenuated virus) from the immunisation schedule, any poliovirus identified by Australian virology laboratories requires further investigation to determine its origin, as it potentially represents an importation event.

It is important that Australia maintains high levels of polio vaccine coverage to avoid a resurgence of poliomyelitis in the event of a wild poliovirus importation. In 2010, China reported an outbreak of polio due to wild type virus in Xinjiang province, which borders Pakistan and from where the importation originated. A total of 21 cases of polio were reported ranging in age from 6 months to 42 years before the outbreak was controlled through mass immunization programs.² The WHO provides a weekly update of the global polio eradication situation including a list of countries reporting cases due to wild poliovirus (<http://www.polioeradication.org/Dataandmonitoring/Poliothisweek.aspx>).

The Australian Immunisation Handbook recommends that individuals who are at continuing risk of infection, such as health care workers should have a polio vaccine booster every 10 years.¹

The last wild poliovirus isolated in India was in January 2011. The country was declared polio-free 12 months later reducing the number of countries that have never interrupted wild poliovirus transmission to three; Afghanistan, Nigeria and Pakistan.² It was only in 2009 that India reported 741 cases of polio due to wild type poliovirus which accounted for 46% of the cases worldwide. India's polio-free status is a significant public health achievement that supports the feasibility of the global polio eradication strategy.

This report summarises the polio surveillance program in Australia for 2010 and 2011, encompassing AFP surveillance in children and virological surveillance.

Methods

AFP Surveillance

AFP surveillance was initiated by the Australian Government in 1995 in collaboration with the Australian Paediatric Surveillance Unit (APSU) as part of Australia's commitment to the WHO poliomyelitis eradication program. Since 2000, AFP surveillance has been co-ordinated by VIDRL in collaboration with the APSU. In late 2007, the Paediatric Active Enhanced Disease Surveillance (PAEDS) surveillance scheme was established as a collaboration between the APSU and the National Centre of Immunisation Research and Surveillance. PAEDS is a hospital based surveillance system for paediatric conditions of public health interest, including AFP, at four tertiary paediatric hospitals in Adelaide, Melbourne, Perth and Sydney.³ In April 2011 the Polo Expert Committee which is responsible for reviewing AFP cases to determine if they are compatible with polio was renamed the Polio Expert Panel (PEP) by the Communicable Disease Network Australia (CDNA).

The strategy adopted for AFP surveillance is as follows:

- Paediatricians reviewing a patient less than 15 years of age who presents with AFP, or clinicians reviewing a patient of any age suspected of poliomyelitis, are requested to notify the NERL (telephone 03 9342 2607, email polio@mh.org.au). Notification of the AFP case is also included on the paediatrician's monthly report card to the APSU (<http://www.apsu.org.au/>). Upon receipt of the notification the AFP National Surveillance Co-ordinator, based at VIDRL, forwards a clinical questionnaire for the clinician to complete.

- Alternatively, AFP cases are ascertained by PAEDS nursing staff from medical records and with parental agreement are enrolled in the surveillance system.

- Two faecal specimens are collected 24 to 48 hours apart and within 14 days of onset of paralysis. The collection of specimens within these time frames enables them to be classified as adequate by WHO.

- The faecal specimens are tested free of charge by the NERL, which is accredited by WHO for this purpose.

- The PEP which is convened by the Department of Health (DoH), reviews the clinical and laboratory data for all notified cases of AFP, irrespective of whether they are an eligible or ineligible case. An eligible case is: an Australian child under 15 years of age with AFP (including Guillain-Barré syndrome) or an Australian of any age with paralytic illness if polio is suspected. Examples of ineligible cases are where the patient is aged 15 years or older, an overseas resident and cases notified in error or later determined to be non-AFP. The PEP classifies cases of AFP as:

- Poliomyelitis due to wild poliovirus, vaccine-derived poliovirus (VDPV) or vaccine associated paralytic poliomyelitis (VAPP)
- Non-polio AFP or
- Non-AFP

A follow-up questionnaire is sent to notifying clinicians if the PEP requires more information regarding the AFP case before a final classification can be made.

After each PEP meeting the Australian AFP data is forwarded to WHO for inclusion in the global AFP surveillance data published in the Weekly Epidemiological Record (available at <http://www.who.int/wer/en/>). Ineligible cases are not reported to WHO.

The WHO AFP surveillance performance indicator for a polio non-endemic country is 1 case of non-polio AFP per 100,000 children aged less than 15 years each year. For Australia in 2009 this equated to 41 cases, based on the Australian Bureau of Statistics (ABS) population data released in December 2008. An AFP surveillance scheme that satisfies the surveillance performance indicator is deemed sufficiently sensitive to detect a wild poliovirus importation in children of that country.

The WHO surveillance performance indicator for laboratory testing is that at least 80% of notified AFP cases have adequate faecal specimens collected and tested in a WHO accredited laboratory.

At the end of each calendar year, a number of AFP notifications remain unclassified if insufficient clinical and laboratory data were available to enable the PEP to review the cases. The PEP classifies the remaining AFP notifications as "polio compatible-zero evidence" if a final review reveals no evidence of clustering amongst the unclassified cases.

Virus Culture

Upon receipt at the NERL, faecal specimens are treated with Minimum Essential Medium containing Hank's salts, chloroform (9.1% v/v) and foetal bovine serum (2%). The suspension is clarified and the supernatant inoculated onto a series of mammalian cell lines. Two WHO recommended cell lines are used for the isolation of poliovirus; L20B (a transgenic mouse epithelial cell line expressing the human poliovirus receptor, CD155) and RD-A (human rhabdomyosarcoma).^{4,5} Up to September 2011 the NERL utilised two additional cell lines for the isolation of poliovirus and non-polio enteroviruses: Buffalo Green Monkey Kidney (BGMK) and human embryonic lung (HEL).

Diagnostic laboratories in Australia are encouraged to refer enteroviruses of unknown serotype to the NERL for further characterisation as poliovirus infection can lead to clinical presentations without paralysis such as aseptic meningitis.

A series of tests known as intratypic differentiation (ITD) are performed on poliovirus isolates to determine whether the virus is a wild poliovirus strain, OPV strain (Sabin-like) or a VDPV. In 2009 WHO introduced diagnostic poliovirus real time reverse transcriptase PCR (rRT-PCR) developed by the Centers for Disease Control and Prevention USA, as the primary ITD method.⁶ The Australian NERL sequences the complete poliovirus VP1 genomic region, which contains a major neutralizing antibody binding site. The VP1 genomic sequence provides valuable biological information, including the number of mutations within a significant region of the OPV virus strain and enables the phylogenetic analysis of wild poliovirus to rapidly determine the likely source of the virus, as utilised in the 2007 importation.^{7,8}

Enterovirus Surveillance

The ERLNA was established primarily as a means of detecting imported poliovirus amongst untyped enteroviruses from clinical specimens. The network consists of 10 public sector diagnostic virology laboratories:

Australian Capital Territory

Canberra Hospital: Prof. Peter Collignon, Dr Karina Kennedy, Ms Jennifer Ridgway

New South Wales

Infectious Diseases and Immunology, the University of Sydney: Prof. Peter McMinn

Queensland

Queensland Health and Scientific Services: Dr Russell Simmons, Dr Bruce Harrower, Dr David Warrilow

South Australia

Microbiology and Infectious Diseases, Flinders Medical Centre: Prof. David Gordon

SA Pathology, Institute of Medical and Veterinary Science: Dr Tuck Weng, Ms Kok, Ms Lyn Payne

Victoria

Department of Microbiology, Royal Children's Hospital: Dr Andrew Daley, Ms Poppy Adamopoulos

National Enterovirus Reference Laboratory, VIDRL: Dr Bruce Thorley, Mr Jason Roberts

Viral Identification Laboratory, VIDRL: Dr Chris Birch, Ms Gina Papadakis

Western Australia

Department of Clinical Microbiology, Sir Charles Gairdner Hospital: Dr Avram Levy, Dr Simon Williams, Dr David Williams, Dr David Speers

Department of Microbiology, Princess Margaret Hospital for Children: Dr Leanne Sammels, Ms Katie Lindsay, Prof. Tony Keil

The NERL encourages members of the ERLNA to perform their own virus typing. It has advised members of the ERLNA on enterovirus detection, supplied laboratory and computer analysis protocols and performed tests in parallel with other laboratories for quality assurance purposes. The NERL receives untyped enteroviruses from three laboratories for typing on a regular basis. The other laboratories perform their own enterovirus typing and report the results to the NERL for inclusion in the National Enterovirus Database.

The NERL screens clinical specimens for enterovirus using a semi-nested RT-PCR directed to highly conserved sequence in the five non-translated region (NTR).⁹ Enterovirus typing is

primarily performed by amplifying a fragment of the VP1 genomic region according to a published method,¹⁰ but the complete nucleotide sequence of VP1 is required to type some enteroviruses. The enterovirus typing RT-PCR is directed to a region of sequence divergence that allows differentiation between enterovirus genomes. As a consequence, the enterovirus sequence based typing assay is not as sensitive as the pan-enterovirus detection assay. This can result in an enterovirus being detected by pan-enterovirus RT-PCR in a clinical specimen without subsequent identification by the VP1 enterovirus typing assay.

Environmental surveillance

The laboratory cell culture protocol implemented by the NERL is based on a two-phase separation procedure published by WHO. Further advice was obtained from the Enterovirus Laboratory at the National Public Health Institute, Finland, a Global Specialised Laboratory in the WHO Polio Laboratory Network.¹¹ In brief, 800 mL of sewage is collected prior to any biological or chemical treatment and referred to the NERL within 24 hours. At the laboratory, 500 mL of the sample is centrifuged and the supernatant vigorously shaken at 4°C with dextran, polyethylene glycol and sodium chloride. The mixture is incubated overnight at 4°C in a separating funnel and the lower organic phase is collected the next day and used to re-suspend any pellet stored after the initial centrifugation. The final solution is clarified as for a faecal specimen and inoculated onto the L20B and RD-A cell lines and observed microscopically for cytopathic effect.

Results

Classification of AFP cases

A total of 57 notifications of AFP were received in 2010 (Table 1) and 78 notifications in 2011 (Table 2). The PEP classified 41 cases as non-polio AFP involving children less than 15 years of age with onset of paralysis in 2010, and 60 cases in 2011 (Tables 1 and 2). This equates to a non-polio AFP rate of 1.0 case per 100,000 children less than 15 years of age in 2010 and a rate of 1.4 per 100,000 in 2011. Thus, the WHO AFP surveillance performance criterion for a polio-free country of one case of non-polio AFP per 100,000 children less than 15 years of age was met in both years (Table 3).

In 2009, the PEP resolved to follow a WHO recommendation and report AFP notifications that could not be classified due to a lack of clinical information as "polio compatible – zero evidence". In 2010-11, a total of three AFP notifications, one each from New South Wales, Queensland and Victoria were reported to WHO as polio compatible-zero evidence (Tables 1 and 2).

Four AFP cases were notified by more than one clinician in 2010 and were regarded as duplicate notifications while nine cases were duplicated in 2011 (Tables 1 and 2). Ten AFP notifications in 2010 and eight cases in 2011, did not meet the criteria for an eligible case, either involving patients greater than 14 years of age or the cases were later reported as non-AFP. The cases involving patients greater than 14 years of age were all classified by the PEP as non-polio AFP but were not reported to the WHO as the global polio surveillance program focuses on AFP in children less than 15 years of age as an age group at high risk of poliovirus infection.

Notification of AFP cases by state and territory

In 2010 and 2011, AFP cases were reported from all jurisdictions in Australia except the Australian Capital Territory (Tables 1 and 2). After excluding duplicate notifications and ineligible cases, the non-polio AFP rates per jurisdiction exceeded the WHO AFP surveillance performance indicator of 1.0 case per 100,000 children in New South Wales, South Australia and Victoria in both years. Western Australia did not reach the surveillance indicator in 2010, with a non-polio AFP rate of 0.5 per 100,000, but it was well exceeded in Western Australia in 2011, with a rate of 2.0 per 100,000. The increase in AFP cases notified in Western Australia in 2011 may indicate a surveillance failure in 2010, and/or the year-to-year variation in incidence of a rare childhood condition in a relatively small population. This is further demonstrated in Tasmania, which did not notify any AFP cases in 2010 but reported a non-polio AFP rate of 2.0 per 100,000 in 2011 (Tables 1 and 2).

Faecal collection from AFP cases

WHO defines adequate specimens for poliovirus culture as being two faecal specimens collected at least 24 hours apart and within 14 days of the onset of paralysis. A further surveillance criterion set by WHO is for adequate faecal collection from 80% of the eligible AFP cases.

In 2010, a total of 54 faecal specimens from 29 of the 41 eligible cases were tested at the NERL. Twelve (29%) of the eligible cases had adequate specimens collected while another 12 (29%) cases had only one specimen collected within 14 days of onset.

In 2011, a total of 69 faecal specimens were received from 36 of the 60 eligible cases. Nineteen (32%) of the non-polio AFP cases had adequate specimens collected, and a further 13 (22%) cases had one specimen collected within 14 days of onset.

Table 1: Notification of AFP cases 2010, by state or territory

State/Territory	Estimated population aged <15 years	Expected number of AFP cases in 2010†	Total number of notifications	Ineligible notifications	Duplicate notifications	Polio compatible-zero evidence	Eligible cases with final classification by PEP	Non-polio AFP rate per 100,000 children
ACT	64,981	1	0	0	0	0	0	0
NSW	1,343,184	13	23	3	4	1	15	1.2
NT	52,857	1	0	0	0	0	0	0
QLD	886,584	8	6	0	0	0	6	0.8
SA	291,569	3	4	0	0	0	4	1.3
TAS	97,579	1	0	0	0	0	0	0
VIC	1,008,841	10	20	5	0	1	14	1.4
WA	438,532	4	4	2	0	0	2	0.5
Australia	4,184,127	41	57	10	4	2	41	1.0

* Australian Bureau of Statistics, estimated population at 30 June 2009. Available at www.abs.gov.au.

† Based on a non-polio AFP rate of 1 case per 100,000 children less than 15 years of age

PEP Polio Expert Panel

AFP Acute flaccid paralysis

Table 2: Notification of AFP cases 2011, by state or territory

State/Territory	Estimated population aged <15 years	Expected number of AFP cases in 2011†	Total number of notifications	Ineligible notifications	Duplicate notifications	Polio compatible-zero evidence	Cases classified by the PEP as non-polio AFP	Non-polio AFP rate per 100,000 children
ACT	66,077	1	0	0	0	0	0	0
NSW	1,355,128	13	24	2	2	0	20	1.5
NT	53,079	1	2	1	0	0	1	1
QLD	901,689	9	5	0	0	1	4	0.4
SA	293,041	3	4	0	0	0	4	1.3
TAS	97,626	1	2	0	0	0	2	2
VIC	1,017,432	10	30	4	5	0	21	2.1
WA	446,058	4	11	1	2	0	8	2
Australia	4,230,130	42	78	8	9	1	60	1.4

* Australian Bureau of Statistics, estimated population at 30 June 2010. Available at www.abs.gov.au.

† Based on a non-polio AFP rate of 1 case per 100,000 children less than 15 years of age

PEP Polio Expert Panel

AFP Acute flaccid paralysis

Australia has never achieved the WHO criterion of collection of adequate specimens from 80% of AFP cases nationally (Figure 2). At the jurisdictional level, Western Australia was the only state to reach the WHO target, with adequate specimens collected from seven of the eight cases (88%) classified in 2011.

Laboratory testing of specimens

AFP cases

In 2010, a total of 54 faecal specimens were referred from 31 cases of AFP involving patients aged less than 15 years, while in 2011, 69 faecal specimens were referred (Table 4). The specimens included faeces, faecal extracts, swabs and cell culture isolates. No poliovirus was isolated from any of these specimens.

Non-polio enteroviruses (NPEV) were reported from three AFP cases in 2010. In the first case, echovirus 19 was isolated from an unimmunised three year old with a diagnosis of anterior horn cell disease confirmed by a MRI consistent with myelitis.¹² The patient had an upper respiratory tract infection three weeks prior to presentation. There was no history of recent travel or contact with an overseas visitor. Initial laboratory investigation was hampered by the CSF being inhibitory by RT-PCR. The local hospital isolated enterovirus by virus culture from a nasopharyngeal aspirate and faeces, which was confirmed by RT-PCR. The original specimens and virus isolates were referred to the NERL and echovirus 19 was identified from the virus isolates by sequencing a fragment of the capsid encoding region of the virus genome.

Echovirus 3 was identified from the first of two faecal specimens referred from the second AFP case involving a 12 year old patient diagnosed with men-

ingitis. The specimen was collected nine days after the onset of symptoms while the second specimen was collected 13 days after onset.

In the third case, enterovirus 68 (EV68) was identified from a single faecal specimen received by the NERL from a patient with spinal cord ischaemia. The virus was detected by RT-PCR from the faecal extract while virus culture was negative.

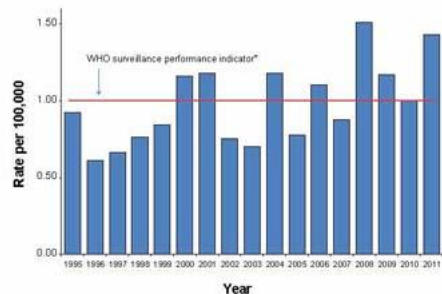
In 2011, a NPEV was detected in two faecal specimens from one AFP case. An atypical enterovirus cytopathic effect was observed in virus culture and was later confirmed by pan-enterovirus RT-PCR. The virus was typed from the stool extract by sequencing a fragment of the VP1 genomic region and identified as coxsackievirus type A24.

In 2010 and 2011, a total of 13 specimens were received from eight cases involving patients greater than 14 years of age, which is outside of the WHO AFP surveillance criterion. No enterovirus was isolated from the specimens.

Enterovirus Surveillance

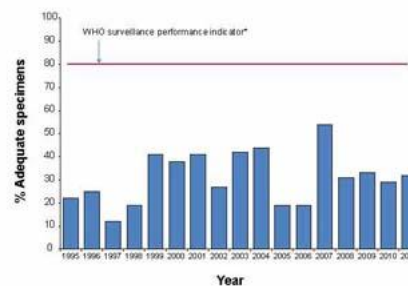
No poliovirus was detected by the ERLNA in 2010 and 2011. The ERLNA typed 234 NPEVs in 2010 and 331 in 2011. Coxsackievirus A6 (CA6) and enterovirus 71 (EV71) were amongst the leading enteroviruses detected in Australia in 2010 and 2011. Both viruses are typically associated with hand, foot and mouth disease in children. B5 was the predominant subgenogroup of EV71 detected in Australia in 2010 and 2011 and was detected in post-mortem specimens from an infant in 2010. Coxsackievirus B1 was detected in both the western and eastern states in 2009 and 2010. Echovirus 25 was the most common enterovirus identified in 2011.

Figure 1: Non-polio AFP rate classified by the PEP 1995 to 2011



* The WHO AFP surveillance performance indicator for a polio non-endemic country is one case per 100,000 children <15 years of age.

Figure 2: Percentage of AFP cases with adequate faecal specimens, 1995 to 2011



* The criterion for the WHO surveillance performance indicator is the collection of two faecal specimens more than 24 hours apart and within 14 days of the onset of symptoms.

Table 3: Surveillance for AFP cases in children less than 15 years, Australia, 2010 to 2011, compared with the WHO performance indicators

WHO surveillance performance indicator for AFP cases in children less than 15 years*	Australia's AFP surveillance performance	
	2010: 41 AFP cases expected	2011: 42 AFP cases expected
Non-polio AFP case rate of 1.0 / 100,000 children	Non-polio AFP rate 1.0 / 100,000 children (41 cases classified)	Non-polio AFP rate 1.43 per 100,000 children (60 cases classified)
More than 80% of classified AFP cases with adequate faecal specimens†	29% (adequate specimens received from 12 / 41 AFP cases)	32% (adequate specimens received from 19 / 60 AFP cases)

* Population data derived from the Australian Bureau of Statistics, estimated population, at 30 June 2009 and 2010. Available at www.abs.gov.au. Based on Australia's population less than 15 years of age.

† Adequate faecal specimens defined as 2 faecal specimens collected at least 24 hours apart and within 14 days of onset of paralysis.

Table 4: Test results for faecal specimens from AFP cases involving children < 15 years of age referred to the NERL, 2010 and 2011

Result	2010	2011
Non-polio enterovirus*	3	2
No enterovirus isolated	51	67
Total	54	69

* In 2010, non-polio enteroviruses identified from three AFP cases were echovirus 19, echovirus 3 and enterovirus 68. In 2011, coxsackievirus A24 was identified from two faecal specimens of one AFP case.

It was mainly associated with fever and most cases originated from Western Australia.

Environmental Surveillance

In 2010, sentinel sites for environmental surveillance for poliovirus were established at Armidale and two sites in Newcastle at Burwood Beach and Shortland. The Burwood Beach site was replaced by Byron Bay in 2011. Fifteen collections (5 from each site) were tested in 2010 and a further nine collections (3 from each site) in 2011. No poliovirus was reported from any of the 500 mL grab samples processed by the NERL.

The sewage extracts were tested in parallel by cell culture and a pan-enterovirus RT-PCR. The pan-enterovirus RT-PCR is a validated in-house test and was utilised to confirm the cell culture results as not all human enteroviruses can infect the RD-A cell line. All enterovirus isolates from cell culture and positive detections by RT-PCR were investigated to determine the virus type by nucleic acid sequencing and the results from the two methods were the same. All samples except one from each of the Newcastle sites and one from Byron Bay were positive for NPEV in virus culture, which serves as an indicator organism for the collection, transportation and laboratory procedures.

Regional reference laboratory activities

The following activities were performed as a Polio Regional Reference Laboratory in 2010 and 2011:

- Brunei Darussalam: Specimens from three AFP cases were received in each year. Sabin poliovirus types 1, 2 and 3 were isolated from one case in 2011, consistent with recent immunisation.
- Pacific Island countries: Specimens from 14 and 9 AFP cases were received in 2010 and 2011, respectively. No poliovirus was isolated from any of these but NPEVs were isolated from 6 of the 20 cases. In 2010, 14 specimens were referred from Fiji to investigate the cause of a hand, foot and mouth disease outbreak. Coxsackievirus A6 was detected in 11 of these. Twenty faecal specimens were referred from a gastroenteritis outbreak in Fiji in 2011. A NPEV was isolated from only one of the specimens, indicating that an enterovirus was unlikely to have caused the outbreak.
- Papua New Guinea: Specimens from 17 and 9 AFP cases were referred in 2010 and 2011, respectively. Sabin-like poliovirus type 3 was isolated from one case in 2010. Fourteen NPEVs were isolated by cell culture or detected by RT-PCR from 14 of the 26 cases.
- Philippines: Four poliovirus type 2 and 3 poliovirus type 3 viruses were referred from AFP cases for ITD in 2010 and all were characterised as Sabin-like. A poliovirus type 3 and a poliovirus type 1 were characterised as Sabin-like from sources other than AFP in 2010 and 2011, respectively.

Table 5: Enterovirus test results at the NERL, Australia, 1995 to 2011

Year	Poliovirus		Non-polio enterovirus	No enterovirus detected	EVID results referred [§]	Total samples reviewed
	Sabin-like	Non-Sabin-like				
1995	190	0	200	13	0	403
1996	224	0	198	9	0	431
1997	124	0	76	0	0	200
1998	52	0	15	4	0	71
1999*	60	1	9	9	0	79
2000	45	0	44	47	0	136
2001*	46	5	33	75	0	159
2002	36	0	21	49	0	106
2003	9	0	15	47	0	71
2004	6	0	26	61	0	93
2005	18	0	10	39	0	67
2006	2	0	6	71	29	108
2007†	0	2	32	115	107	256
2008	0	0	20	92	77	189
2009‡	1	0	63	78	113	255
2010	0	0	170	39	108	317
2011	0	0	174	61	205	440

* Untyped enterovirus or uncharacterised poliovirus isolates were referred for further testing after completion of a laboratory inventory. The six isolates tested as non-Sabin-like and were subsequently identified as wild type poliovirus prototype strains and were destroyed.

† Wild poliovirus type 1 was imported from Pakistan.

‡ A Sabin-like poliovirus type 1 was identified from an unimmunised infant.

§ Enterovirus identification results include retrospective data made available via the ERNLA.

Quality Assurance Programs

In 2010 and 2011, the NERL passed the annual quality assurance panels for poliovirus isolation by cell culture and poliovirus RT-PCR for ITD and vaccine derived poliovirus. In 2011, WHO introduced a trial poliovirus sequencing proficiency panel with the intention of making this an annual test from 2012. The NERL maintained accreditation as a Polio Regional Reference Laboratory after WHO conducted a two day on-site review of laboratory procedures and documentation in October 2010. Accreditation is assessed in the interim years by the submission of an annual checklist and subject to passing the annual laboratory proficiency tests.

Discussion

Clinical surveillance for cases of AFP in children is a sensitive means of detecting imported cases of poliomyelitis in a polio-free country by targeting an age group that is at high risk of infection if not immunised. This occurred in China when an outbreak of polio was reported in the far western province of

Xinjiang due to an importation of a wild poliovirus type 1 from Pakistan in August 2011.² Initially four polio cases were confirmed through AFP surveillance, in children aged between four months and two years, who had had onset of paralysis the preceding month. A further 17 cases of polio were confirmed up to October 2011 and included two fatalities. Polio cases were also reported in adults. The outbreak was stopped by concerted vaccination campaigns targeting people up to 39 years of age.² This serves as a salutary reminder that polio can quickly spread in a population with inadequate polio vaccination coverage at any age.

Australia met the WHO AFP surveillance performance indicator of at least 1.0 non-polio AFP case per 100,000 children less than 15 years of age in 2010 and 2011, reporting a rate of 1.0 and 1.4 per 100,000 respectively. This was the fourth year in a row that Australia has met the indicator used by WHO to assess whether a national AFP surveillance system is likely to detect an imported case of polio in a child. At the state and territory level, New South Wales, South Australia and Victoria exceeded the performance indicator rate in both years, while the Northern Territory, Tasmania and

Western Australia only did so in 2011. Queensland did not reach the surveillance indicator in either year despite it being the only jurisdiction where AFP is notifiable. Six of Australia's eight jurisdictions met the WHO performance indicator in 2011, a result that reinforces the overall sensitivity of the national AFP surveillance system.

Another important aspect of AFP surveillance is the testing of stool specimens to exclude poliovirus as the causative agent. Notwithstanding the strong performance of AFP case ascertainment in Australia in recent years, the number of AFP cases with adequate stool specimens has averaged 31% for the last four years. This compares to the WHO surveillance performance indicator of 80%, a target that Australia has never met. Australia's standard of performance against this surveillance indicator is not unusual for developed nations. The reasons for poor rates of faecal specimen collection are manifold. This could include an unwillingness or inability of the patient to provide a faecal specimen, prioritising laboratory tests of specimens from other sites such as cerebrospinal fluid and relying upon neurological diagnostic procedures such as magnetic resonance imaging and nerve conduction studies.

Virological surveillance for poliovirus was introduced in Australia to complement the clinical surveillance program for AFP cases in children. It has two components, the typing of enteroviruses through the ERLNA and environmental surveillance by testing grab samples of sewage at sentinel sites. No poliovirus was isolated through either of these surveillance systems providing additional data in support of Australia's continued polio-free status. In addition to testing for poliovirus, enterovirus typing facilitates the detection of NPEVs of public health importance, such as EV71, and will enable the epidemiology of enteroviruses circulating in Australia to be better understood. The predominant subgenogroup of EV71 detected in Australia in 2010 and 2011 was B5 and, based on reports from the ERLNA this subgenogroup was mainly associated with fever and hand, foot and mouth disease. EV71 B5 was also detected in faecal material from a post-mortem sample from an infant suffering from a suspected viral illness in 2010 but causality cannot be confirmed due to the detection of the virus from a non-sterile site. A report from New South Wales in 2011 linked EV71 meningoencephalitis with the death of a 63 year old male who had received rituximab as treatment for non-Hodgkin's lymphoma.¹³ The EV71 was typed as subgenogroup C2. Anti-CD20 monoclonal antibody therapy can deplete B cells, which are required to clear enterovirus infection.

As a result of the broader focus on enterovirus surveillance in support of poliovirus surveillance, the Polio Reference Laboratory was renamed the National Enterovirus Reference Laboratory from July 2011.

Acknowledgements

The Australian National Enterovirus Reference Laboratory is funded by the Australian Government DoHA, the Victorian Government Department of Health and VIDRL. We thank the APSU and the PAEDS team for their ongoing collaboration with acute flaccid paralysis surveillance and members of the Polio Expert Panel. We would like to acknowledge the important contribution of the notifying clinicians, laboratories (both national and international) and water authorities for their assistance with the work described in this report.

Author details

Mr Jason Roberts, Senior Medical Scientist¹
Ms Linda Hobday, Medical Scientist¹
Mrs Aishah Ibrahim, Medical Scientist¹
Mr Thomas Aitken
Dr Bruce Thorley, Senior Medical Scientist, Laboratory Head¹

1. National Enterovirus Reference Laboratory, Victorian Infectious Diseases Reference Laboratory, North Melbourne, Victoria, Australia.

Corresponding author: Mr Jason Roberts, National Enterovirus Reference Laboratory, Victorian Infectious Diseases Reference Laboratory (VIDRL), Locked Bag 815, Carlton, South Victoria 3053. Telephone: +61 3 9342 2607; Facsimile: +61 3 9342 2665; E-mail: jason.roberts@mh.org.au

References

1. Poliomyelitis. In: Australian Government Department of Health and Ageing. The Australian Immunisation Handbook. 10th edn. Canberra: Commonwealth of Australia; 2013. p. 338-344.
2. Progress Toward Interruption of Wild Poliovirus Transmission—Worldwide, January 2011–March 2012. *Morbidity and Mortality Weekly Report*, 2012; 61(19):353-357.
3. Pym M, Adams J, Booy R, Buttery J, Elia S, Elliott E et al. The development and trial of paediatric active enhanced disease surveillance (PAEDS): a new surveillance mechanism for Australia. *J Paed Child Health* 2008;44:A16.
4. Wood DJ, Hull B. L20B Cells Simplify Culture of Polioviruses from Clinical Samples. *J Med Virol* 1999, 58:188-192.
5. World Health Organization. Polio Laboratory Manual, 4th edition. *Department of Immunization, Vaccines and Biologicals* 2004. WHO/IVB/04.10.
6. Kilpatrick DR, Yang CF, Ching K, Vincent A, Iber J, Campagnoli R, Mandelbaum M, De L, Yang SJ, Nix A. . 2009. Rapid group-, serotype-, and vaccine strain-specific identification of poliovirus isolates by real-time reverse transcription PCR using degenerate primers and probes containing deoxymosine residues. *J Clin Microbiol* 2009;47:1939-1941.

-
7. Stewardson AJ, Roberts JA, Beckett CL, Prime HT, Loh PS, Thorley BR, Daffy JR.. An Imported Case of Poliomyelitis in Melbourne, Australia. *Emerging Infect Diseases* 2009;15:63-65.
 8. Carnie JA, Lester R, Moran R, Brown L, Meagher J, Roberts JA, Thorley BR. Public Health Response to Imported Case of Poliomyelitis, Australia. *Emerg Infect Dis* 2009;15:1733-1737.
 9. Roberts JA, Thorley BR. Enterovirus. In: Schuller M, Sloots TP, James GS, Halliday CL, Carter IWJ, editors. *PCR for Clinical Microbiology: An Australian and International Perspective*. 1st edn. Dordrecht: Springer; 2010. p. 229-233.
 10. Nix WA, Oberste MS, Pallansch MA. Sensitive, seminested PCR amplification of VP1 sequences for direct identification of all enterovirus serotypes from original clinical specimens. *J Clin Microbiol* 2006;44:2698-2704.
 11. World Health Organization. Guidelines for environmental surveillance of poliovirus circulation. Geneva: Department of Vaccines and Biologicals WHO/V&B/03.03; 2003.
 12. Kesson AM, Ming CC, Troedson C, Thorley BR, Roberts JA. Echovirus 19 associated with a case of acute flaccid paralysis. *Journal of Paediatrics and Child Health*. doi: 10.1111/jpc.12043.
 13. Ahmed R, Buckland M, Davies L, Halmagyi GM, Rogers SL, Oberste S, Barnett MH. Enterovirus 71 meningoencephalitis complicating rituximab therapy. *J Neurol Sci* 2011;305:149-151.

LETTERS

Letter to the editor: Application of Bayesian methods to the inference of phylogeny for enterovirus surveillance

J A Roberts (jason.roberts@mh.org.au)^{1,2}, A Hung², B R Thorley^{1,2}

1. National Enterovirus Reference Laboratory Australia, Victorian Infectious Diseases Reference Laboratory, Victoria, Australia
2. School of Applied Sciences and Health Innovations Research Institute, RMIT University, Melbourne, Australia

Citation style for this article:

Roberts JA, Hung A, Thorley BR. Letter to the editor: Application of Bayesian methods to the inference of phylogeny for enterovirus surveillance. *Euro Surveill.* 2013;18(9):pii=20409. Available online: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20409>

Article submitted on 19 February 2013 / published on 28 February 2013

To the editor:

We read with interest the article by Niesters et al. describing a pilot study to share enterovirus sequence data within the Netherlands for epidemiological investigation [1]. As the authors note, nucleotide sequence data can extend the identification of virus serotype to the tracing of disease transmission patterns through phylogenetic analysis. The article referred to use of a web-based typing tool for enteroviruses and noroviruses based on BLAST analysis followed by a neighbor-joining phylogeny. We believe such analyses can be further enhanced through the inclusion of temporal and geographical discrete variables for the inference of phylogeny (phylogeography). In particular, the application of Bayesian inference to phylogeography offers several advantages, such as the capacity to explicitly account for parameter uncertainty, reducing potential model bias, especially where data are scarce. This approach was referred to in the same issue of *Eurosurveillance* by Carriço et al. for bacterial molecular epidemiology [2] and has been successfully applied to virological studies for avian influenza A(H5N1) and rabies [3]. Carriço et al. make reference to the computational demands of such methods [2] and indeed, this is an important consideration. From our experience, the use of commodity graphics processing units combined with appropriate parallel threading software extensions can overcome some of the limitations imposed by high computational demand at low cost [4].

Australia established an Enterovirus Reference Laboratory Network – primarily for poliovirus surveillance but also to detect other enteroviruses of public health significance, such as enterovirus 71 (EV71) – through the sharing of enterovirus sequence data. EV71 infection often manifests benignly as hand, foot and mouth disease in infants but has been associated with fatal neurological disease, particularly in the Asia-Pacific region [5]. The correlation of specific EV71 subgenogroups with increased neurological presentation,

such as C₄ infection recently in Cambodia, China and Vietnam, highlights the benefit of a phylodynamic analysis that accounts for temporal and geographical parameters. We feel that the application of Bayesian methods to the inference of phylogeny will play an important role in the elucidation of global chains of enterovirus transmission.

Conflict of interest

None declared.

Authors' contributions

Jason A Roberts conceived the idea to respond to the recent articles on molecular epidemiology in *Eurosurveillance* and drafted the letter. The text was revised and edited by Andrew Hung and Bruce R Thorley, and all authors approved the final version of the letter.

References

1. Niesters HG, Rossen JW, van der Avoort H, Baas D, Benschop K, Claas EC, et al. Laboratory-based surveillance in the molecular era: the TYPENED model, a joint data-sharing platform for clinical and public health laboratories. *Euro Surveill.* 2013;18(4):pii=20387. Available from: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20387>
2. Carriço JA, Sabat AJ, Friedrich AW, Ramirez M, on behalf of the ESCMID Study Group for Epidemiological Markers (ESGEM). Bioinformatics in bacterial molecular epidemiology and public health: databases, tools and the next-generation sequencing revolution. *Euro Surveill.* 2013;18(4):pii=20382. Available from: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20382>
3. Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. *PLoS Comput Biol.* 2009;5(9):e1000520.
4. Ayres DL, Darling A, Zwickl DJ, Beerli P, Holder MT, Lewis PO, et al. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Syst Biol.* 2012;61(1):170-3. <http://dx.doi.org/10.1093/sysbio/syr100>
PMid:21963610 PMCID:3243739
5. Solomon T, Lewthwaite P, Perera D, Cardoso M, McMinn P, Ooi MH. Virology, epidemiology, pathogenesis, and control of enterovirus 71. *Lancet Infect Dis.* 10(11):778-90.

Letter

Electron microscope detection of an endogenous infection of retrovirus-like particles in L20B cells

Jason A. Roberts, Bruce R. Thorley, Leesa D. Bruggink
and John A. Marshall*

Victorian Infectious Diseases Reference Laboratory, 10 Wreckyn Street, North Melbourne, Victoria 3051, Australia

*To whom correspondence should be addressed. E-mail: john.marshall@mh.org.au

Abstract L20B cells are a cell line commonly used for the isolation of poliovirus. The current study indicates that L20B cells are chronically infected with a retrovirus-like particle that replicates in the cytoplasm and buds through the plasma membrane. The findings indicate that care is needed in the use of L20B cells for certain virus isolation studies and emphasize the importance of electron microscope studies as an adjunct to the development of diagnostic virology protocols.

Keywords L20B cells, retrovirus-like particles, poliovirus

Received 14 November 2012, accepted 24 December 2012; online 16 January 2013

L20B cells are a cell line derived from mouse L cells that have been transfected with a cDNA clone of the poliovirus receptor [1]. The cell line is widely used for the detection of poliovirus and can also detect some other enteroviruses [2,3]. Since 1999, World Health Organization Global Polio Network laboratories have been using L20B cells for poliovirus isolation [3].

The current study documents a thin sectioning electron microscope analysis of uninfected and poliovirus-infected L20B cells. The findings indicate that both uninfected and poliovirus-infected L20B cells can contain high concentrations of a virus other than poliovirus. The characteristics of the endogenous retrovirus-like particles in L20B cells are given and the implications of these findings for virological studies using L20B cells are discussed.

Poliovirus (Sabin poliovirus serotype 1)-infected L20B cells were grown by standard procedures to a cytopathic effect of 20–25% [4]. Infected and control uninfected L20B cells were then processed for thin sectioning electron microscopy, which was carried out essentially as described by Shehu-Xhilaga *et al.* [5].

Retrovirus-like particles were common in both poliovirus-infected and uninfected L20B cells. The particles were round or ovoid in shape with a distinct core and were commonly seen within the cytoplasm (Fig. 1) or as a budding form (Fig. 2). Intracytoplasmic forms had a diameter of 52.0 ± 8.3 nm [mean (m) \pm standard deviation (SD), $n = 8$], with a core diameter of 23.9 ± 3.4 nm (m \pm SD, $n = 8$). Budding forms showed an additional outer layer measuring 8.4 ± 1.6 nm (m \pm SD, $n = 5$) in thickness. Budding forms were considerably larger than intracytoplasmic forms, with a diameter (excluding the outer layer) of 75.1 ± 2.3 nm (m \pm SD, $n = 3$) and a core diameter of 46.7 ± 10.4 nm (m \pm SD, $n = 3$).

It is notable that Laue *et al.* [6] have previously referred, without any other comment, to 'retrovirus-infected L20B cells' but the current study appears to be the first to document in detail the presence of retrovirus-like particles in L20B cells. The findings indicate that care is needed in the use of enterovirus-infected L20B cells for morphological and biochemical studies and emphasize the

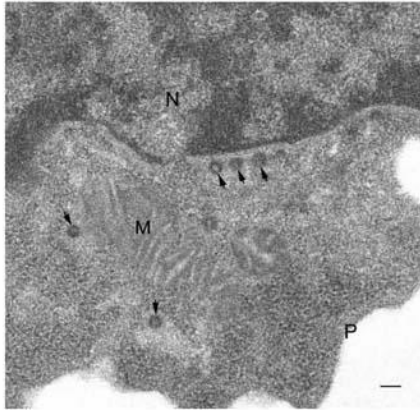


Fig. 1. Thin sectioning electron micrograph showing intracytoplasmic retrovirus-like particles (arrows) associated with L20B cells. N, nucleus; M, mitochondrion; P, plasma membrane. Scale bar: 100 nm.

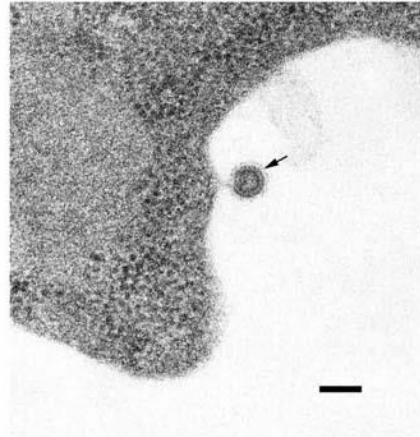


Fig. 2. Thin sectioning electron micrograph showing a budding retrovirus-like particle (arrow) associated with L20B cells. Scale bar: 100 nm.

importance of electron microscope studies as an adjunct to the development of diagnostic virology protocols. Further studies of the characteristics of L20B cells are warranted.

References

- 1 Pipkin P A, Wood D J, Racaniello V R, and Minor P D (1993) Characterisation of L cells expressing the human poliovirus receptor for the specific detection of polioviruses in vitro. *J. Virol. Methods* **41**: 333-340.
- 2 Wood D J and Hull B (1999) L20B cells simplify culture of polioviruses from clinical samples. *J. Med. Virol.* **58**: 188-192.
- 3 Nadkarni S S and Deshpande J M (2003) Recombinant murine L20B cell line supports multiplication of group A coxsackieviruses. *J. Med. Virol.* **70**: 81-85.
- 4 World Health Organization (2004) *Polio Laboratory Manual*. WHO/IVB/04.10. 4th ed (WHO Department of Immunization, Vaccines and Biologicals, Geneva).
- 5 Shehu-Xhilaga M, Hill M, Marshall J A, Rappes J, Crowe S M, and Mak J (2002) The conformation of the mature dimeric human immunodeficiency virus type 1 RNA genome requires packaging of Pol protein. *J. Virol.* **76**: 4331-4340.
- 6 Laue M, Niederwöhmeier B, and Bannert N (2007) Rapid diagnostic thin section electron microscopy of bacterial endospores. *J. Microbiol. Methods* **70**: 45-54.



INSTRUCTIVE CASE

Echovirus 19 associated with a case of acute flaccid paralysisAlison M Kesson,^{1,3} Chong Ming Choo,¹ Christopher Troedson,² Bruce R Thorley⁴ and Jason A Roberts⁴Departments of ¹Infectious Diseases and Microbiology and ²Neurology, The Children's Hospital at Westmead, Westmead, ³Discipline of Paediatrics and Child Health, Sydney Medical School, The University of Sydney, Sydney, New South Wales and ⁴National Poliovirus Reference Laboratory, Melbourne, Victoria, Australia

Abstract: Acute flaccid paralysis can be caused by many members of the enterovirus genus, most notably the three polioviruses types 1 to 3. We report the case of acute flaccid paralysis caused by echovirus 19. The Western Pacific region has been declared polio free by the WHO since 2000. Australia is now using inactivated polio vaccine in the National Immunization Schedule. This vaccine does not carry the extremely rare risk of vaccine associated acute flaccid paralysis but it does leave our newly vaccinated population open gastrointestinal infection with polioviruses and the risk of circulation of the wild-type virus. Continued surveillance of cases of acute flaccid paralysis is to detect polioviruses is essential until poliovirus is completely eradicated.

Key words: acute flaccid paralysis, poliovirus, surveillance

Human enteroviruses are members of the Picornaviridae family and comprise greater than 100 members divided into four species A to D, based on RNA sequence homology within the capsid protein-encoding region.

Enterovirus infection occurs world-wide and can result in a variety of diseases, with a clinical spectrum ranging from asymptomatic infection to very severe disease affecting many body systems with variable clinical manifestations.¹ The prototypic severe enterovirus infection of the nervous system that causes acute flaccid paralysis (AFP) is poliomyelitis. The World Health Organization (WHO) Global Polio Eradication Initiative (GPEI), which commenced in 1988, has resulted in a greater than 99% reduction in global cases of poliomyelitis, with the Western Pacific being declared polio free in 2000. In 2009, 1604 cases of AFP with wild poliovirus identified were reported, with four countries – Afghanistan, India, Nigeria and Pakistan –

remaining endemic. However, globally, there continues to be reports of sporadic cases of AFP due to non-polio enteroviruses and, more recently, large outbreaks of hand, foot and mouth disease due to enterovirus 71, with severe neurological sequelae including AFP in eastern European and Asian countries.^{2,3}

Case Report

A 3-year-old Caucasian, Australian-born girl presented with a 2-day history of fever to 38.8°C and a 24-h history of painless inability to move her left upper arm and shoulder, but with preservation of some finger movements. A medical history revealed that she had been unwell, with an upper respiratory tract infection 3 weeks prior to presentation. She had not received any of her routine childhood immunisations and was unimmunised against polioviruses. She had started day care 2 months prior to this illness. There was no history of recent travel or contact with an overseas visitor; however, 1 week prior to this illness, her unimmunised sister had an upper respiratory tract infection with fever and sore throat.

Physical examination showed weakness confined to the left upper limb, with power at the shoulder and elbow both 0 of 5. There was some degree of wrist flexion and extension and preservation of finger movements, with an absence of biceps, triceps and brachioradialis reflexes. The muscle tone was flaccid, and sensation was preserved and normal. The rest of her neurological examination was normal. She had a mild generalised erythematous rash, and her other systems were normal to examination.

A lumbar puncture was performed, and the cerebrospinal fluid (CSF) showed $2 \times 10^6/L$ polymorphs, $12 \times 10^6/L$ monocytes and $1450 \times 10^6/L$ red blood cells. CSF protein was 0.22 mmol/L, and glucose was 2.4 mmol/L, both within normal limits. A reverse transcription polymerase chain reaction (RT-PCR) for enteroviruses was performed on the CSF, and this was indeterminate due

Key Points

- 1 Acute flaccid paralysis can be caused by enteroviruses other than polioviruses 1, 2 or 3.
- 2 Inactivated poliovirus vaccine, unlike live attenuated Sabin poliovirus vaccine, does not prevent the replication of polioviruses in the gastrointestinal tract and the risk of transmission to another person. They prevent acute flaccid paralysis in immunised patients.
- 3 On going surveillance of cases of acute flaccid paralysis to detect polioviruses is essential.

Correspondence: Professor Alison M Kesson, Department of Infectious Diseases and Microbiology, The Children's Hospital at Westmead, LMB 4001, Westmead, NSW 2145, Australia. Fax: +61 2 98453291, email: alisonk2@health.nsw.gov.au

Conflict of interest: None declared.

Accepted for publication 3 January 2012.



Fig. 1 Sagittal T1-weighted magnetic resonance image of the cervical spine showing increased signal of the anterior horn cells.

to the presence of PCR inhibitors. Her serum immunoglobulins and immunoglobulin G (IgG) subclasses were all within the normal range for age, and her enterovirus complement fixation titre was 256, consistent with recent infection.

A magnetic resonance image scan of her upper spine was performed, and this showed mild swelling of the cervical spinal cord, extending from C2/3 to C6/7, with subtle high signal changes centrally, all changes being consistent with myelitis (Fig. 1).

Viral cultures of nasopharyngeal aspirate and faeces were performed 4 days after the onset of paralysis, and an enterovirus was isolated from both specimens. These cultures were labelled with fluorescent antibodies to polio 1, 2 and 3, enterovirus 71, Coxsackie A16, echo 4, 6, 9, 11, 30 and 33, and Coxsackie B1 to 6, and all were negative. The specimens and viral isolates were then referred to the Australian National Polio Reference Laboratory for serotype identification via a fragment of the VP1 gene using the CODEHOP semi-nested RT-PCR method.⁴ The resultant sequence data were compared with a local database containing prototype VP1 sequence data, and the virus was identified as an echovirus 19. Sequences for comparative analysis were acquired by using the BLAST search algorithm (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>). Phylogenetic analysis indicated a common ancestor with an environmental isolation from China (Fig. 2).

The patient was treated with 2 g/kg of intravenous immunoglobulin (IVIG) with some improvement. On discharge from hospital 1 week after admission, her power for elbow extension had improved to 4 of 5, and her power for elbow and shoulder flexion had improved to 3 of 5. She had physiotherapy and rehabilitation as an outpatient but had residual weakness of her upper limb after 6 months.

Discussion

Poliovirus vaccination offers no protection against non-polio enterovirus infection. The inactivated poliovirus vaccine (IPV) provides the recipient with systemic neutralising antibodies, but mucosal immunity is not as potent as with the oral poliovirus vaccine. Thus, wild poliovirus infection in individuals who have received IPV may result in replication of poliovirus in the gastrointestinal tract.⁵ Oral Sabin polio vaccine (OPV) induces both systemic and mucosal immunity and can significantly reduce the likelihood of wild poliovirus replication within the gastrointestinal tract and reduce transmission of wild poliovirus to other individuals.⁶ The rationale behind the WHO GPEI is that poliovirus can only replicate in humans and that, by universal immunisation with OPV, poliomyelitis can be totally eradicated globally.

AFP is characterised by onset of painless weakness, with no loss of sensation in the affected area. Possible differential diagnoses include the Guillain-Barré syndrome, acute traumatic injuries, rabies, acute transverse myelitis, epidural abscess, toxin of *Clostridium botulinum* and tick-bite paralysis. Non-polio enteroviruses are known to cause sporadic AFP, and several serotypes from all four species (A to D) of human enterovirus have been anecdotally associated with paralytic disease.² This is the second case of AFP associated with echovirus 19 reported in the literature, and unlike the other case, our patient was not immunosuppressed.^{6,7} Epidemiological evidence suggests that non-polio enteroviruses cause less severe disease than polioviruses with less residual paralysis or muscle atrophy.²

Our patient was treated empirically with IVIG, as a neutralising antibody against a specific enterovirus has been successful in the management of patients with hypogammaglobulinaemia and chronic enterovirus meningoencephalitis⁸; however, in this situation, the treatment has not been curative. Our patient did recover some power in her left upper limb, but complete recovery had not occurred after 6 months. We did not use Pleconaril in this patient, because this product is no longer available.

In 2000, the WHO declared the Western Pacific Region, including Australia, poliovirus free, and in November 2005, Australia changed from using OPV to IPV in the national immunisation programme provided for all children born in Australia. This change to IPV was in part due to the very small but serious risk of vaccine-associated AFP. Due to this change, Australia now has more than 1 million infants and young children who have potentially reduced mucosal immunity to polioviruses and in whom wild polioviruses, if reintroduced into Australia, could circulate undetected before a clinical case is detected. It is imperative that clinical and laboratory surveillance for AFP continues, with attempts to rapidly detect and accurately identify all enteroviruses associated with AFP so as to ensure that poliovirus has not inadvertently re-entered Australia with the potential of recurrent endemic disease. The last reported case of imported poliovirus in Australia was in 2007, from a 22-year-old student with a previous history of vaccination with OPV.

The Australian AFP clinical surveillance programme was established in 1995 as a commitment to the WHO GPEI. All potential cases of AFP are notified, and relevant clinical specimens are sent to the National Polio Reference Laboratory for

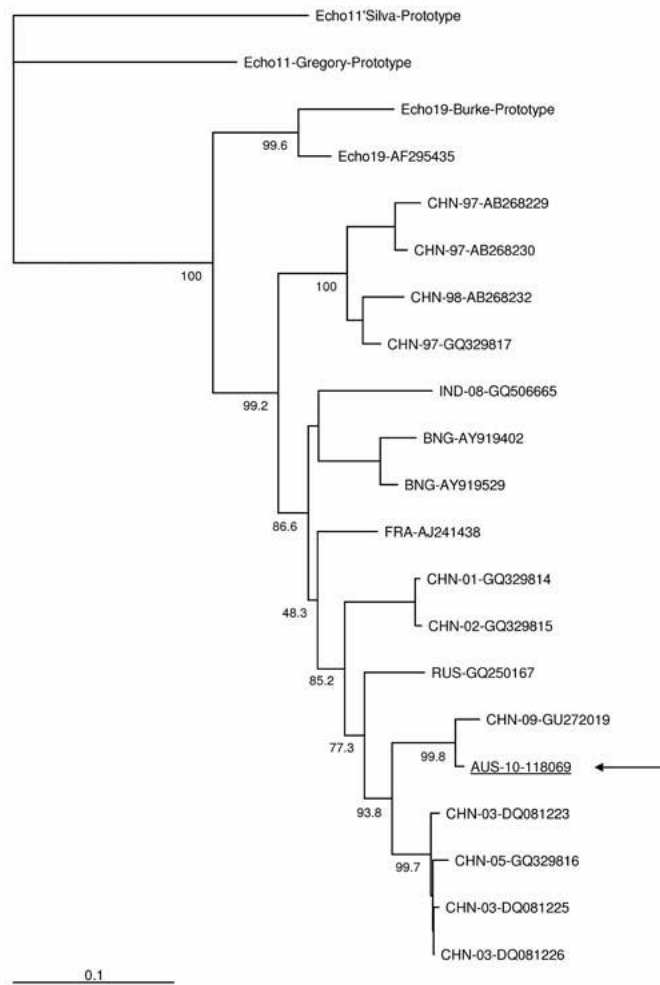


Fig. 2 VP1 phylogenetic tree constructed using sequence data derived from GenBank, created using the PHYLIP DNA distance algorithm (University of Washington, Seattle, WA, USA) and neighbour-joining with bootstrap analysis performed using 1000 pseudoreplicates. The scale bar indicates nucleotide substitutions per site. Sequences are indicated by country, year and sequence accession number or other IDs (shown where available). AUS, Australia; BNG, Bangladesh; CHN, China; FRA, France; IND, India; RUS, Russia.

virus isolation and identification of enteroviruses. Continued active surveillance for poliovirus is essential if this disease is to be eradicated globally.

Acknowledgement

We would like to thank Mr Ken McPhie of the Institute for Clinical Pathology and Medical Research, Westmead Hospital, for the rapid immunofluorescent studies of the enterovirus cultures.

Multiple Choice Questions

- 1 Acute flaccid paralysis is not caused by
 - A poliovirus
 - B sabin oral poliovirus vaccine
 - C salk inactivated poliovirus vaccine
 - D echovirus 19
 - E enterovirus 71

Answer: C. Inactivate poliovirus cannot replicate and cause disease.

- 2 Ongoing replication of enteroviruses in the gastrointestinal tract with failure to clear infection can occur in patients with
- A T-cell immunodeficiencies
 - B NK-cell immunodeficiencies
 - C terminal complement deficiency
 - D hypogammaglobulinaemia
 - E high-dose steroids

Answer: D. Chronic enterovirus infections are associated with hypogammaglobulinaemia.

- 3 Enterovirus infections have not been associated with
- A thyroiditis
 - B pleurodynia
 - C meningitis
 - D acute flaccid paralysis
 - E myocarditis

Answer: A. Enterovirus infections have been shown to cause all of the syndromes listed, except for thyroiditis.

References

- 1 Cherry JD, Krogstad P. Enteroviruses and parechoviruses. In: Feigin R, Cherry J, Demmler G, Kaplan S, eds. *Textbook of Pediatric Infectious Diseases*, 6th edn. Philadelphia, PA: Saunders, 2009, 2110–70.
- 2 Dietz V, Andrus J, Olive J-M, Cochi S, de Quadros C. Epidemiology and clinical characterisation of acute flaccid paralysis associated with non-polio enterovirus isolation: the experience in the Americas. *Bull World Health Organ* 1995; **73**: 597–603.
- 3 Tee KK, Lam TTY, Chan YF *et al.* Evolutionary genetics of human enterovirus 71: origin, population dynamics, natural selection, and seasonal periodicity of the VP1 gene. *J. Virol.* 2010; **84**: 3339–50.
- 4 Nix WA, Oberste MS, Pallansch MA. Sensitive, seminested PCR amplification of VP1 sequences for direct identification of all enterovirus serotypes from original clinical specimens. *J. Clin. Microbiol.* 2006; **44**: 2698–704.
- 5 Onorato IM, Modlin JF, McBean AM, Thoms ML, Losonsky GA, Bernier RH. Mucosal immunity induced by enhanced-potency inactivated and oral polio vaccines. *J. Infect. Dis.* 1991; **163**: 1–6.
- 6 Modlin J. Poliovirus. In: Mandell GL, Bennett JE, Dolin R, eds. *Mandell, Douglas and Bennett's Principles and Practice of Infectious Diseases*, 7th edn. Philadelphia, PA: Churchill Livingstone Elsevier, 2010, 2345–51.
- 7 Starlin R, Reed N, Leeman B, Black J, Trulock E, Mundy L. Acute flaccid paralysis syndrome associated with echovirus 19, managed with pleconaril and intravenous immunoglobulin. *Clin. Infect. Dis.* 2001; **33**: 730–2.
- 8 Quartier P, Foray S, Casanova JL, Hau-Rainsard I, Blanche S, Fischer A. Enteroviral meningoencephalitis in X-linked agammaglobulinemia: intensive immunoglobulin therapy and sequential viral detection in cerebrospinal fluid by polymerase chain reaction. *Pediatr. Infect. Dis. J.* 2000; **11**: 1106–8.



Investigation of a predicted N-terminal amphipathic α -helix using atomistic molecular dynamics simulation of a complete prototype poliovirus virion

Jason A. Roberts^{a,b,*}, Michael J. Kuiper^c, Bruce R. Thorley^{a,b}, Peter M. Smooker^b, Andrew Hung^b

^a World Health Organization Poliomyelitis Regional Reference Laboratory, Victorian Infectious Diseases Reference Laboratory, North Melbourne, Australia

^b School of Applied Sciences and Health Innovations Research Institute, RMIT University, Melbourne, Australia

^c Victorian Life Sciences Computation Initiative, Parkville, Australia

ARTICLE INFO

Article history:

Accepted 22 June 2012

Available online 6 July 2012

Keywords:

Molecular dynamics

Poliovirus

Enterovirus

Virus

Simulation

Structural virology

ABSTRACT

The wild type 1 poliovirus capsid was first described in atomic detail in 1985 using X-ray crystallography. Numerous poliovirus capsid structures have been produced since, but none resolved the spatial positioning and conformation of a predicted N-terminal α -helix of the capsid protein VP1, which is considered critical to virus replication. We studied the helical structure under varying conditions using *in silico* reconstruction and atomistic molecular dynamics (MD) simulation methods based on the available poliovirus capsid atom coordinate data. MD simulations were performed on the detached N-terminal VP1 helix, the biologically active pentamer form of the pre-virion structure, reconstructed empty virus capsids and a full virion containing the poliovirus RNA genome in the form of a supercoiled structure. The N-terminal α -helix structure proved to be stable and amphipathic under all conditions studied. We propose that a combination of spatial disorder and proximity to the genomic RNA made this particular structure difficult to resolve by X-ray crystallography. Given the similarity of our *in silico* model of poliovirus compared to X-ray crystallography data, we consider computational methods to be a useful complement to the study of picornaviruses and other viruses that exhibit icosahedral symmetry.

© 2012 Elsevier Inc. All rights reserved.

1. Background

Poliovirus is classified within the enterovirus genus of the family *Picornaviridae* which includes more than 90 human enteroviruses [1]. The single stranded positive sense RNA genome is enclosed by a non-enveloped capsid composed of 60 copies of four viral proteins: VP1, VP2, VP3 and VP4 (Fig. 1). One copy of each VP protein forms a protomer subunit and five individual protomers form a pentamer structure. The complete capsid has an icosahedral formation made from 12 pentamers. Integral to the capsid structure are myristate molecules covalently bound to the N-terminus of VP4 and an endogenous lipid occupying a hydrophobic pocket accessible from the exterior surface of VP1 that is common to enteroviruses. The RNA genome is polyadenylated and VPg, a virus encoded protein, is covalently attached to the 5' terminus.

The wild poliovirus type 1 (Mahoney strain) capsid structure was resolved to 2.9 Å resolution by X-ray crystallography in 1985 [2]. However, certain regions could not be resolved from the

electron density maps: residues 1–20 of the N-terminus of VP1, residues 1–7 of the N-terminus of VP2, residues 234–238 of the C-terminus of VP3 and residues 1–12 of the N-terminus of VP4. The N-termini of VP1, VP2 and VP4 are each located on the inner surface of the virion and it was postulated that they could not be resolved by X-ray crystallography due to spatial orientation compared to the symmetric icosahedral structure [2]. Prior to assembly into the final capsid structure, biologically active proteins in the form of pentamers are present, comprising five protomers with VP1, VP3 and a precursor protein VP0, which is cleaved during the assembly of infectious 160S particles to form VP2 and VP4. This process is hypothesised to involve cleavage of a scissile bond in the presence of RNA, essentially locking the capsid in its final and infectious conformation [3].

Since the initial atomic description of the poliovirus capsid in 1985, more than twenty coordinate files of poliovirus 1, 2 and 3 capsid structures have been submitted to the Worldwide Protein Data Bank and include complete virus and empty capsid structures. However, the N-terminus of VP1 was not resolved by any of these investigations. This may have been due to spatial disorientation or molecular movement of the protein chain. Proximity to the RNA core within complete virus particles may also have rendered the polypeptide unresolvable by X-ray crystallography although this was considered not to be a contributing factor when the first crystal structure was published in 1985 [4]. The amino

* Corresponding author at: National Enterovirus Reference Laboratory, WHO Poliomyelitis Regional Reference Laboratory, Victorian Infectious Diseases Reference Laboratory, 10 Wreckyn Street, North Melbourne 3051, Australia. Tel.: +61 3 9342 2607; fax: +61 3 9342 2665. E-mail address: jason.roberts@mh.org.au (J.A. Roberts).

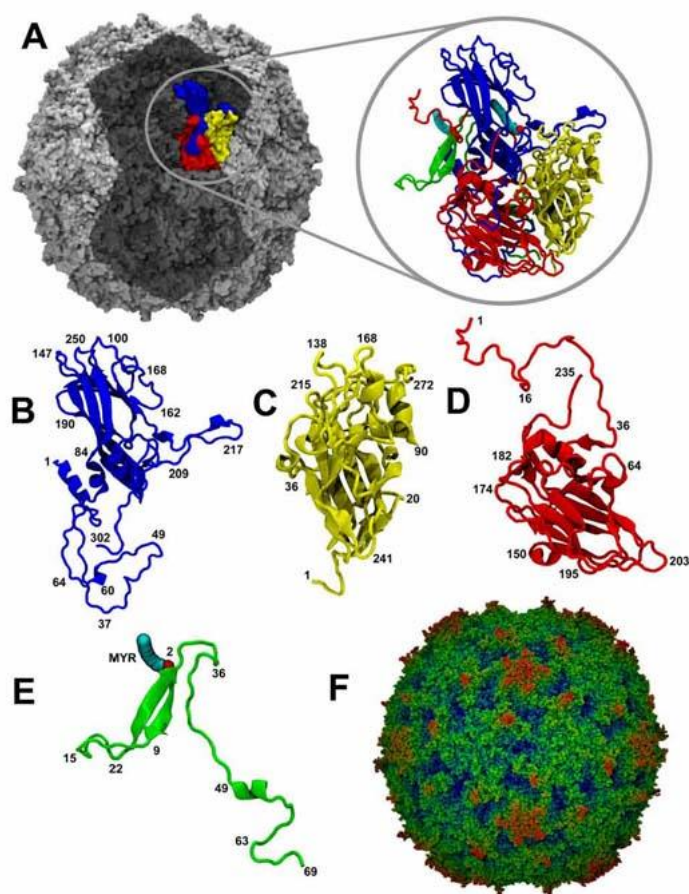


Fig. 1. Structural features of poliovirus. (A) Coloured section shows the relative positioning of a single protomer (1 of 60) within the icosahedral virus capsid. Dark grey shading indicates the relative position of pentamers intersecting at the 2-fold axis of symmetry. The breakout image on the right of (A) shows a cartoon representation of a protomer coloured by chain [26]. (B–E) Individual protomer subunits with selective amino acid residues numbered (B) VP1 = blue, (C) VP2 = yellow, (D) VP3 = red, (E) VP4 = green. (F) All-atom Van der Waals representation of poliovirus virion with radial colouring depicting the relative distance from the centre of the particle with blue = 120 Å to red = 150 Å.

acid sequence of the N-terminus of VP1 in many enteroviruses, including poliovirus and rhinoviruses, reveals a periodicity of hydrophobic residues that may form an amphipathic helix with polar residues along one half of the surface and non-polar residues on the opposite surface [5]. The complete VP1 crystal structure of human rhinovirus 16 was resolved as an amphipathic helix formation at the N-terminus providing support for this structure in other enteroviruses [6].

The N-terminus of VP1 is reversibly exposed to the exterior of native virus particles that sediment at 160S and is capable of binding liposomes, which is believed to be an intermediate step of the cell entry process of poliovirus infection [7,8]. Binding of the poliovirus receptor, CD155, leads to the irreversible externalization of the N-terminus of VP1 and loss of VP4 forming the 135S virus particle [9,10]. The intimate involvement of the N-terminus of VP1 in the poliovirus infection cycle necessitate

resolving its structure within the inner surface of the native capsid to gain a better understanding of the molecular processes involved.

The computational power available from supercomputers now enables molecular dynamics (MD) simulations of multi-million atom systems, facilitating the *in silico* investigation of macromolecular structures including viruses [11]. We have utilised MD software to perform atomistic simulations of the complete wild poliovirus type 1 Mahoney strain, including the native RNA sequence, which equated to approximately four million atoms. The X-ray crystallographic co-ordinates of wild poliovirus type 1 published in 2001 [12] were used to reconstruct the molecular structure and the missing VP polypeptide sequences were reconstituted according to the published prototype nucleic acid sequence (GenBank NC002058). Given the reported involvement of the N-terminus of VP1 in cellular infection by poliovirus we

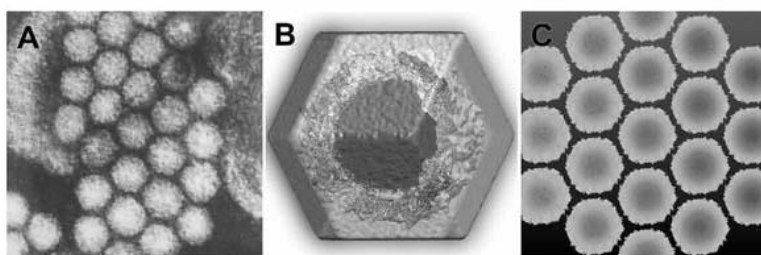


Fig. 2. (A) Transmission electron micrograph of a picornavirus sample showing honeycomb-like aggregation of virus particles. (B) MD simulation of full poliovirus virion suspended in saline solution with transparent density representation of water used to show periodic cell boundaries. (C) Transverse section of periodic cell reflections indicating the similar "honeycomb-like" appearance of the MD simulation of a full poliovirus virion to electron microscopy data.

report an investigation of the proposed amphipathic α -helical structure under various conditions to better understand its intrinsic stability.

2. Methods

Poliovirus capsid reconstructions were based on GenBank sequence ID NC002058. PSIPred predictions [13] derived from the amino acid sequence data and comparative information determined from enterovirus crystal structures available in the Research Collaboratory for Structural Bioinformatics protein data bank (RCSB PDB) were used to select an appropriate template for the reconstruction of the capsid and associated lipids. X-ray crystal structure of poliovirus type 1 Mahoney strain (PDB file 1HXS) at a 2.2 Å resolution [12] was selected as the template.

Reconstruction of the virus was achieved using the Visual Molecular Dynamics (VMD) [14] software package incorporating Compute Unified Device Architecture extensions [15]. A significant element lacking from the original crystal structure was the amino terminus of the VP1 protein common to members of the genus *Enterovirus* and reported to have an α -helical structure [5]. The 19 missing residues constituting the region containing the α -helix were reconstructed using the molefacture plug-in in VMD. Spatial alignment of the helix was based on the positioning of partial coordinate data in the original 1HXS PDB file. Topology files and CHARMM parameter data for the associated myristate and palmitate molecules were determined using the SWISSparam server (<http://swissparam.ch/>) [16].

Icosahedral symmetry data required to reconstruct the 60 virus protomers into a full capsid structure were available in the form of metadata within the original crystal structure. Using the "mono2poly" tcl script (http://www.ks.uiuc.edu/Research/vmd/script_library/scripts/mono2poly/) a single protomer was positioned according to the biological parameters in order to

assemble the 240 proteins and 120 lipids required for a full poliovirus capsid.

In order to recreate a structure analogous to a full wild poliovirus virion, the encoding genomic RNA from GenBank sequence ID NC002058 was split into 10 fragments of approximately 750 bp each and the 3'NTR was polyadenylated ($n=20$). The VPg protein was replicated using the PDB file 2BBP and covalently attached to the 5'NTR of the genome. Each RNA of the 10 RNA fragments was generated in a helical structure approximately 200 nm in length using the "make-na" server (<http://structure.usc.edu/make-na/index.html>). Steered molecular dynamics was used to compress each fragment to 300 nm. Once compressed the fragments were arranged in a supercoil formation. Steered molecular dynamics was used to direct the supercoiled RNA into a sphere approximately 300 Å in diameter then solvated and neutralised in a magnesium chloride solution. The RNA was steered further to 187 Å diameter and encapsidated by the previously constructed poliovirus capsid resulting in a final structure totalling 1.05 million atoms.

An additional 3 million atoms were required to replicate a 0.154 mM sodium chloride solution with cell edge padding at 16 Å on each axis with cubic cell dimensions of 344.8 Å. Due to the inefficient nature of a cubic system for the simulation of an icosahedral structure, rhombic-dodecahedron periodic boundary conditions were chosen to reduce the amount of water required to solvate the virus and better replicate the *in vitro* aggregation of multiple picornavirus particles (Fig. 2).

MD simulation was performed using supercomputers at the Victorian Life-Sciences Computation Initiative (VLSI) and the Nanoscale Molecular Dynamics (NAMD) software package [17] (<http://www.ks.uiuc.edu/Research/namd/>). Initially, 50,000 steps of energy minimisation were performed. This was followed by equilibration using a constant particle number, pressure and temperature ensemble (NPT) for 0.5 ns to 310 °K at 1013.25 mbar (one atmosphere) using 256 cores of an SGI Altix XE supercomputer.

Table 1

Table outlining specific details of each simulation including periodic boundary conditions (PBC) and total atoms used.

Simulation	Detached helix	Pentamer	Empty capsid	Complete virion
Total atoms	8586	388,521	3,938,950	2,754,142
PBC	Cubic	Cubic	Cubic	Rhombic dodecahedron
Amino acids	20	4390	52,680	52,680
Nucleotides (RNA)	0	0	0	7640
Fatty acids	0	10	120	120
Water molecules	2760	106,629	1,038,857	563,929
Sodium ions	10	370	3728	2581
Magnesium ions	0	0	0	3155
Chloride ions	8	309	3011	1607

Once equilibrated the virus simulation was moved to a BlueGene/P supercomputer running NAMD 2.8 beta 1 with memory optimisation extensions on 2048 cores. NPT simulations of a detached α -helix (8586 atoms), pentamer (388,521 atoms), empty capsid (3.94 million atoms) and full virion (2.75 million atoms) were run for 10 ns each (Table 1). Trajectory data for each simulation were then analysed on local computers for structural changes including root mean square deviations (RMSD) and ion permeation (refer to supplementary Fig. 1).

The fit-averaged structure of VP1 within a MD simulation of a pentamer, empty capsid and full virion were compared to the following PDB coordinate files of wild poliovirus type 1 crystal structures: 1HXS, 1AL2, 1AR6, 1AR7, 1AR8, 1AR9, 1ASJ, 1PO1, 1PO2, 1POV, 1VBD.

3. Results

Initial *in silico* reconstructions of the solvated poliovirus virions were in the order of four million atoms using cuboidal periodic boundary conditions (PBC). The adoption of a rhombic-dodecahedral PBC allowed the atom count to be reduced to 2.75 million atoms with a simulation time of 0.79 days/ns (1.27 ns/day) when using 2048 cores of a BlueGene/P supercomputer. After 10 ns of simulation each of the four models (detached α -helix, pentamer, empty capsid and full virion) were transferred locally for analysis.

The full virion structure after 10 ns of MD simulation is shown in Fig. 3. In order to quantify the stability of the simulations, the root mean square deviation (RMSD) values for α -carbons were calculated for pentamer, empty capsid and the full virion (Fig. 4) with a modest increase in RMSD indicating the simulations were stable.

The full virion required more time to stabilise compared to the empty capsid. This was evidenced also by the variation in radius of the two different capsids over the course of the 10 ns simulations (Fig. 5). The empty capsid showed no more than 2% deviation in radius compared to the full virion, which showed an initial increase in radius up until 6.5 ns at which point the virus stabilised. The final virion diameter was approximately 29.5 nm which is consistent with the observed properties of poliovirus virions *in vitro* with a diameter of approximately 30 nm [1] (Fig. 2).

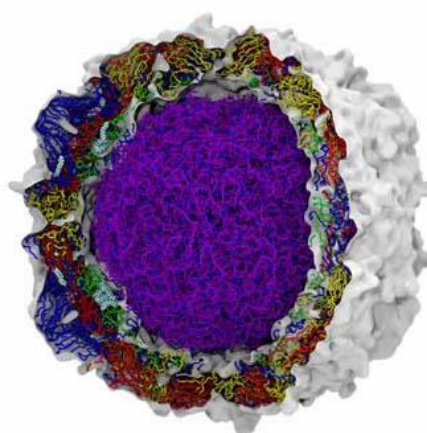


Fig. 3. Poliovirus type 1 (Mahoney strain) after 10 ns of atomistic MD simulation. Total simulation size approximately 2.8 million atoms. Water and surrounding ions have been removed from this representation. Volumetric (density) map representation is represented (white) with protein backbone shown as tubes and coloured by subunit VP1 = Blue, VP2 = Yellow, VP3 = Red, VP4 = Green. Lipids are represented as Van der Waals models with colouring by name of element. RNA is shown using a "liquorice" representation (purple) with associated magnesium ions (Blue).

We also examined the stability of VP1-4 and compared their regions of flexibility with temperature factors from known crystal structures. At the 10 ns time point, each of the 60 individual protomers constituting the virus capsid were transposed to the initial protomer coordinates as defined by the 1HXS template crystal structure. The 60 protomers now occupying the same space were then averaged to give a single "fit-averaged" structure.

The individual fit averaged protomer structures for the pentamer, empty capsid and full virion were compared to the original template coordinates and deviation from the original structure for the VP1 protein was calculated using the RMSD method. RMSD values ranged from 0.873 to 1.099 when compared to the original 1HXS crystal template structure and from

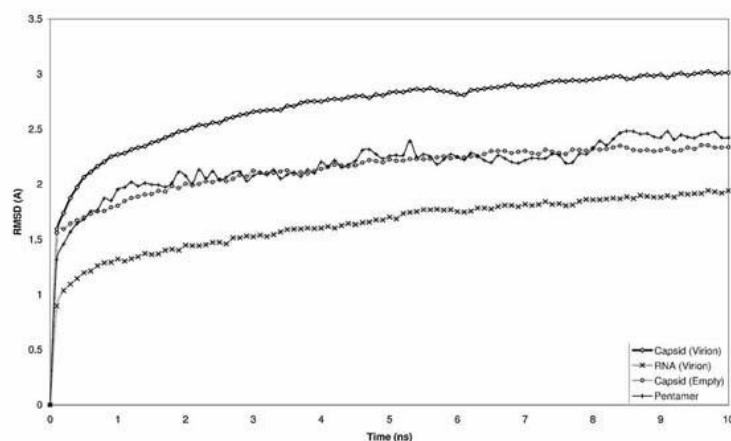


Fig. 4. Root Mean Square Deviation (RMSD) of the poliovirus empty capsid, full virion, spherical RNA alone and an individual pentamer over the 10 ns simulation time.

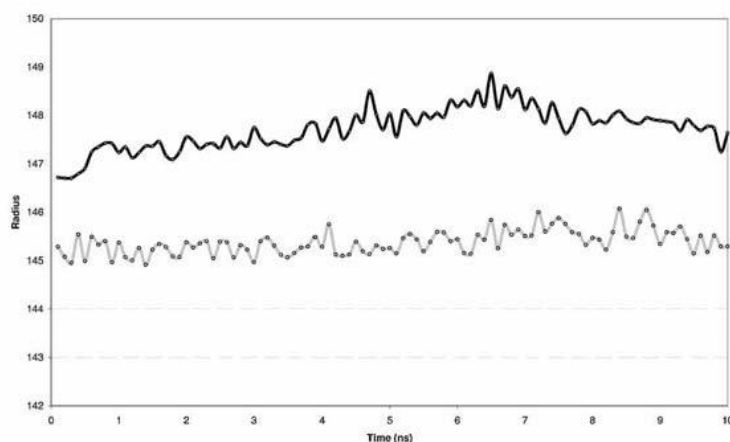


Fig. 5. Average capsid radius (Å) over 10 ns simulation time for empty capsid (grey) and full virion (black).

1.125 to 1.604 for the ten other crystal structures listed in the Methods that were available in the RCSB PDB for comparison (Fig. 6)

Root mean square fluctuation (RMSF) values were calculated for each residue of the 60 protomers for the period 7.5 ns to 10 ns. The average RMSF for each residue was then applied to a fit averaged protomer structure (Fig. 7). This method indicated the relative flexibility of individual amino acid residues. The empty capsid protomers showed significantly increased flexibility in the VP1 N-terminal α -helix and surrounding structures when compared to the full virion (Fig. 7).

The individual amino acids of the capsid proteins were analysed for their proximity to the genomic RNA. The RNA interactions were predominately confined to the amino termini of the capsid proteins VP1, VP2 and VP4 with the exception of a single instance (1 of 60) of residues VP1-267R, VP2-256M and VP2-257C (Fig. 8).

Using the protomer-fit averaging method described above, all 60 protomers were overlaid according to the original 1HXS crystal coordinate file and observed relative to the genomic RNA molecule (Fig. 9). It was noted that the N-terminal α -helix was in close contact with the RNA and partially concealed within the surface of the RNA.

The region obscured by the RNA showed a strong correlation to data missing from the 11 available poliovirus type 1 crystal

structures. Two other regions highlighted in Fig. 9, the C-terminus of VP3 and the N-terminus of VP2, which were also missing from the available crystal structures, displayed high RMSD values using the protomer-fit averaging method at the 10 ns time point (Fig. 10).

When comparing the RMSD values of the VP1 protein, particularly with respect to the original 1HXS template structure, it was shown that the α -helix displays higher RMSD values in the empty capsid than that of the full virion (Fig. 11). This is consistent with the α -helix in the empty capsid having increased molecular flexibility when RNA is not in close proximity.

The stability of the putative α -helix of the VP1 N-terminus was analysed in isolation from the remainder of the capsid structure. The detached amino acids (residues one to 20 of VP1) maintained a helical structure for the duration of the 10 ns simulation (Fig. 12a). Significant stabilising elements of this helix are salt bridges between residues ARG15-ASP11 and ARG15-GLU16 (Fig. 12b and supplementary Figs. 2–5). The distribution of amino acid residue types is indicative of an α -helix that is amphipathic by nature and matches previous descriptions (Fig. 12c) [6].

Positive ions, specifically sodium ions, were observed entering the virus core of the full virion and chloride ions were noted leaving the virus core predominately *via* small fissures that formed at the 2-fold axis of symmetry over the 10 ns MD simulation (supplementary Fig. 1). Fissures at the 2-fold axis of symmetry recently have been shown using cryo-electron microscopy to be involved in

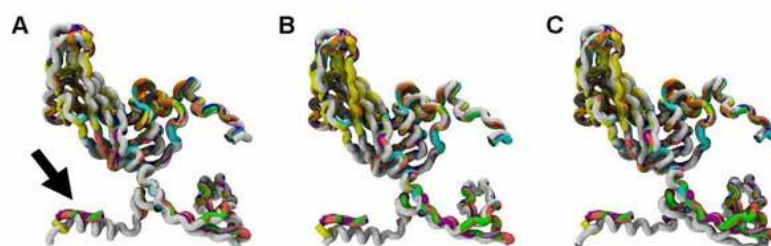


Fig. 6. View of fit-averaged VP1 structural protein (White) overlaid on ten available poliovirus type 1 crystal structure coordinates (coloured) and the original template crystal structure 1HXS (blue) from the RCSB PDB. Arrow indicates the position of the reconstructed amphipathic α -helix after 10 ns simulation of (A) pentamer ($n=5$), (B) empty capsid ($n=60$) and (C) full virion ($n=60$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

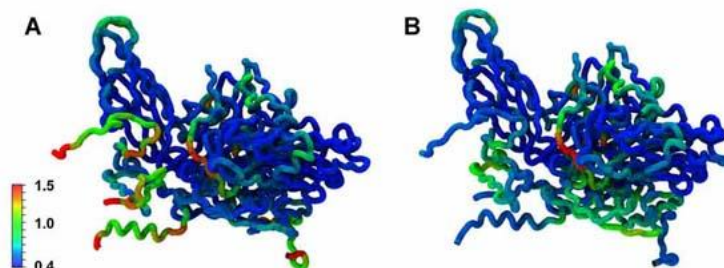


Fig. 7. Root mean square fluctuation RMSF values calculated over 7.5 to 10 ns time period for all protomers *in situ* (A) empty capsid and (B) full virion.

the release of RNA in response to heat treatment [18,19]. Magnesium ions within the virus core were not prone to any significant movement and remained intimately associated with the RNA, a phenomenon observed previously in full virus simulations of satellite tobacco mosaic virus [11]. It was observed that ion movement was not limited to the two-fold axis of symmetry but also occurred via a small pore-like structure at the base of the canyon near the entrance to the hydrophobic pocket, this pore-like structure has been observed previously in viruses related to poliovirus such as rhinovirus [20].

4. Discussion

We are investigating MD simulation methodology to build complete human enteroviruses, including the native RNA sequence.

We believe that all-atom simulations of enteroviruses will provide insights to significant molecular interactions and biological properties of the virion that cannot be discerned from the modelling of individual structural units or the use of static models. The advantage of full virus simulations is weighed against the increase in the total atom count, which may limit the simulation time. This was addressed by adopting a rhombic dodecahedral formation for the solvated full virus simulation with a one third reduction in the total atom count.

Initially we aimed to validate the atomistic MD simulation of a complete enterovirus by constructing wild poliovirus type 1 (Mahoney strain) and comparing various structural elements with published data. We chose this virus due to the abundance of X-ray crystallography and biochemical analyses published over more than thirty years. We noted that specific regions of the

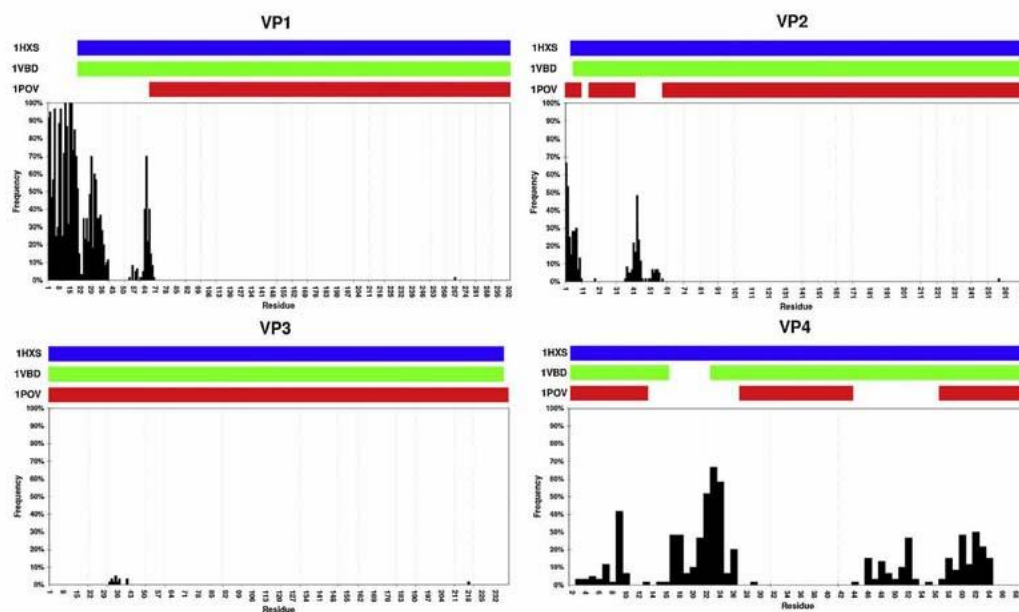


Fig. 8. Bar graph depicting the frequency of amino acids found within a distance of 5 Å from the viral RNA for all 60 capsid protomers. Shaded bars at the top of each chart indicate available coordinate data by residue for the full virion template files 1HXS (blue) and 1VBD (green) and empty capsid 1POV (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

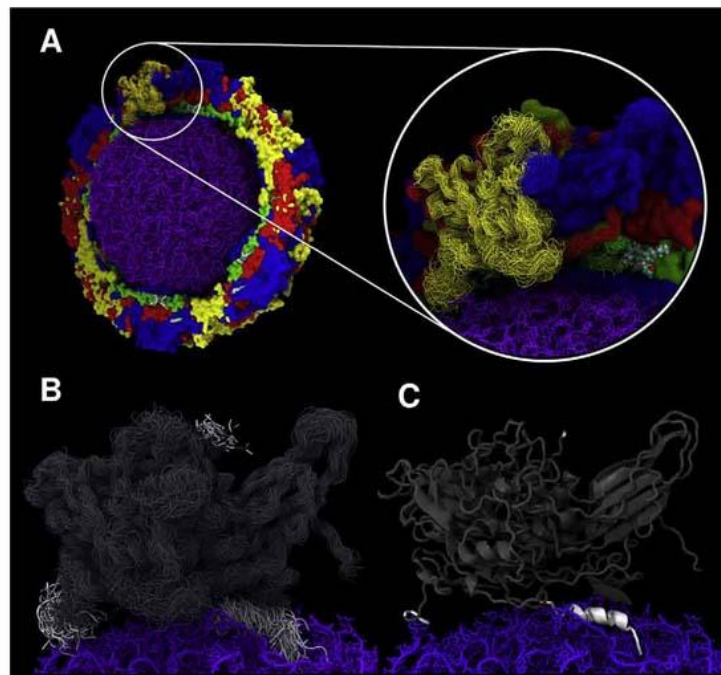


Fig. 9. (A) cutaway representation of the full poliovirus virion highlighting relative positioning of a protomer in relation to the RNA genome (purple). The fit averaged protomer ($n = 60$) is coloured by chain as in previous figures and represented as a backbone structure at the 10 ns time point. (B) Overlaid backbone structure of all 60 fit averaged protomers, structures missing from original X-ray crystal data are highlighted in white. (C) Cartoon representation of the structure derived from the fit averaging procedure. Missing crystal data are highlighted in white.

virus internal structure could not be resolved by X-ray crystallography and wanted to investigate whether this could be due to interaction of the genomic RNA with the capsid proteins by a simulation of the full poliovirus virion. Although the virus capsids display an icosahedral symmetry the structure of underlying genomic RNA has not been determined due to the amorphous nature of the virus core. We thought it important to incorporate the RNA in this structure to better reproduce an infectious particle. An additional advantage of simulating the entire virion is that we also avoid the potential for any simulation artefacts that may arise by simulating a portion of the capsid through symmetry. This could be especially relevant if there is an event that breaks capsid symmetry such as capsid distortion prior to RNA release.

The availability of 60 individual hydrophobic pockets representing 60 independent measurement points will prove useful in terms of statistical and biological relevance for future experiments involving antiviral drug-binding events. The core of the fivefold axis of symmetry displayed variation in flexibility in the presence of RNA when compared to an empty capsid as shown in Fig. 7. It is not known what influence this variation in flexibility would have on the study of drug binding in the hydrophobic pocket. Inclusion of the encoding RNA molecule may also serve to assist researchers studying viral uncoating mechanisms and subsequent RNA release.

The full virus and the individual components (pentamer, empty capsid, spherical RNA) were stable after 10 ns simulation when

assessed by RMSD and the average capsid radius of the empty capsid and full virion. The N-terminus of the VP1 protein of poliovirus type 1 has been hypothesised to form an amphipathic α -helix based on the periodicity of the amino acid sequence [5]. The putative amphipathic helical structure was intrinsically stable when the final 20 amino acid residues were simulated as an independent entity, however, a fit averaged structure of the 60 protomer units indicated increased flexibility of the N-terminal region of VP1 when it was part of the empty capsid structure compared to the full virion. At completion of the full virion simulation, the N-terminus of VP1 protruded internally of the capsid in close proximity to the genomic RNA, an orientation that would occlude it from X-ray analysis.

The N-terminus of VP1 is reversibly displayed on the exterior of the virus capsid through virus breathing, a feat that would necessitate molecular flexibility. It was postulated that binding of poliovirus to the cellular receptor, CD155, results in the irreversible externalization of the N-terminus of VP1 as part of cell membrane penetration to form a pore complex for egress of the genomic RNA [5,8]. The orientation of the N-terminus of VP1 on the internal surface of the capsid prior to receptor binding has been the subject of much speculation [5]. The amphipathic helix of the N-terminus of VP1 of human rhinovirus 16 was reported to be located on the interior surface of the capsid along the icosahedral fivefold axis, the orientation described in this study [6]. Recent published data of enterovirus 71 indicated that the positioning of the VP1 α -helix was oriented toward the 2 fold axis of symmetry [21]. In addition, antibodies directed to the poliovirus N-terminus of VP1 were

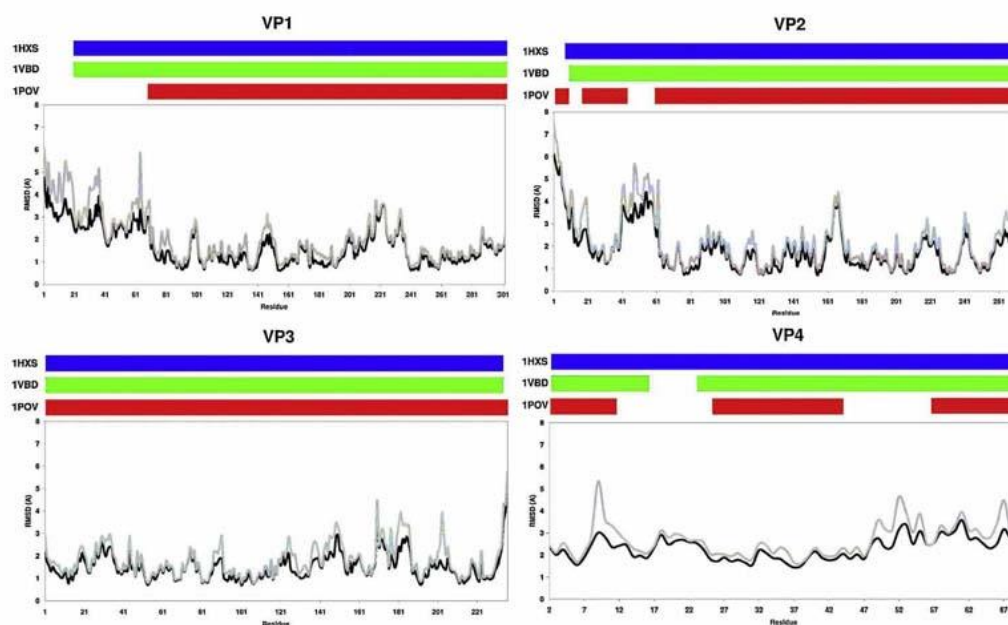


Fig. 10. RMSD of individual residue for fit-averaged protomers ($n = 60$) at the 10 ns time point for empty capsid (grey) and full virion (black). Shaded bars at the top of each chart indicate available coordinate data by residue for the full virion template files 1HXS (blue) and 1VBD (green) and empty capsid 1POV (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

detected by cryo-EM analysis to be orientated towards the two-fold axis of symmetry [22]. Further studies of this process using the full poliovirus MD model described in this report as a basis will be performed.

Our results show that the detached helix is just as stable as those in full capsids (supplementary Figs. 6–9). The mechanistic implication is that unlike many toxins and viral fusion peptides, the VP1 N-terminal helix may not be required to unfold and refold upon membrane insertion, but rather exists in a pre-formed helix that can readily insert into the membrane [23]. The presence of conserved glycine and proline motifs within the N-terminus of the VP1 protein may act as molecular hinges facilitating the required

flexibility for expulsion VP1 during the transition from 160S to 135S conformations [8,24,25].

The production of a stable all-atom MD simulation of wild poliovirus that has been validated against crystallographic data lays the foundation for the modelling of other enteroviruses and picornaviruses or other similar viruses exhibiting icosahedral symmetry. The recent availability of increased computational power should allow significant increases in simulation times approaching microseconds and enhance future simulations involving the repositioning of the VP1 toward the two-fold axis of symmetry and the investigation of RNA models folded using predicted stem loop structures.

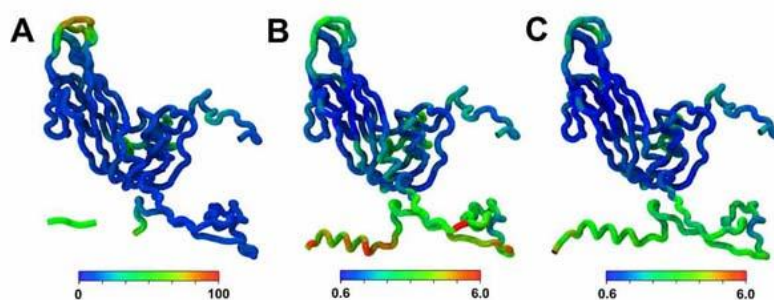


Fig. 11. (A) VP1 protein of the 1HXS template crystal structure coloured by beta value. Figures B and C are coloured "RMSD by residue" of fit averaged VP1 structural proteins ($n = 60$) at 10 ns time point (B) empty capsid and (C) full virion.

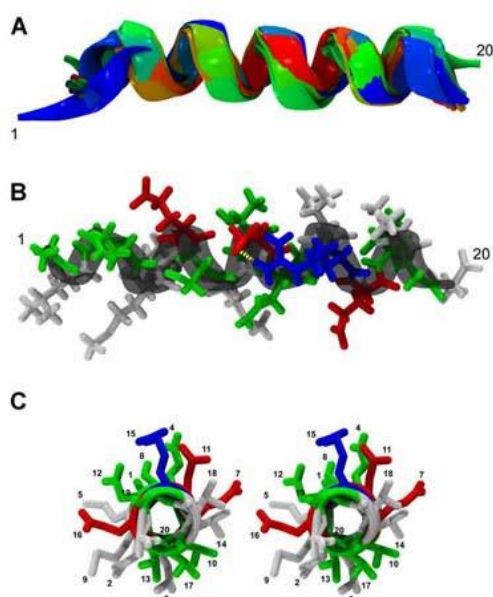


Fig. 12. Detached α -helix from the N-terminus of VP1 after 10 ns simulation. (A) Conformational changes captured at 1 ns time points, (1 ns = blue, 10 ns = red). (B) Licorice representation of amino acid side chains, showing salt bridge between ARG and ASP (centre), (hydrophobic = white, polar = green, basic = blue, acidic = red, yellow = Hydrogen-bond). (C) Stereo representation of side chain charge distribution around the amphipathic helix.

Acknowledgements

We would like to thank Dallas Wilson, Thomas Aiken and Mike Catton of the Victorian Infectious Diseases Reference Laboratory for supportive contributions at various stages of this work. We thank Cyril Reboul, Grischa Meyer and Matthew Downton for their kind assistance with the generation of the rhombic dodecahedral modelling conditions. We thank the VIDRL Electron Microscopy Unit for their kind provision of the picornavirus transmission electron microscopy image.

We would like to acknowledge and thank nVidia™ for providing access to a Tesla C2070 GPU for the purposes of this study.

This research was supported by a Victorian Life Sciences Computation Initiative (VLSI) grant number VR0069 on its Peak Computing Facility at the University of Melbourne, an initiative of the Victorian Government, Australia.

The National Enterovirus Laboratory and WHO Regional Poliomyelitis Reference Laboratory are supported by funding from the Victorian State Government, the Department of Health and Ageing, Australia and the World Health Organization.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jmgl.2012.06.009>.

References

- [1] N.J. Knowles, et al., Family Picornaviridae, in: A.M.Q. King, et al. (Eds.), *Virus taxonomy*, Ninth report of the International Committee on Taxonomy of Viruses, Academic Press, 2012, pp. 855–880.
- [2] J. Hogle, M. Chow, D. Filman, Three-dimensional structure of poliovirus at 2.9 Å resolution, *Science* 229 (1985) 1358.
- [3] R. Basavappa, D. Filman, R. Syed, O. Flore, J. Icenogle, J. Hogle, Role and mechanism of the maturation cleavage of VP0 in poliovirus assembly: structure of the empty capsid assembly intermediate at 2.9 Å resolution, *Protein Science* 3 (1994) 1651–1669.
- [4] D. Baltimore, Picornaviruses are no longer black boxes, *Science* 229 (1985) 1366.
- [5] J.M. Hogle, Poliovirus cell entry: common structural themes in viral cell entry pathways, *Annual Review of Microbiology* 56 (2002) 677.
- [6] A.T. Hadfield, W. Lee, R. Zhao, M.A. Oliveira, L. Minor, R.R. Rueckert, et al., The refined structure of human rhinovirus 16 at 2.15 Å resolution: implications for the viral life cycle, *Structure* 5 (1997) 427–441.
- [7] C.E. Fricks, J.M. Hogle, Cell-induced conformational change in poliovirus: externalization of the amino terminus of VP1 is responsible for liposome binding, *Journal of Virology* 64 (1990) 1934–1945.
- [8] D. Bubeck, D.J. Filman, N. Cheng, A.C. Steven, J.M. Hogle, D.M. Belnap, The structure of the poliovirus 135S cell entry intermediate at 10 Å resolution reveals the location of an externalized polypeptide that binds to membranes, *Journal of Virology* 79 (2005) 7745–7755.
- [9] S. Curry, M. Chow, J.M. Hogle, The poliovirus 135S particle is infectious, *Journal of Virology* 70 (1996) 7125–7131.
- [10] Y. Huang, J.M. Hogle, M. Chow, Is the 135S poliovirus particle an intermediate during cell entry, *Journal of Virology* 74 (2000) 8757–8761.
- [11] P.L. Freddolino, A.S. Arkhipov, S.B. Larson, A. McPherson, K. Schulten, Molecular dynamics simulations of the complete satellite tobacco mosaic virus, *Structure* 14 (2006) 437–449.
- [12] S.T. Miller, J.M. Hogle, D.J. Filman, Ab initio phasing of high-symmetry macromolecular complexes: successful phasing of authentic poliovirus data to 3.0 Å resolution, *Journal of Molecular Biology* 307 (2001) 499–512.
- [13] L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, *Bioinformatics* 16 (2000) 404.
- [14] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *Journal of Molecular Graphics* 14 (1996) 33–38.
- [15] J.E. Stone, J.C. Phillips, P.L. Freddolino, D.J. Hardy, L.G. Trabuco, et al., Accelerating molecular modeling applications with graphics processors, *Journal of Computational Chemistry* 28 (2007) 2618–2640.
- [16] V. Zoete, M.A. Cuendet, A. Grosdidier, O. Michielin, SwissParam: a fast force field generation tool for small organic molecules, *Journal of Computational Chemistry* (2011).
- [17] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, et al., Scalable molecular dynamics with NAMD, *Journal of Computational Chemistry* 26 (2005) 1781.
- [18] H.C. Levy, M. Bostina, D.J. Filman, J.M. Hogle, Catching a virus in the act of RNA release: a novel poliovirus uncoating intermediate characterized by cryo-electron microscopy, *Journal of Virology* 84 (2010) 4426.
- [19] M. Bostina, H. Levy, D.J. Filman, J.M. Hogle, Poliovirus RNA is released from the capsid near a twofold symmetry axis, *Journal of Virology* 85 (2011) 776.
- [20] D. Garriga, A. Pickl-Herk, D. Luque, J. Wruss, J.R. Castón, D. Blas, et al., Insights into minor group rhinovirus uncoating: the X-ray structure of the HRV2 empty capsid, *PLoS Pathogens* 8 (2012) e1002473.
- [21] X. Wang, W. Peng, J. Ren, Z. Hu, J. Xu, Z. Lou, et al., A sensor-adaptor mechanism for enterovirus uncoating from structures of EV71, *Nature Structural & Molecular Biology* (2012).
- [22] J. Lin, L. Lee, M. Roivainen, D.J. Filman, J.M. Hogle, D.M. Belnap, Structure of the fab-labeled ‘breathing’ state of native poliovirus, *Journal of Virology* (2012).
- [23] G.S. Hong, C.P. Chen, M.H. Lin, J. Krüger, C.F.W. Becker, R.H.A. Fink, W.B. Fischer, Molecular dynamics simulations and conductance studies of the interaction of VP1 N-terminus from poliovirus and gp41 fusion peptide from HIV-1 with lipid membranes, *Molecular Membrane Biology* (2012) 1–17.
- [24] S.P. Sansom, M. Weinstein, H. Hinges, swivels and switches: the role of prolines in signalling via transmembrane α -helices, *Trends in Pharmacological Sciences* 21 (2000) 445–451.
- [25] S. Trabulo, A.L. Cardoso, M. Mano, M.C.P. De Lima, Cell-penetrating peptides—mechanisms of cellular uptake and generation of delivery systems, *Pharmaceuticals* 3 (2010) 961–993.
- [26] D. Frishman, P. Argos, Knowledge-based protein secondary structure assignment, *Proteins: Structure, Function, and Bioinformatics* 23 (1995) 566–579.



Characterization of Poliovirus Variants Selected for Resistance to the Antiviral Compound V-073

Hong-Mei Liu,^a Jason A. Roberts,^b Deborah Moore,^a Barbara Anderson,^a Mark A. Pallansch,^a Daniel C. Pevear,^c Marc S. Collett,^d and M. Steven Oberste^a

Division of Viral Diseases, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, USA^a; Victorian Infectious Diseases Reference Laboratory, Melbourne, Australia^b; VenatoRx Pharmaceuticals, Inc., Malvern, Pennsylvania, USA^c; and ViroDefense, Inc., Rockville, Maryland, USA^d

V-073, a small-molecule capsid inhibitor originally developed for nonpolio enterovirus indications is considerably more potent against polioviruses. All poliovirus isolates tested to date ($n = 45$), including wild, vaccine, vaccine-derived, and laboratory strains, are susceptible to the antiviral capsid inhibitor V-073. We grew poliovirus in the presence of V-073 to allow for the identification of variants with reduced susceptibility to the drug. Sequence analysis of 160 independent resistant variants (80 isolates of poliovirus type 1, 40 isolates each of types 2 and 3) established that V-073 resistance involved a single amino acid change in either of two virus capsid proteins, VP1 (67 of 160 [42%]) or VP3 (93 of 160 [58%]). In resistant variants with a VP1 change, the majority (53 of 67 [79%]) exhibited a substitution of isoleucine at position 194 (equivalent position 192 in type 3) with either methionine or phenylalanine. Of those with a VP3 change, alanine at position 24 was replaced with valine in all variants ($n = 93$). The resistance phenotype was relatively stable upon passage of viruses in cell culture in the absence of drug. Single-step growth studies showed no substantial differences between drug-resistant variants and the virus stocks from which they were derived, while the resistant viruses were generally more thermally labile than the corresponding drug-susceptible parental viruses. These studies provide a foundation from which to build a greater understanding of resistance to antiviral compound V-073.

In 1988, the World Health Organization (WHO) launched the Global Polio Eradication Initiative (GPEI). The Initiative has relied exclusively on the oral polio vaccine (OPV), an inexpensive and easily administered live, attenuated vaccine. OPV is generally safe and has been highly effective under most circumstances. However, at a low frequency, about 1 per 750,000 vaccinees, OPV itself can cause paralysis (vaccine-associated paralytic poliomyelitis [VAPP]) (1). Normally, OPV viruses are excreted in the stool of healthy vaccinated individuals for several weeks. Should these excreted vaccine-related polioviruses continue to circulate (termed circulating vaccine-derived polioviruses [cVDPVs]), virus reversion to neurovirulence can occur and result in paralytic disease and outbreaks. Moreover, when individuals with a primary immune deficiency, such as agammaglobulinemia, receive OPV, the virus may replicate persistently, accumulating genetic changes associated with reversion to neurovirulence. The resulting immunodeficiency-associated VDPVs can be excreted for years, posing a risk of paralysis for the infected individual and compromising efforts to eradicate the virus. As wild poliovirus transmission is eliminated, and VAPP and VDPV cases continue to occur, the risks of OPV use will outweigh its benefits. Thus, part of the GPEI strategy involves the global cessation of OPV use when it is determined wild poliovirus transmission has ceased (2).

After reviewing current and post-OPV risks, the National Research Council of the National Academies concluded that at least one, preferably two, polio antiviral drugs be developed as a supplement to the tools currently available for control of polio outbreaks posteradication (4, 5).

V-073, a small-molecule capsid inhibitor originally developed for nonpolio enterovirus indications (3), was recently found to be considerably more potent against polioviruses (14). V-073 is currently being advanced for possible use in the management of po-

liovirus chronic infections and laboratory incidents. While V-073 is active against all poliovirus isolates tested to date (14), due to the quasispecies nature of RNA viruses, there likely exist drug-resistant variants at low levels in drug-susceptible virus populations. Treatment-emergent drug resistance could potentially have medical and public health implications. On the other hand, drug-resistant virus variants may be enfeebled or otherwise more benign. Thus, it is important to understand drug resistance and its potential consequences.

Here, we describe the selection and isolation of V-073-resistant poliovirus variants in cell culture, the genetic basis for resistance, and several features of the variants, including cell culture growth, phenotypic stability in cell culture, and thermal stability of virion infectivity. The antigenic, immunogenic, and pathogenic features of V-073-resistant polioviruses are reported elsewhere (9).

MATERIALS AND METHODS

Viruses. Type 1 VDPV isolates 10235 (cVDPV-1 isolated from the Dominican Republic, GenBank accession no. AF405625) and 10224 (cVDPV-1 isolated from the United States), type 2 isolate 10230 (cVDPV-2 isolated from Egypt, GenBank accession no. AF448783), and type 3 isolate 10805 (cVDPV-3 isolated from Iran, GenBank accession no. EU684056) were obtained from the Centers for Disease Control and Prevention (CDC) Polio and Picornavirus Laboratory collection and propa-

Received 9 March 2012. Returned for modification 16 April 2012.

Accepted 7 August 2012.

Published ahead of print 13 August 2012.

Address correspondence to M. Steven Oberste, soberste@cdc.gov.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AAC.00539-12

gated on LLC-MK₂ cells at 37°C (types 1 and 2) or 36°C (type 3) in minimal essential medium (MEM) with Earle's salts (Invitrogen, Carlsbad, CA), supplemented with 2% fetal bovine serum (FBS; Atlas Biologicals, Fort Collins, CO, or Thermo Scientific, Lafayette, CO).

Plaque assays. LLC-MK₂ monolayer cells were grown in six-well cell culture plates and washed with MEM (Invitrogen, Grand Island, NY). Tenfold serial dilutions of each virus were prepared in MEM, and 200 μ l was inoculated into each well. After infection at room temperature for 45 to 60 min, 2.5 ml of overlay solution was added, consisting of 0.45% agarose (SeaKem LE agarose; Lonza, Rockland, ME) in MEM with 2% FBS. After the agar overlay had gelled, the plates were incubated at 36°C or 37°C, as described above, for 36 to 46 h. The overlay was removed, and the cells were stained with a solution of 0.4% formaldehyde and 0.03% crystal violet for 30 min. The plates were then washed once with water and allowed to dry overnight. Plaques were counted, and virus titer was determined.

Selection of poliovirus variants with reduced susceptibility to V-073. For each virus, 10 independent substrains were obtained by plaque purification. The substrains were amplified in LLC-MK₂ cells in MEM with 2% FBS. For drug-resistant variant selection, serial dilutions of each plaque-purified virus were inoculated into LLC-MK₂ cells in six-well plates and covered under agar overlay with either no drug, drug at a final concentration of 10 times the median effective concentration (EC_{50}) for that particular virus ($10 \times EC_{50}$), or drug at 50 times the EC_{50} ($50 \times EC_{50}$). Drug (V-073) was provided by ViroDefense, Inc. The susceptibility of these strains to V-073 was reported recently (14). For each parental virus, two plaques from each selection condition were selected and propagated in LLC-MK₂ cells under the selection condition. Drug resistance frequency was calculated by dividing the titer of the surviving plaques in the presence of $50 \times EC_{50}$ V-073 by the titer of parental plaque-purified virus in the absence of drug.

Drug susceptibility assay. Susceptibility of parental and drug-resistant variants to V-073 and determination of the EC_{50} s were performed as described previously (15). Briefly, drug and virus were combined with LLC-MK₂ cells in 96-well plates in a cross-titration format to ensure reaching endpoints for both drug and virus titrations, with duplicate wells for each drug-virus concentration. After 3 days of incubation at 37°C, the plates were stained with crystal violet, washed three times with water, and allowed to dry overnight. Viral cytopathic effect was measured by reading the absorbance at 590 nm. EC_{50} s were derived by analyzing dose-response absorbance values by four-parameter curve fitting using Prism 5.04 (GraphPad Software, Inc., La Jolla, CA).

Identification of amino acid changes in V-073-resistant variants. Viral RNA was extracted and purified by using the QiaAmp viral RNA mini-kit (Qiagen, Inc., Valencia, CA). The VP3 and VP1 capsid genes were amplified by reverse transcription-PCR. PCR products were sequenced with the PCR primers and additional primers within each amplicon. The sequences of PCR and sequencing primers used are available upon request. Sequences were assembled using Sequencher (version 4.8; Gene Codes Corp., Ann Arbor, MI) and analyzed using the Wisconsin Sequence Analysis Package, version 11.0 (Accelrys, San Diego, CA). The sequences were submitted to the GenBank sequence database under accession numbers JN105289 to JN105295.

Stability of V-073-resistant phenotype. V-073-resistant viruses representing the two predominant variant classes were subjected to 10 cycles of passage in LLC-MK₂ cell cultures in the absence of drug. Viruses were initially inoculated onto 10^5 LLC-MK₂ cells in 24-well plates at a multiplicity of infection (MOI) of 0.01 and then passaged every 3 days by inoculating 50 μ l of a 10^{-4} virus dilution onto 10^5 fresh LLC-MK₂ cells in 0.5 ml of MEM supplemented with 2% FBS. At the tenth passage, the EC_{50} of the virus populations was determined. The distribution of resistant and susceptible viruses was also determined by plaque assay in the presence or absence of V-073 at $10 \times$ the respective EC_{50} s, as described above.

Single-step growth curves. Single-step growth of the parental and drug-resistant variants was carried out in LLC-MK₂ cell monolayers in

24-well plates. Cells were washed with MEM and infected at an MOI of 10 with the test virus in replicate plates. After the plates were incubated for 45 min at room temperature, the inoculum was removed, and the cells were washed with MEM without serum, and then 0.5 ml of MEM supplemented with 2% FBS was added (in the absence of V-073), and the plates were incubated at 37°C in an atmosphere of 5% CO₂. At 0, 2, 4, 6, 8, and 10 h postinoculation, a plate was frozen at -70°C . Plates were freeze-thawed three times, and the virus titers in the supernatants were determined by plaque assay in the absence of V-073.

Virion thermal stability. The kinetics of heat inactivation was measured by incubating variant and parental viruses at 46°C for 0, 5, 15, 30, or 60 min. The titers of the treated viruses were determined by plaque assay in the absence of V-073. For each time point, stability at 46°C was expressed as a ratio relative to the titer of virus obtained at 0 min.

Poliovirus sequence analyses. VP1 and VP3 sequences of the drug-resistant variants were compared to a database of sequences composed of circulating wild and vaccine-derived polioviruses using a customized search program implemented in MATLAB R2010a (The MathWorks, Natick, MA).

Structural modeling. Protomer homology models representing each parental VDPV strain were created using the SWISS-model server (<http://swissmodel.expasy.org/>) (17) using the template PDB coordinate file 1EAH as a reference. Point mutations for each subsequent protomer were created using the "Mutator" plug-in associated with the Visual Molecular Dynamics (VMD) software package (8). Custom parameter data for the associated V-073 molecule and covalently bound myristic acid were obtained using the SWISS-param server (<http://swissparam.ch/>) (18).

Each protomer was assembled into a pentamer configuration using the biological assembly data contained in the 1EAH crystal structure metadata file. Pentamers were then solvated using a TIP3 water model in a cuboidal system with 12-Å padding on each axis and ionized with sodium and chloride ions at a concentration representing 0.154 M. Energy minimization and molecular dynamics simulation for 1 nanosecond at 37°C and one atmosphere pressure, was performed using the Nanoscale Molecular Dynamics (NAMD) software package (16) on an SGI Altix supercomputer housed at the Victorian Life Sciences Computation Initiative, Melbourne, Australia. Postsimulation pentamers were deconstructed into five individual protomers and fit-aligned to the original 1EAH coordinates and a final average protomer structure obtained. Simulation data analysis and image generation was achieved using VMD.

RESULTS

Isolation of V-073-resistant poliovirus variants from susceptible virus populations. Previously, we showed that 45 distinct poliovirus isolates, including wild, vaccine, vaccine-derived, and laboratory strains of poliovirus, were all susceptible to the capsid inhibitor V-073, with an MIC for 90% of isolates (MIC_{90}) of 0.076 μM (14). However, due to the quasispecies nature of RNA viruses, there likely exist drug-resistant variants at low levels in drug-susceptible virus populations. Indeed, poliovirus variants with reduced susceptibility to V-073 were isolated from drug-susceptible virus populations by virus growth in cell culture in the presence of V-073. Ten independent wild-type (parental) virus plaques were picked for each of four polioviruses, two VDPV type 1 isolates (strains 10224 and 10235), one VDPV type 2 isolate (strain 10230), and one VDPV type 3 isolate (strain 10805). The plaques were each amplified by a single passage in LLC-MK₂ cells to allow generation of a quasispecies population ("pool") from which drug-resistant variants could be selected. Viruses from each parental pool were propagated in LLC-MK₂ cells in the absence or in the presence of V-073 at either $10 \times$ or $50 \times$ the EC_{50} for the respective parental virus. Surviving plaques were quantified and their titers compared to that of their parental virus. Under these

TABLE 1 Frequency of V-073-resistant variants in drug-susceptible poliovirus populations

Virus	Parent EC ₅₀ (μ M)	Resistance frequency $\times 10^{-5}$ (95% CI) ^a	
		Selection at 10 \times EC ₅₀	Selection at 50 \times EC ₅₀
VDPV-1 10224	0.069	30.4 (24.0–38.4)	5.61 (4.25–7.41)
VDPV-1 10235	0.018	3.20 (2.24–4.57)	3.31 (2.50–4.37)
VDPV-2 10230	0.036	26.5 (18.2–38.7)	11.8 (7.48–18.6)
VDPV-3 10805	0.029	42.7 (18.2–100)	30.8 (15.6–61.0)

^a 95% CI, 95% confidence interval.

selection conditions, the frequency of V-073-resistant variants in populations of the polioviruses defined by selection at 10 \times the parental virus EC₅₀ ranged from 3.20 to 42.7 $\times 10^{-5}$ (Table 1). Similar frequencies were observed when selection was conducted at 50 \times the EC₅₀ (3.31 to 30.4 $\times 10^{-5}$).

Genetic basis of resistance. To determine the genetic basis of V-073 resistance, two surviving plaques from each of the parental pools (20 drug-resistant variants from each selection condition [10 \times and 50 \times EC₅₀] derived from each of four parental viruses), 160 variants in all, were isolated. Initially, the complete capsid region of two drug-resistant variants from type 1, type 2, and type 3 VDPVs were sequenced and compared to their respective parental virus sequence. Amino acid changes in the variants that correlated with the resistance phenotype were restricted to capsid protein coding regions for VP3 and VP1. This is consistent with the observation that VP3 and VP1 are the only proteins in contact with the drug when bound to the virus (10). Therefore, we subsequently sequenced the VP3 and VP1 coding regions of all variants and their parental virus.

For each independently selected resistant virus, a single amino acid change relative to its parental virus sequence in either VP1 or

VP3 was associated with the resistant phenotype (Table 2). Two amino acid positions were predominantly associated with drug resistance and together represented 91% (146 of 160) of all resistant variants. In VP1, the isoleucine at position 194 (equivalent position 192 in type 3 polioviruses) was replaced by either a methionine or phenylalanine in 53 of 160 (33%) of the resistant variants analyzed, with an apparent preference for phenylalanine (38 of 54 [70%]). In VP3, the alanine at position 24 was replaced exclusively with valine in 93 of 160 (58%) of the resistant viruses. Two minor resistant variants were also observed, each representing 4% (6 of 160) of the total viruses analyzed, one found exclusively in the type 1 virus 10224, in which the isoleucine at VP1 position 183 was replaced with threonine, and the other, from the type 3 virus 10805, in which the phenylalanine at VP1 position 237 (position 236 in type 3) was replaced with leucine. The amino acid change at VP1 position 237 was also found in 6 out of 40 resistant variants of another type 2 VDPV (data not shown). There appeared to be no consistent amino acid substitution bias associated with the two drug selection concentrations.

Stability of the resistance phenotype. For further characterization of viruses resistant to V-073, we focused on variants exhibiting the two predominant amino acid substitutions, VP1 I194M/F and VP3 A24V. One representative of each variant for each of the three poliovirus serotypes (6 resistant variants in all), together with their respective drug-susceptible parents, were characterized. The susceptibilities to V-073 of the parental viruses and their resistant variants, expressed as EC₅₀s, are provided in Table 3. The reduction in susceptibility of the variants relative to their parents ranged from 38-fold to >556-fold. To ascertain the stability of the resistance phenotype, the six resistant variants were cultured in the absence of V-073. After 10 passages in the absence of drug selection pressure, the proportion of the virus population

TABLE 2 Amino acid changes associated with resistance to V-073

Virus	EC ₅₀	Amino acid change (no. of strains) ^a						No. of viruses analyzed
		VP1			VP3			
		A, 88	P, 161	I, 183	L, 194 ^b	F, 237 ^c	A, 24	
VDPV-1 10224	10 \times		S (2)	T (6)	M (1)		V (8)	20
	50 \times	T (1) ^d			F (3)		V (19)	20
VDPV-1 10235	10 \times				F (9)		V (11)	20
	50 \times				F (11)		V (8)	20
VDPV-2 10230	10 \times				M (11)		V (9)	20
	50 \times				M (2)		V (18)	20
VDPV-3 10805	10 \times				F (3)	L (4)	V (13)	20
	50 \times				F (12)	L (2)	V (6)	20
Totals	10 \times (%)		2 (2)	6 (8)	27 (34)	4 (5)	41 (51)	80
	50 \times (%)	1 (1)	0 (0)	0 (0)	26 (32)	2 (2)	52 (65)	80
	Total (%)	1 (1)	2 (1)	6 (4)	53 (33)	6 (4)	93 (58)	160

^a Values for VP1 and VP3 are expressed as amino acid changes (number of changes) except as noted for the totals in column 1. Column subheadings indicate the "wild amino acid, residue number."

^b Equivalent residue 192 in type 3.

^c Equivalent residue 236 in type 3.

^d One virus had changes at both residues, VP1₁₈₃ and VP3₂₄.

TABLE 3 Stability of the V-073 resistance phenotype upon cell culture passage^a

Virus	Amino acid change	EC ₅₀ (μM)		Passage 10 (% resistant) ^b
		Passage 1	Passage 10	
VDPV-1 10235 (parent)		0.018	NA ^c	NA
10235.1	VP1 I194F	>10	>10	82
10235.3	VP3 A24V	1.4	0.95	69
VDPV-2 10230 (parent)		0.036	NA	NA
10230.4	VP1 I194 M	1.5	3.0	43
10230.8	VP3 A24V	>10	0.31	93
VDPV-3 10805 (parent)		0.029	NA	NA
10805.1	VP1 I192F	>10	>10	97
10805.5	VP3 A24V	1.1	1.0	37

^a Passage was performed in the absence of drug selection.

^b Resistance was determined at 10× the parental EC₅₀.

^c NA, not applicable.

that maintained the resistance phenotype was determined by plaque assay in both the absence and presence of V-073 at 10× the EC₅₀ for the respective virus. As shown in Table 3, after 10 passages in the absence of drug selection pressure, the percentage of the virus population retaining the resistant phenotype ranged from 100% (strains 10230.8 and 10805.1) to 37% (strain 10805.5), suggesting that the resistant phenotype was generally stable in cell culture and not strongly selected against. The EC₅₀s for the V-073-resistant viruses in the passage-10 virus populations were determined and found to be largely the same as the passage 0 virus for five of the six resistant viruses studied (Table 3). However, in the strain 10230.8 passage-10 population, the EC₅₀ was reduced from >10 μM to 0.31 μM, suggesting a recovery of drug susceptibility when propagated in the absence of drug selection, but it was still 10-fold less susceptible than its parental virus.

Virus growth in cell culture. To assess the virological consequences of drug resistance, studies were performed in which the resistant variant viruses were compared to their respective drug-susceptible parental virus population. In the first of these,

a single step growth curve in cell culture of the predominant VP1 and VP3 variants in the absence of drug was conducted in parallel with the respective parental virus. There appeared to be no substantial difference in the single step growth curve profiles for any of the resistant variants relative to their respective parental virus, using an MOI of 10 (Fig. 1). All variants replicated to approximately the same peak titer and with approximately the same kinetics as their respective parental strain, indicating there was no major deficit in viral replicative capacity in cell culture. Similar results were also observed at MOI of 1 or 0.1 (data not shown).

Thermal stability of virus infectivity. It has been reported previously that drug resistance to enterovirus capsid inhibitors results in a more thermally labile virion (12). Thermolability is thought to be due to the presence of bulkier amino acids in the pocket, impacting the binding of natural factors that bind into the pocket and stabilize the virion in the extracellular space. To investigate whether the changes necessary to confer V-073 resistance affect virion stability, V-073-resistant variants were heated in culture medium without FBS at 46°C for various periods of time, after which the surviving poliovirus infectivity was determined. Four of the six resistant variants were substantially more heat labile than their respective parental virus, while two variants (10235.1 and 10230.4) showed only a slight reduction in thermal stability (Fig. 2).

Structural modeling of drug-resistant variant viruses complexed with V-073. To better understand the mechanism of V-073 resistance, we modeled the amino acid changes responsible for resistance in all three polio serotypes, based on the known crystal structure of V-073 complexed with poliovirus type 2 (Fig. 3). The isoleucine at VP1_{194(192)}} and alanine at VP3_{24}} both have hydrophobic interactions with the dichlorophenyl portion of the drug. The observed substitutions in VP1_{194(192)}} and VP3_{24}} disrupt this interaction and cause deformation of the drug-binding pocket. The amino acid residues of the minor variants (Table 2) are also predicted to impinge on the drug-binding pocket (data not shown). The phenyl group of Phe_{236}} residue in VP1 virus appears to interact with the chloro-methoxyphenyl portion of the drug, and this interaction is ablated in the type 3 virus with Leu substituted at residue 236.

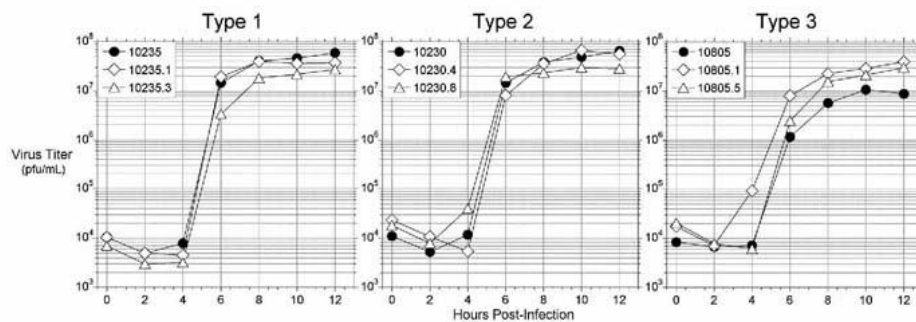


FIG 1 Single-step growth curves. Cultures of LLC-MK₂ cells were infected at an MOI of 10 with drug-susceptible parent virus and two drug-resistant variants for each poliovirus type (types 1, 2, and 3) described in Table 3. The amount of infectious virus present at various times postinfection was quantified by plaque assay. Solid circles, parental viruses; diamonds, VP1 I194 variants (VP1 I192 for type 3); triangles, VP3 A24 variants.

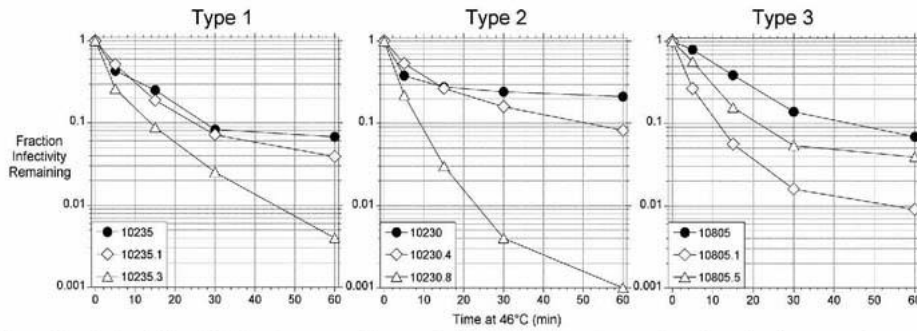


FIG 2 Thermal inactivation of virus infectivity. Drug-susceptible parental virus and two drug-resistant variants for each poliovirus type (types 1, 2, and 3) described in Table 3 were exposed to 46°C for various times, and the remaining infectivity determined by plaque assay. Solid circles, parental viruses; diamonds, VP1 I194 variants (VP1 I192 for type 3); triangles, VP3 A24 variants.

DISCUSSION

The National Research Council of the National Academies has recommended that at least one, preferably two, polio antiviral drugs be developed as a supplement to the tools currently available

for control of polio outbreaks posteradication (4, 5). Pursuant to this recommendation, poliovirus-specific capsid inhibitor V-073 is being advanced clinically to assess the potential utility of poliovirus antiviral drugs in the treatment of chronic poliovirus infec-

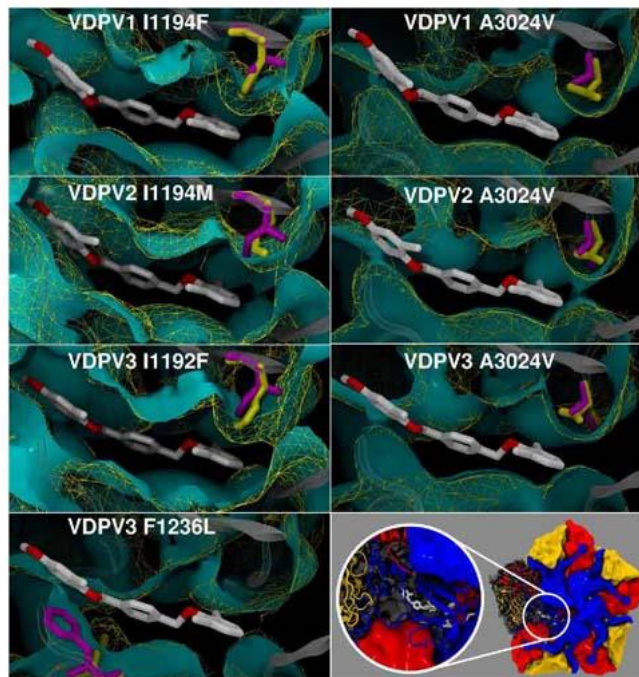


FIG 3 Location in the poliovirus type 2 structure of predominant amino acid substitutions that confer resistance to V-073, based on the crystallographic structure of V-073 bound to poliovirus type 2 (10). Superolateral cutaway view of electron density maps showing V073 bound within the hydrophobic pocket of VP1. Density mapping at 1 Å resolution and 2 Å radius representing the original parental strain (cyan) with a wireframe overlay (yellow) of the resistant virus. Amino acid differences are shown as a liquorice representation with parental (magenta) and variant (yellow). The final frame in the figure shows the position of V073 (white) relative to the overall pentamer structure with VP1 (blue), VP2 (yellow), and VP3 (red); VP4 is not visible in this representation.

tions and management of polio incidents. As with the application of any antiviral drug, the issue of treatment-emergent drug resistance presents a potential obstacle to implementation. It is important to understand the potential for and consequences of antiviral resistance.

We show here that poliovirus variants with reduced susceptibility to V-073 can be isolated in cell culture from otherwise drug-susceptible virus populations. The frequency of these variants in virus populations was estimated at 3.20×10^{-5} to 42.7×10^{-5} (geometric mean, 16.8×10^{-5}). The frequency of drug resistance observed with V-073 is similar to that reported for other capsid inhibitors. For example, the frequency of HRV14 resistance to WIN 52084 was about 4×10^{-5} (7), and that of coxsackievirus B3 to pleconaril was 5×10^{-5} (6).

It was reported previously that poliovirus type 3 variants selected for resistance to capsid inhibitors may also be dependent on drug for their growth (11–13). Drug dependence is characterized by plating indices (ratio of the plaque titer in the presence of drug to the plaque titer in the absence of drug) of about 400 to >2,000, while nondependent drug-resistant variants have values close to 1 (i.e., similar plaque titers with or without drug) (11). The variants studied here had plating indices that ranged from 0.42 to 1.24, indicating a lack of drug dependence. It is unclear why we did not observe drug-dependent viruses in our study; however, it has been reported that the frequency of such viruses may vary widely among different enteroviruses (11). Of course, the physicochemical properties of the drug compound may also influence the type of variants detected and V-073 is chemically distinct from the compounds used to select drug-dependent variants, despite their common mechanism of action.

As described previously for other picornavirus capsid inhibitors, the basis for resistance lies in single amino acid substitutions involving residues that line the drug-binding pocket (6, 7, 12). Here, sequence analysis of 160 independent resistant variants (80 isolates of poliovirus type 1, 40 isolates each of types 2 and 3) established that V-073 resistance also involved single amino acid changes, predominantly at two sites, at isoleucine residue 194 in VP1 (192 in poliovirus type 3) or in VP3 at alanine 24. Minor sites where an amino acid change conferred resistance include proline₁₆₁, isoleucine₁₈₃, and phenylalanine₂₃₆, all in VP1. The locations of the amino acid substitutions are consistent with those observed in polioviruses and rhinoviruses that are resistant to other capsid-binding antiviral compounds (7, 12). For example, the I1192F change confers resistance to WIN 51711 in PV3 (12) and V1188M, lying in a similar position in human rhinovirus 14, confers resistance to the related compound, WIN 52084 (7).

The resistant viruses studied in this report were derived in the laboratory from drug-susceptible virus populations by selection in cell culture. When a collection of naturally occurring polioviruses, including wild polioviruses and VDPVs, were tested for V-073 susceptibility, all of the 42 polioviruses were susceptible, with EC_{50} s ranging from 0.003 to 0.126 μ M (14). In that study, all polioviruses possessed the drug-susceptible genotype encoding isoleucine at residue 194 (192 in type 3) in VP1 and alanine at residue 24 in VP3. When the CDC database of 4,500 VP1 sequences and 250 VP3 sequences, representing all known poliovirus genotypes, including sequences for recently circulating lineages, was inspected, there was an absolute conservation of these two residues. Thus, while we are able to select for and isolate V-073-resistant variants from drug susceptible virus stocks in the

laboratory, the above observations suggest that changes at these positions are rare in nature.

Our initial cell culture investigations into the stability and fitness of V-073-resistant variants presented here revealed mixed results. Single cycle cell culture growth of resistant variants was indistinguishable from that of their corresponding parental viruses, an observation described previously for other picornavirus capsid inhibitors (6, 11). Furthermore, the drug resistance phenotype appeared relatively stable upon cell culture passage of the resistant viruses studied; however, one of the six variants studied appeared to reacquire some level of drug susceptibility. Finally, the thermal stability data suggested generally reduced virion stability of the resistant viruses compared to their susceptible parents, a feature previously noted for picornavirus capsid inhibitor resistance (6, 11).

To explain the lack of drug-resistant viruses among a broad panel of poliovirus isolates, as well as the lack of genotypic representation in the database, additional assessments of V-073 resistance are necessary. It may be that the parameters of cell culture do not reflect fitness penalties that would be associated with resistance to V-073 in infections in a living host. Indeed, when evaluated in mice, we found that laboratory-derived V-073-resistant variants exhibit clear attenuation of their replicative capacity and neurovirulence (9). Further study of V-073 resistance, particularly in the context of treatment-emergent resistance in clinical studies with V-073, is warranted.

ACKNOWLEDGMENTS

We thank Walter Dowdle for his continued support and Aaron Curns for advice on statistical analyses.

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of Centers for Disease Control and Prevention.

ViroDefense, Inc., received financial support from The Task Force for Global Health, Inc., through grants from the WHO and Rotary International.

M.S.C. is a principal in and D.C.P. is a consultant to ViroDefense, Inc., the corporate sponsor of drug candidate V-073. There are no other declarations among the authors.

REFERENCES

- Alexander LN, et al. 2004. Vaccine policy changes and epidemiology of poliomyelitis in the United States. *JAMA* 292:1696–1701.
- Anonymous. 2005. Cessation of routine oral polio vaccine (OPV) use after global polio eradication: framework for national policy makers in OPV-using countries, p 12. World Health Organization, Geneva, Switzerland.
- Buontempo PJ, et al. 1997. SCH 48973: a potent, broad-spectrum, anti-enterovirus compound. *Antimicrob. Agents Chemother.* 41:1220–1225.
- Collett MS, Neyts J, Modlin JF. 2008. A case for developing antiviral drugs against polio. *Antivir. Res.* 79:179–187.
- Committee on Development of a Polio Antiviral and Its Potential Role in Global Poliomyelitis Eradication. 2006. Exploring the role of antiviral drugs in the eradication of polio. National Research Council, Washington, DC.
- Groarke JM, Pevear DC. 1999. Attenuated virulence of pleconaril-resistant coxsackievirus B3 variants. *J. Infect. Dis.* 179:1538–1541.
- Heinz BA, et al. 1989. Genetic and molecular analyses of spontaneous mutants of human rhinovirus 14 that are resistant to an antiviral compound. *J. Virol.* 63:2476–2485.
- Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. *J. Mol. Graphics* 14:33–38.
- Kouivaskala DV, et al. 2011. Immunologic and pathogenic properties of poliovirus variants selected for resistance to antiviral drug V-073. *Antivir. Ther.* 16:999–1004.
- Lentz KN, et al. 1997. Structure of poliovirus type 2 Lansing complexed with antiviral agent SCH48973: comparison of the structural and biological properties of three poliovirus serotypes. *Structure* 5:961–978.

11. Mosser AG, Rueckert RR. 1993. WIN 51711-dependent mutants of poliovirus type 3: evidence that virions decay after release from cells unless drug is present. *J. Virol.* 67:1246–1254.
12. Mosser AG, Sgro J-Y, Rueckert RR. 1994. Distribution of drug resistance mutations in type 3 poliovirus identifies three regions involved in uncoating functions. *J. Virol.* 68:8193–8201.
13. Mosser AG, Shepard DA, Rueckert RR. 1994. Use of drug-resistance mutants to identify functional regions in picornavirus capsid proteins. *Arch. Virol. Suppl.* 9:111–119.
14. Oberste MS, et al. 2009. In vitro antiviral activity of V-073 against polioviruses. *Antimicrob. Agents Chemother.* 53:4501–4503.
15. Pevear DC, Tull TM, Seipel ME, Groarke JM. 1999. Activity of pleconaril against enteroviruses. *Antimicrob. Agents Chemother.* 43: 2109–2115.
16. Phillips JC, et al. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.
17. Schwede T, Kopp J, Guex N, Peitsch MC. 2003. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* 31: 3381–3385.
18. Zoete V, Cuendet MA, Grosdidier A, Michielin O. 2011. SwissParam, a fast force field generation tool for small organic molecules. *J. Comput. Chem.* 32:2359–2368.

The association of picornaviruses with gastroenteritis



*Bruce R. Thorley and
Jason A. Roberts*

Enterovirus Reference Laboratory
Victorian Infectious Diseases
Reference Laboratory
10 Wreckyn Street
North Melbourne, VIC 3051
Tel (08) 9842 2607
Fax (08) 9842 2605
Email bruce.thorley@mh.org.au
and jason.roberts@mh.org.au

Many members of the family *Picornaviridae* survive passage through the stomach and some are known to replicate in the intestinal tract. Yet these picornaviruses are not considered as a leading cause of acute gastroenteritis. Further investigation of gastroenteritis cases not associated with viruses known to cause diarrhoea may serve to delineate further between picornaviruses causing gastroenteritis and an incidental finding.

Picornaviruses are non-enveloped, single-stranded, positive-sense RNA viruses (Figure 1) classified in the order *Picornavirales*, family *Picornaviridae*. The 12 picornavirus genera cause a wide range of diseases in both humans and animals: *Aphthovirus*, *Astrovirus*, *Cardiovirus*, *Enterovirus*, *Erbovirus*, *Hepatitis virus*, *Kobovirus*, *Parechovirus*, *Sapovirus*, *Senecavirus*, *Teschovirus* and *Tremovirus* (Table 1)¹. While adenoviruses, astroviruses, caliciviruses and rotaviruses are accepted as the main causal agents of viral gastroenteritis, a variety of picornaviruses is also routinely detected in faecal specimens. However, it can be difficult to prove a causal association between gastroenteritis of unknown aetiology and a family of viruses that are often transmitted by the faecal-oral route but result in asymptomatic infection.

Enteroviruses

Despite human enteroviruses deriving their name from one of their principal sites of replication as enteric viruses, any diarrhoeal symptoms are usually considered secondary to more major diseases resulting from spread to other organs via the

bloodstream. The range of clinical presentation of enterovirus infection includes poliomyelitis, meningitis, encephalitis, myocarditis and hand, foot and mouth disease. A viral meningitis outbreak associated with echovirus types 6 and 9 in Western Australia reported gastrointestinal symptoms in 83% (55/66) of the echovirus 9 cases and 63% (37/59) of the echovirus 6 cases². Diarrhoea and/or vomiting was reported in 10% (7/70) of the confirmed echovirus 6 and 9 infections with non-meningitis

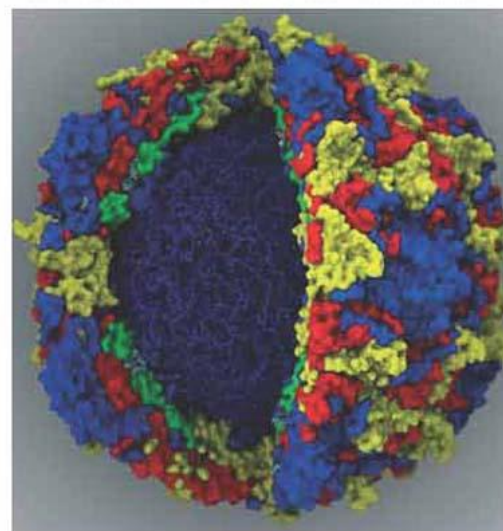


Figure 1. Computational model of the Mahoney strain of wild poliovirus type 1, family *Picornaviridae*, genus *Enterovirus* after molecular dynamics simulation. The non-enveloped virus capsid is composed of four structural proteins-VP1 (blue), VP2 (red), VP3 (yellow) and VP4 (green)-and associated lipids. The single stranded positive sense RNA genome (purple) is depicted within the capsid associated with ions.

illness. Enterovirus 71 is most commonly associated with hand, foot and mouth disease but cases with neurological involvement have been reported in many countries including Australia. While meningitis or encephalitis was reported as the primary diagnosis in 15 of 24 hospitalised patients with confirmed EV71 infection in Denmark, nine of the 15 patients had symptoms of diarrhoea and gastroenteritis was the main diagnosis for another four patients⁵.

Human rhinoviruses are now classified in the genus *Enterovirus*¹. A recent publication reported the detection of species C rhinovirus in faecal specimens from three patients with diarrhoea in the absence of respiratory symptoms⁴. Other pathogens – parechovirus, rotavirus and *Salmonella* group B – were detected individually from the three specimens so the role of species C rhinoviruses with gastroenteritis is unclear. Nevertheless, the detection of species C rhinovirus in faecal specimens warrants further investigation to determine whether they have an association with diseases other than respiratory infections.

Aichi virus

Aichi virus was initially isolated in 1989 from the prefecture of the same name in Japan. Sequencing the virus genome resulted in its classification as a member of the picornavirus genus *Kobuvirus*¹. Aichi virus has been identified in 1–2% of sporadic cases of gastroenteritis worldwide and up to 20% of faecal specimens from outbreak cases in Japan, mainly associated with eating oysters⁵. However, a recent report from Germany provided evidence of an association between Aichi virus and acute diarrhoea via human-to-human transmission that did not involve food⁶. Bovine kobuvirus has been officially classified by the International Committee on Taxonomy of Viruses but has not been associated with gastroenteritis. In contrast, porcine kobuvirus has been associated with diarrhoea in South Korea and other countries⁷ but its taxonomic classification remains unassigned¹.

Parechoviruses

Human parechovirus types 1 and 2 were initially classified as echovirus types 22 and 23, respectively, but were reclassified in the genus *Parechovirus* and now include types 1–16¹. Parechovirus was detected in 16.3% of 1,824 faecal specimens from children screened by RT-PCR in the Netherlands from 2004 to 2008, without investigating a possible disease association⁸. While in Thailand parechovirus was detected in 14.6% of 82 faecal specimens from children with acute gastroenteritis that

were negative for adenovirus, astrovirus, norovirus, rotavirus and sapovirus supporting an association with the disease⁹.

Novel picornaviruses

In recent years, the screening of faecal specimens for novel viruses by nucleic acid-based methods has produced interesting results. A new picornavirus, human common stool-associated picornavirus or cosavirus, was identified from cases of acute flaccid paralysis (AFP) reported as negative for poliovirus (non-polio AFP) in Pakistan. The virus was determined to have a high prevalence in Pakistan but was only detected in one patient after extensive screening of faecal specimens referred for enteric bacteriology in the United Kingdom¹⁰. Three groups independently identified a new picornavirus that is related to Aichi virus: kobu-like virus associated with stool and sewage or Klassevirus was detected in paediatric gastroenteritis faecal specimens from Australia and the USA and a sewage sample from Spain¹¹; and from paediatric gastroenteritis specimens by a second group¹²; while the closely related stool Aichi-like virus or salivirus was identified from non-polio AFP cases from Nigeria¹³. These three picornaviruses remain unassigned by the International Committee on Taxonomy of Viruses (Table 1)¹.

The screening of gastroenteritis specimens that are negative for the known causal agents by recently developed techniques such as pyrosequencing¹⁴ may identify other novel picornaviruses; however, epidemiological studies in a number of countries will be required to elucidate the extent of their involvement with gastroenteritis.

In summary, the large group of picornaviruses do have an association with gastroenteritis. Aichi virus, parechoviruses and certain species of enterovirus are reported to have strong associations with viral gastroenteritis. Further investigation of gastroenteritis involving the novel picornaviruses described in this article is required and it will be intriguing to determine whether species C rhinovirus could also have a role.

References

1. Knowles, N.J. *et al.* (2012) Family *Picornaviridae*. In *Virus taxonomy Ninth report of the International Committee on Taxonomy of Viruses*. (King, A.M.Q. *et al.*, eds), pp. 855–880, Academic Press.
2. Ashwell, M.J.S. *et al.* (1996) Viral meningitis due to echovirus types 6 and 9: epidemiological data from Western Australia. *Epidemiol. Infect.* 117, 507–512.
3. Badran, S.A. *et al.* (2011) Clinical and virological features of enterovirus 71 infections in Denmark, 2005 to 2008. *Scand. J. Infect. Dis.* 43, 642–648.

Table 1. *Picomaviridae* genera and type species assigned by the International Committee on Taxonomy of Viruses¹.

Genus	Type Species
<i>Aphthovirus</i>	<i>Foot-and-mouth disease virus</i>
<i>Avihepatovirus</i>	<i>Duck hepatitis A virus</i>
<i>Cardiovirus</i>	<i>Encephalomyocarditis virus</i>
<i>Enterovirus</i>	<i>Human enterovirus C</i>
<i>Erbovirus</i>	<i>Equine rhinitis B virus</i>
<i>Hepatovirus</i>	<i>Hepatitis A virus</i>
<i>Kobuvirus</i>	<i>Aichi virus</i>
<i>Parechovirus</i>	<i>Human parechovirus</i>
<i>Sapelovirus</i>	<i>Porcine sapelovirus</i>
<i>Senecavirus</i>	<i>Seneca Valley virus</i>
<i>Teschovirus</i>	<i>Porcine teschovirus</i>
<i>Tremovirus</i>	<i>Avian encephalomyelitis virus</i>
Unassigned	Human cosavirus Human klassevirus or salivirus Porcine kobuvirus

- Lau, S.K.P. *et al.* (2012) Detection of human rhinovirus C in fecal samples of children with gastroenteritis. *J. Clin. Virol.* 53, 290–296.
- Reuter, G. *et al.* (2011) Kobuviruses – a comprehensive review. *Rev. Med. Virol.* 21, 32–41.
- Drexler, J.F. *et al.* (2011) Aichi virus shedding in high concentrations in patients with acute diarrhea. *Emerg. Infect. Dis.* 17, 1544–1548.
- Park, S.J. *et al.* (2010) Molecular detection of porcine kobuviruses in pigs in Korea and their association with diarrhea. *Arch. Virol.* 155, 1803–1811.
- Benschop, K. *et al.* (2008) High prevalence of human parechovirus (HPeV) genotypes in the Amsterdam region and identification of specific HPeV variants by direct genotyping of stool samples. *J. Clin. Microbiol.* 46, 3965–3970.
- Pham, N.T.K. *et al.* Diversity of human parechoviruses isolated from stool samples collected from Thai children with acute gastroenteritis. *J. Clin. Microbiol.* 48, 115–119.
- Kapoor, A. *et al.* (2008) A highly prevalent and genetically diversified *Picomaviridae* genus in South Asian children. *Proc. Natl Acad. Sci. USA* 105, 20482–20487.
- Holtz, L.R. *et al.* (2009) Klassevirus I, a previously undescribed member of the family *Picomaviridae*, is globally widespread. *Virology* 396, 86.
- Greninger, A.L. *et al.* (2009) The complete genome of klassevirus – a novel picornavirus in pediatric stool. *Virology* 396, 82.
- Li, L. *et al.* (2009) A novel picornavirus associated with gastroenteritis. *J. Virol.* 83, 12002–12006.
- Roberts, J.A. and Thorley, B.R. (2010) New approaches to enterovirus identification. *Microbiol. Australia* 31, 138–141.

Biographies

Bruce Thorley is the Head of the Australian National Enterovirus Reference Laboratory, which is accredited by the World Health Organization as a Polio Regional Reference Laboratory and located at the Victorian Infectious Diseases Reference Laboratory. He is also the Chief Investigator for the national Acute Flaccid Paralysis Surveillance program that investigates polio-like illness in children. In recent years he has focused on broadening surveillance for poliovirus in Australia by establishing an Enterovirus Reference Laboratory Network and testing environmental samples from sentinel sites.

Jason Roberts is a Senior Medical Scientist at the Victorian Infectious Diseases Reference Laboratory with a background in molecular diagnostic assay development. He is a consultant virologist for the Australian Polio Expert Panel and acts a temporary advisor for the WHO Polio Laboratory Network. His research interests relate to neurotropic RNA viruses, in particular the characterisation and molecular modelling of enteroviruses. He is currently collaborating with RMIT University, the Victorian Partnership for Advanced Computing and the Victorian Life Sciences Computation Initiative to recreate ‘in silico’, novel enterovirus and poliovirus proteins to determine the mechanisms of specific mutations associated with neurovirulence.