Integrative Structural Biology of Viruses

Helen Mary Elizabeth Duyvesteyn
DPhil in Structural Biology
Wolfson College, University of Oxford



Supervisors:

Professor David Ian Stuart

Co-Supervisors:

Professor Juha Huiskonen
Dr Allen Orville

'Trinity' Term 2019 Word count: 49569

Declaration

I hereby declare that, aside for where specific reference is made to the work of others, the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This thesis is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgements. This dissertation contains fewer than 50,000 words, exclusive of appendices, figure legends and tables.

Helen Mary Elizabeth Duyvesteyn

September 2019

Abstract

Picornaviruses include significant infectious agents and are challenging structural targets owing to their pathogenicity. They share a common architecture, but subtle differences in their structure mediate tissue tropism, receptor binding and uncoating cues. This thesis focuses on members of the Enterovirus and Aphthovirus genera, and aims to contribute structural understanding of their lifecycle and the forces that underpin and modulate capsid stability to ultimately inspire novel therapeutics and vaccines. I take advantage of improvements in cryoEM which enable structural biologists to routinely solve complex structures at close to atomic resolution and resolve sample heterogeneity.

EV-F3 is a non-pathogenic enterovirus. Presented crystallographic structures or EV-F3 highlight a potentially druggable site, and soaks with siallactose suggest a cellular receptor. Cryo-EM investigations show intermediate states, confirming that heating can cause transitions to altered particle states seen in other enteroviruses. Long-wavelength X-ray data indicate the presence of K^+, but not Ca^{2+} ions within the capsid.

I present structural characterisation of recombinant VLP vaccine candidates for poliovirus (PV) and foot-and-mouth disease virus (FMDV). For FMDV, I show a continuum of cryo-EM structures with a potential S-S bridge at the 2-fold interface and varied 2-fold structures. A pentamer subassembly is characterised and analysis of metal ion composition in FMDV VLPs gives no evidence for bound Ca^{2+} ions.

Single-particle cryo-EM experiments of yeast-derived thermostabilised PV-1 and PV-3 VLPs show both C and D antigenic states to be present for PV-1 and demonstrate a predominance of D-antigenic states for PV-3.

Finally, preliminary investigations into aphthovirus uncoating include attempts to express full-length $\alpha_v\beta_6$ integrin receptor and a preliminary structure of FMDV bound to $\alpha_v\beta_3$ head-group, demonstrating VLP affinity for virus receptors. I demonstrate the potential of caged protons for time-resolved structural studies of FMDV uncoating and present a novel expanded particle structure of an aphthovirus.

Acknowledgements

The past four years have flown past, and have, for the most part, been really quite fun. My journey into the world of structural biology and virology would never have been so enjoyable without the support of friends, mentors and family.

First and foremost, I wish to throughly thank my supervisor, Dave Stuart, for guiding me through my PhD, and facilitating my exploration of a diverse range of methodologies. I am extremely grateful for the freedom and opportunities that he has given me to do really fun science, tolerance of my very stupid mistakes, and the frequent times that I have ended up careering down an infinitely empty rabbit hole and he has steered me back to perspective.

Other members of the Stuart group have also been exceptionally patient and supportive as I have got (and am still getting) to grips with virology, general biology and structural biology. Liz Fry has been a fantastic, kind and patient member, often working in the background of the Stuart group, always happy to help with queries, however stupid, and mediate sample requests and delivery. Jingshan Ren, master of crystallographic refinement, has been extremely helpful, and I am indebted to his recent efforts sorting out PDB depositions for the glutathione paper while I was in 'the thesis hole'. Thanks also to Mo Bahar for initiating me into the 'wet lab'.

Many thanks to 'team admin' for all of our fun, ridiculous chats, but, also for all of their behind-the-scenes work. Claire Morris and Agata Krupa have helped me on numerous occasions with navigating Dave's diary logistics for urgent meetings, and Mel Matthews has always been extremely helpful with my many queries about finances, particularly recently when reviewing my grant situation.

Data processing, particularly for EM data, would never have been so painless without the ever-ready rescomp computing cluster. At times when jobs have been urgent, especially for XFEL trips, Robert Esnouf has been rapid and on point for meeting our needs. I also wish to thank Jun Dong for her IT support, especially when ordering new computing equipment.

Thanks to Margaret Jones and Ollie Hayward, who have done a sterling job keeping the lab ticking, and dealing with my multitude of order requests, such that I was able to focus on my own work. Special thanks also to Robin Parsons and Ben Knight for keeping things running smoothly in

OPIC. Tom Walter and Karl Harlos have been instrumental in my crystallography successes to date; fishing many of the more successful crystals. Without their fishing expertise, I would certainly have lost many hours, numerous crystals, and my (relative) sanity. Tom Walter has also, very patiently, helped with a number of my crystal tray and thermofluor set ups, providing sound advise, and patiently guiding in times of stress.

Many of my microscope sessions would have been an abject failure without the help of the EM support staff, particularly Dimple Kotecha and Bilal Qureshi, who have been exceptional at keeping the microscopes happy, and patiently responding to my weekend-session-panic WhatsApp messages. Thank you also to the thermo engineers, Abhay Kotecha in particular, for the excellent introductory training that they provided.

A number of members of Diamond Light Source have provided much assistance with my experiments, especially those at 'best beamline', I24 and Armin and Ramona at I23. SACLA trips would never have been as smooth or fun without team I24. Robin Owen and John Beale have also given me much guidance during my time using the microspectrometer.

During my brief stay at Pirbright, I was made to feel most welcome by the Toby Tuthill group, especially Jamie Kelly, who introduced me to the basics of virology. Fun discussions of all things picorna served as an excellent introduction to my PhD.

I wish to also thank numerous people at cicBIOGUNE in Bilbao, particularly Nicola Abrescia and his group. Isaac Santos has become a true friend, especially following numerous unsuccessful microscopy screening sessions with PRD1. Thanks also to Hani Yosef who patiently steered me through the basics of mammalian cell culture, and always provided words of encouragement after the n-ty-th empty western blot.

My friends have also been instrumental in maintaining my (relative) sanity over the past few years. Foremost, I wish to thank Helen Ginn for being an exceptionally awesome friend, 'long-distance housewife', mentor, advisor, and 'root'.

Two further key players in the maintenance of my relative mental stability are the 'Emilys'; Em Griffiths and 'L'Em' (Emily Kingston) and their respective husbands, Rob and Matt. Thanks especially to Em, who, in addition to Helen Ginn, has kept me going during recent months with regular Skype sessions and amusing messages. I also wish to remember my friend and old housemate, Robin Virgo who could have been writing up at the same time as me had he lived to see his PhD

through.

PhD life would never have been so enjoyable without the fun, supportive and helpful Strubi community. I feel extremely privileged to be based at a place with so many like-minded, highly knowledgeable, friendly individuals. I would like to particularly mention Max Renner (for our endless chats putting the world to rights, and boringly similar views), Rob Stass (with whom I have had many fun adventures during our mid-PhD years), Reinis Ruza (who has shared many interesting publications and provided sound purification advice and protocols in times of stress), Liz Allen, Alice Stelfox, Serban Ilca, Weng Ng, Claire Coupland, Vojtech Prazak, Rachel Wooley, and Daming Zhou. I wish also to acknowledge members of 'office 2.0', my more recent workspace; Mattia Bertinelli (especially for his excellent IT support and general 'positivity'), Rhys Pryce (for his beautifully annoying sense of humour, and supportive nature) and to Vicky, my 'work wife', for being an awesome friend and excellent rant giver and receiver, especially with recent dramas.

Finally, thanks to my family; Mum, Dad, my brother Rob, and Grandma Mary, for cheering me on from Dorset these past few years.

Note to Reader

Conventions

- Virus amino acid residues will be numbered with the first digit defining the capsid protein (VP) to which it belongs, consistent with convention. For instance, residue 41 of VP2 will be residue 2041.
- Non-standard abbreviations are listed in the 'Abbreviations' section, but, where clarity was deemed necessary, some abbreviations are also explained within the main text.
- Amino acids will be referred to by their single letter codes when referencing a particular residue and sequence number. In cases where more clarity is needed, such as when stating an amino acid without a residue number, the three-letter abbreviation will be used.
- Unless otherwise stated, picornavirus viral proteins (VPs) will be coloured as follows: VP1 blue; VP2 green; VP3 magenta; VP4 yellow or yellow/orange.
- 2-, 3- and 5-fold icosahedral symmetry axes are shown, where appropriate, as an ellipse, triangle and pentamer shapes, respectively.
- Surface loops in picornavirus structures are referred to by the two β strands that they link e.g. the loop that connects the 'G' and 'H' strands will be referred to as the GH-loop.

General Contributions

- The successful crystals used for the majority of crystallographic structures presented in this thesis were fished by either Karl Harlos and Tom Walter. Assistance with remote data collections was provided by my supervisor David Stuart, as well as Elizabeth Fry and Jingshan Ren.
- Tom Walter helped with the preparation of crystallisation trays for FMDV A22cc VLP and EV-F3 virus sample.
- Data collected at the I23 beamline at Diamond Light Source were set up by Armin Wagner, and samples were mounted by Ramona Durnam.

- All cryo-EM data collections were set up by myself.

Chapter-specific Notes

Chapter 2

- Sections for methods performed by someone other than myself are marked with an asterisk, '*'.

Chapters 3 and 4

- Material for the room temperature EV-F3 empty and mature crystal structures were provided by David Rowlands (Astbury Centre for Structural Molecular Biology, ASCMB, Leeds) and collected and initially refined by Jingshan Ren and Elizabeth Fry.
- Material for the EV-F3 GSH story (Chapter 4, Section 4.1.3) has been accepted for publication (see Chapter 9).

Chapter 5

- Vaccinia-derived VLPs were prepared by Alison Burman at the Pirbright Institute, UK.

Chapter 6

- PV-1 and PV-3 yeast VLPs were prepared by Lee Sherry (ACSMB, Leeds) and purified by Mohammad Bahar (Strubi, Oxford).
- Cryo-EM grids for PV-1 and PV-3 were prepared by Mohammad Bahar.

Chapter 7

- Ectodomain integrin constructs for α_v and β_6 were kindly donated by Alison Burman, Pirbright Institute, UK, and that of β_3 from Thomas Bowden, Strubi.
- Some of the initial work for integrin expression was performed at the Centre for Cooperative Research in Biosciences (CIC bioGUNE), Bilbao, Spain, under the supervision of Nicola Abrescia.
- ERAV was first purified under the guidance of James Kelly from the Toby Tuthill group at The Pirbright Institute.

- Tom Walter assisted with the provisional cryo-EM screening of ERAV empty particles grids, conducted on a T12 microscope.
- Microspectrometer work was performed with the assistance of John Beale, based at I24 beamline, Diamond Light Source UK, with advice given by Robin Owen.

Non-Standard Abbreviations

Å	Ångstrom (10^{-10} m)
AEV	avian encephalomyelitis virus
AIrn	FMDV A-Iran subtype.
AIrncc	FMDV A-Iran subtype with S-S bridge at residue 2093.
AiV	aichivirus
APS	active-pixel sensor
Asiacc	FMDV Asia1 serotype with S-S bridge at residue 2093.
AsiaShcc	FMDV Asia1, subtype Shamir with a S-S bridge at residue 2093.
BEV	bovine enterovirus
bp	(nucleotide) base pairs
BSA	bovine serum albumin
BTB	bromothymol blue
CAR	coxsackievirus and adenovirus receptor
CCD	charged coupled device
CMC	carboxymethyl cellulose
CMOS	complementary metal oxide semiconductor
cDNA	complementary deoxyribonucleic acid
CNS	Crystallography NMR system (refinement suite)
CVB	coxsackievirus B
CVA	coxsackievirus A
CV-A10	coxsackievirus A10
CV-A16	coxsackievirus A16
CV-A24v	coxsackievirus A24v
CV-B3	(CVA-B3) coxsackievirus B3
CPE	cytopathic effect
CTF	contrast transfer function
DLS	Diamond Light Source
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
Dox	Doxycycline hydrochloride
E-[1-9]	Echovirus, for instance E-9 is echovirus-9
E.coli	Escherichia coli
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
EM	electron microscopy
ERAV	equine rhinitis A virus
EV	enterovirus
EV-A71	enterovirus A71
EV-D68	enterovirus D68
GMEM	Glasgow modified Eagle's medium
GMEM-BHK21	Glasgow modified Eagle's medium for BHK21 cells
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
FBS	fetal bovine serum
FMDV	foot-and-mouth disease virus
FSC	Fourier Shell Correlation
GSH	glutathione, reduced
GSSG	glutathione, oxidised
GUI	Graphical User Interface
HIV	Human Immunodeficiency Virus
HRP	horseradish peroxidase
hr	hour(s)
HRV	human rhinovirus, for instance HRV14.

IRES	internal ribosome entry site
LLG	log-likelihood
MAbs	monoclonal antibodies
MES	2-(N-morpholino)ethanesulfonic acid
MQ	ultrapure water from a Milli-Q (Sigma) water purification system.
min	minute/minutes
MOI	multiplicity of infection
MOPS	3-(N-morpholino)propanesulfonic acid
MOPV	monovalent oral poliovirus vaccine
NCS	non-crystallographic symmetry
OD	optical density
OIE	The World Organisation for Animal Health
OPV	oral polio vaccine
OTur	OTurkey serotype, FMDV
OTurcc	OTurkey serotype, FMDV with disulphide bridge at residue 2093.
PaSTRy	particle stability thermal release assay
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque forming units, how many particles are capable of forming plaques per unit volume.
p.i.	post infection
PV	poliovirus.
PV-1	poliovirus type 1. The same system applies for types 2 and 3.
PVr	poliovirus receptor, CD155.
rpm	revolutions per minute
RNA	ribonucleic acid
RMSD	root mean square deviation
Tris	Tris-(hydroxymethyl) aminomethane
S	Svedberg unit for sedimentation coefficient.
SAXS	small-angle X-ray scattering
SCARB2	reverse transcription-polymerase chain reaction
SDS-PAGE	sodium dodecyl sulfate polyacrylamide
SEC	size-exclusion chromatography
s	second(s)
Sf9	spodoptera frugiperda 9 cell
SVDV	Swine vesicular disease virus
TAE	Tris-acetate-EDTA
TEM	transmission electron microscopy
Tm	melting temperature.
TMEV	Theiler's Encephalomyelitis Virus
UTR	untranslated region
VLP	viral like particles
VP	viral protein
WHO	World Health Organisation
XFEL	X-ray Free Electron Laser

Contents

1 Biological Introduction 16

1.1 Picornaviruses-an overview 16

1.1.1 Genomic Organisation and Replication 16

1.1.2 Structure and Assembly 18

1.1.3 Capsid dynamics: Uncoating and Genome Release 22

1.1.4 Receptor binding and Entry 23

1.1.5 Metal Ions and Capsid Stability 30

1.1.6 Enteroviruses: Poliovirus and Bovine Enterovirus 32

1.2 Aphthoviruses: Foot and Mouth Disease Virus 37

1.2.1 Overview 37

1.2.2 Genome and Structure 37

1.2.3 Antigenic Sites 39

1.2.4 Receptor binding 42

1.3 Therapeutic Applications 42

1.3.1 The Quest for Broad-Spectrum Antivirals 42

1.4 Vaccination 47

1.4.1 Virus Like Particles 48

1.5 Aims 48

2 Materials and Methods 50

2.1 Molecular Biology 50

2.1.1 Solutions and Media 50

2.1.2 Cloning 50

2.1.3 Transient Expression . . . 55
2.1.4 Stable Cell line generation . . . 59
2.1.5 Expression of Truncated $\alpha_v\beta 3$ Integrin as a FMDV receptor . . . 62
2.2 Virus and VLP Production and Purification . . . 63
2.2.1 Equine Rhinitis Virus A . . . 63
2.2.2 EV-F3 . . . 64
2.2.3 Yeast PV VLP Production* . . . 65
2.2.4 Vaccinia FMDV VLP Production* . . . 66
2.2.5 Virus Preparation for EM . . . 67
2.3 Picornavirus Concentration . . . 67
2.4 Protein Crystallisation and Structure Determination . . . 67
2.4.1 Crystallisation . . . 67
2.4.2 Crystal Harvesting and Data Collection . . . 68
2.4.3 Detectors . . . 70
2.4.4 Data Processing . . . 70
2.4.5 Anomalous X-ray Diffraction for Metal Ion determination . . . 71
2.5 Transmission Electron Microscopy . . . 73
2.5.1 Negative Stain Microscopy . . . 74
2.5.2 Single Particle Cryo-Electron Microscopy . . . 74
2.5.3 Data Acquisition . . . 79
2.5.4 Data Processing . . . 82
2.5.5 Model Fitting . . . 85
2.6 Time-Resolved Cryo-EM . . . 87
2.6.1 Caged compound Control Experiments . . . 88
2.7 Thermofluor Assays . . . 89
2.8 Illustrations . . . 91

3 EV-F3: from Entry to Uncoating 92

3.1 Introduction . . . 92
3.2 Results . . . 93
3.2.1 EV-F3 production and purification . . . 93

3.2.2 Crystallographic structures of Mature and Immature EV-F3 . . . 95
3.2.3 A continuum of Uncoating Intermediates . . . 121
3.3 Discussion . . . 138

4 EV-F3: Co-factor and ligand interactions 140
4.1 Understanding Capsid Stability: Metal ion Binding sites . . . 141
4.1.1 Assessing the Thermostability of EV-F3 Capsids . . . 141
4.1.2 Anomalous Data Collection . . . 143
4.1.3 Glutathione mediates capsid stability at a druggable site . . . 159
4.1.4 Investigating Sialic Acid as an EV-F3 Receptor . . . 171
4.2 Discussion: Implications for Drug Development . . . 183

5 Characterising FMDV VLP Vaccine Candidates 185
5.1 Introduction . . . 185
5.2 Results . . . 189
5.2.1 Investigations into S-S bridging in FMDV VLPs . . . 189
5.2.2 10 mM DTT treated A22cc . . . 216
5.2.3 Investigating the Metal Ion composition of FMDV VLPs . . . 238
5.2.4 Thermofluor analysis of VLPs . . . 246
5.3 Discussion . . . 251

6 Thermostabilised Poliovirus VLPs as Vaccine Candidates 253
6.1 Introduction . . . 253
6.2 PV-1 SC6b . . . 256
6.2.1 Data Processing . . . 257
6.2.2 Structure Analysis . . . 257
6.3 PV-3 SC8 . . . 268
6.3.1 Data Collection . . . 268
6.3.2 Data Processing . . . 268
6.4 Discussion . . . 284
6.4.1 Rationalising the PV-1 SC6b:PV-3 SC8 C/D antigenic divide . . . 284
6.4.2 Assessing the Significance of Internal Capsid Density . . . 285

6.4.3 Grid preparation 286

7 Developing tools to probe Aphthovirus Uncoating 287

7.1 Introduction 287

7.2 Results 290

7.2.1 Cloning and Expression of $\alpha_v\beta_6$ Integrin Receptor 290

7.2.2 Expression of Truncated $\alpha_v\beta_3$ 301

7.2.3 Investigating the use of caged protons for on-grid pH drop 308

7.2.4 Additive Stabilisation of FMDV intermediates 312

7.2.5 A Low Resolution ERAV Expanded Particle 315

7.3 Discussion 320

8 Conclusions and Future Directions 323

9 Publications 328

10 Appendix 375

10.1 Description of contents 375

Chapter 1

Biological Introduction

1.1 Picornaviruses-an overview

The family *Picornaviridae* (order *Picornavirales*, realm *Riboviria*) encompasses an array of clinically pertinent animal and human viruses that cause a spectrum of diseases, including the common cold, hepatitis, meningitis and paralysis (Racaniello, 2007). As its prefix *pico* indicates, member viruses are relatively small (circa 300 Å capsid diameter) with a sedimentation rate (S_{20W}) of 140-165S. To date, 47 genera have been identified, spanning 110 species, categorised based on genetic, biological and biophysical characteristics (Zell et al., 2017). Searching *Picornaviridae* under '[organism]' on the NCBI GenBank yields ca. 122,000 hits this year, by contrast to the ca. 100,000 entries noted by Zell (2018) last year (Clark et al., 2015). This thesis will focus on the genera *Enterovirus* and *Aphthovirus*, a small proportion of the *Picornaviridae*, see Figure 1.1.

Residues will be numbered with the first digit defining the capsid protein to which it belongs, consistent with convention. Thus, residue 41 of VP2 will be residue 2041.

1.1.1 Genomic Organisation and Replication

Picornaviral genomes are highly compact and non-segmented; consisting of a single, positive-sense RNA molecule of circa 6.5-10.1 kb. With the exception of canine picodicistroviruses of genus *Dicipivirus* (Woo et al., 2012), the genome comprises one open reading frame (ORF) that is translated into one polyprotein, before processing by one or more viral proteases into over 25 different polypeptides, Figure 1.2 (Zell, 2018). The polyprotein is composed of three regions; P1 (5' region),

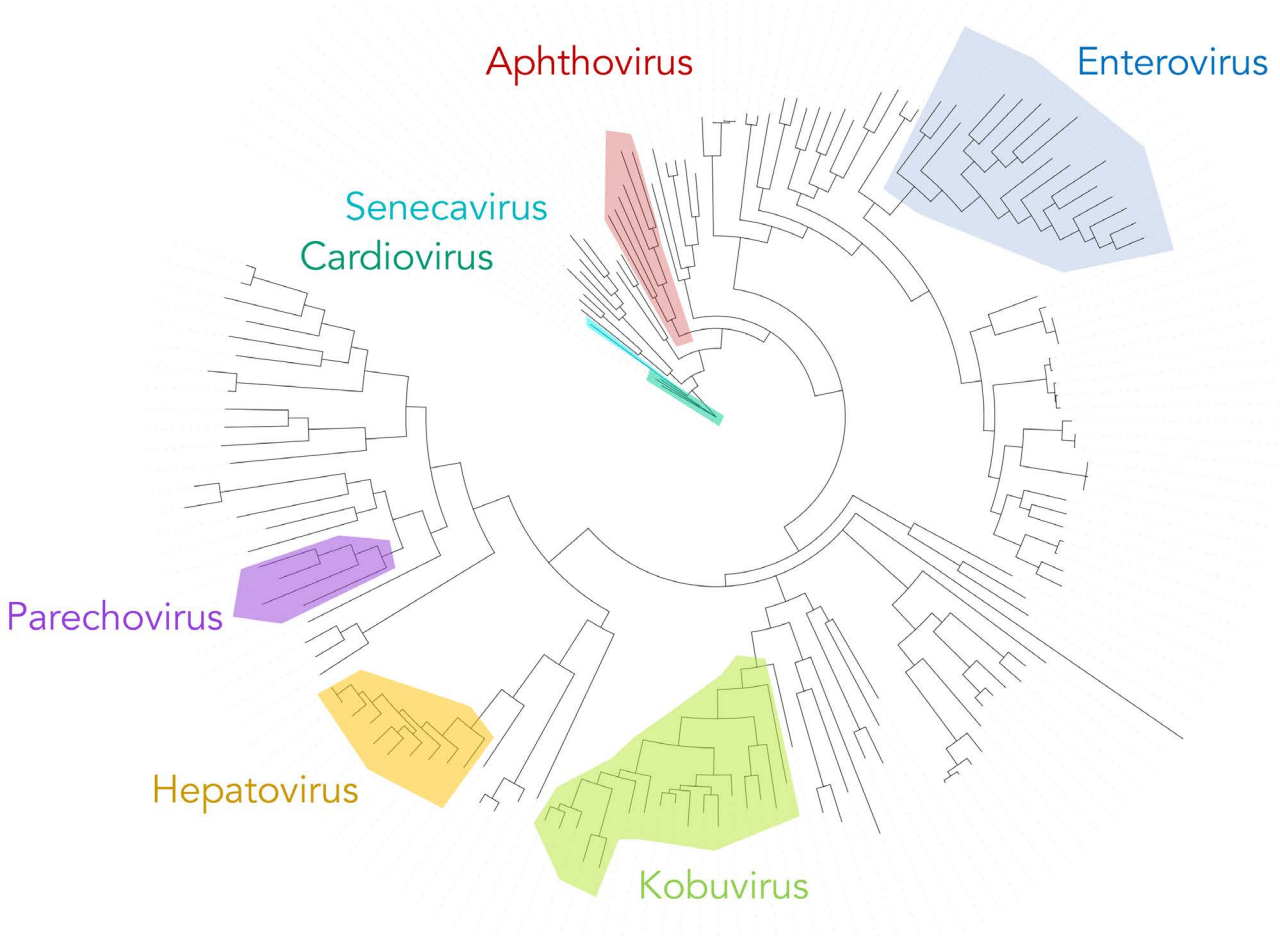


Figure 1.1: Phylogenic tree illustrating the scale of the Picornaviridae family. Sequences were sourced from the ICTV Virus Taxonomy website (https://talk.ictvonline.org/ictv-reports/ictv_online_report/positive-sense-rna-viruses/picornavirales/w/picornaviridae). Pertinent genii are shaded and labelled. Individual virus labels are omitted for clarity.

P2 (middle region) and P3 (3′ region), which are flanked by untranslated regions (UTRs) containing extensive secondary structure that is paramount for genomic function. Small protein(s), VPg, covalently appended to the 5′ end, acts to prime protein synthesis initiation, while the 3′ region contains a poly(A) tail of variable length (Agol, 1991; Gu et al., 2014).

Structural proteins (1AB, 1C and 1D, or VP0, VP3 and VP1) are encoded within the P1 region, while non-structural proteins, 2A-2C and 3A-3D, including proteases, and RNA-dependent polymerase, occur in the P2 and P3 regions (Figure 1.2) and are essential for replication.

Non-structural proteins (NSPs) are beyond the scope of this thesis, but will be briefly described. Perhaps the most conserved NSP, yet least understood, is the 2C protein, which is believed to be a superfamily 3 (SF3) helicase and play a role in enteroviral RNA replication (Teterina et al., 2006; Guan et al., 2017). The 3C region encodes the protease, 3Cpro, that is responsible for the maturation cleavage of precursor polyproteins (Cheah et al., 1990). 3D encodes an RNA-dependent RNA polymerase (RdRp), which has similarities to many other DNA and RNA polymerases (Sankar and Porter, 1992). Replication relies on the reorganisation of cytoplasmic membranous organelles, which is mediated by these viral polymerases and helicases (Zell, 2018).

Significant cross-picornavirus variations occur within NSPs, with some encoding two or three 2A proteins, and others two or three copies of 3B (VPg) (Figure 1.2). Aphthoviruses and cardioviruses, by contrast to other genera, contain a poly(C) tract close to the 5′ end of the RNA that is crucial for viral replication (Zell, 2018). Additionally, all aphthoviruses and cardioviruses, as well as hepatitis A virus, Aichi virus and echovirus 22, contain a leader polypeptide preceding the structural region (Zell, 2018).

1.1.2 Structure and Assembly

During capsid assembly, the P1 polyprotein is cleaved into VP0, VP1 and VP3 (with molecular weights in the range 24-41 kD). Encapsulation of positive-sense viral RNA, and cleavage of VP0 into VP2 and VP4, results in an icosahedrally symmetric 'mature' viral capsid. 'Empty', or 'immature', particles, resembling the mature virion in structure and antigenicity, but devoid of viral RNA, and often with the VP0 uncleaved, have been found in most picornavirus infected cells (excepting cardioviruses). Their role, if any, is unknown, and it has been speculated that they could be decoy for the host immune system, and/or a low-energy pentamer reservoir protected against

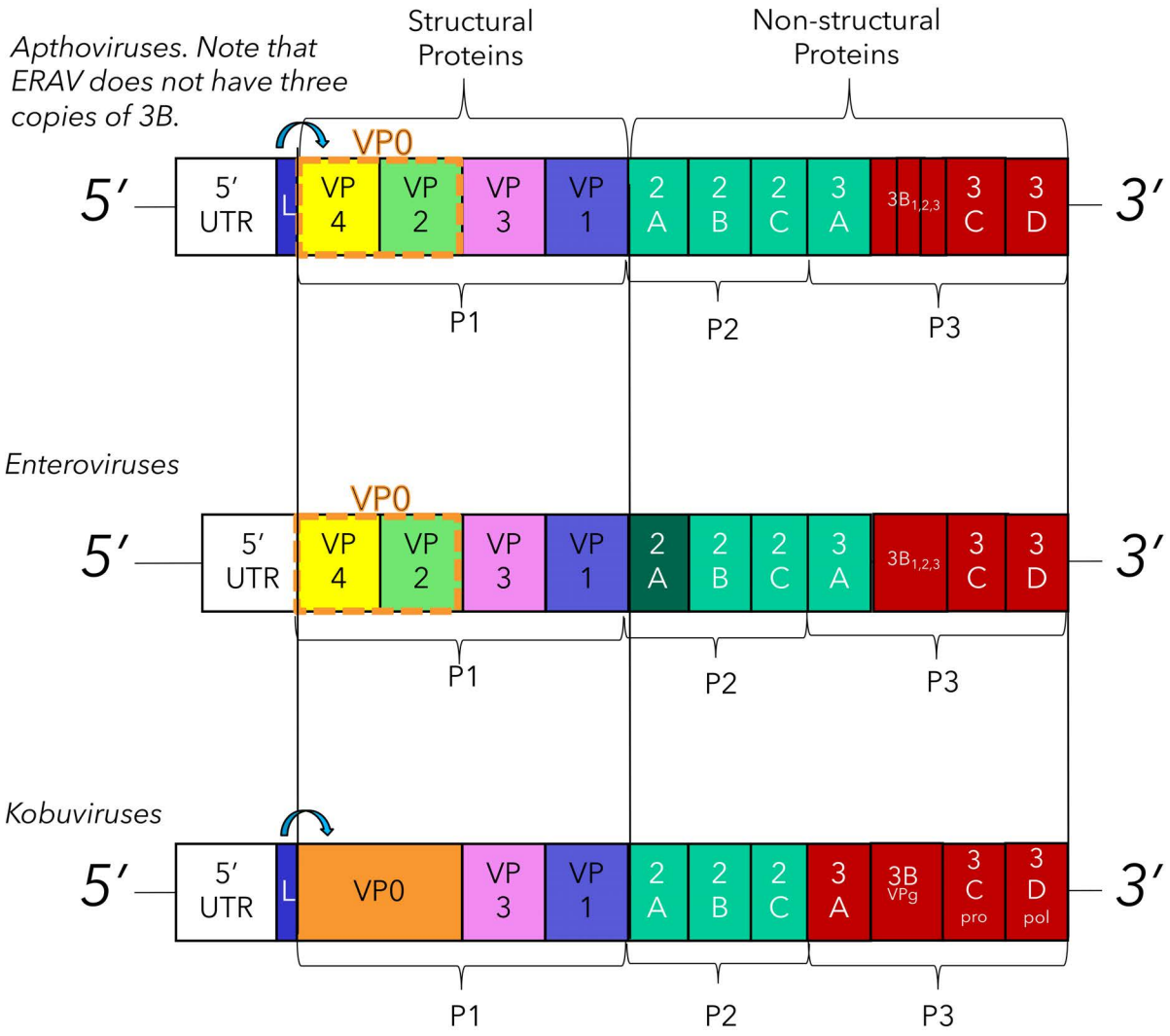


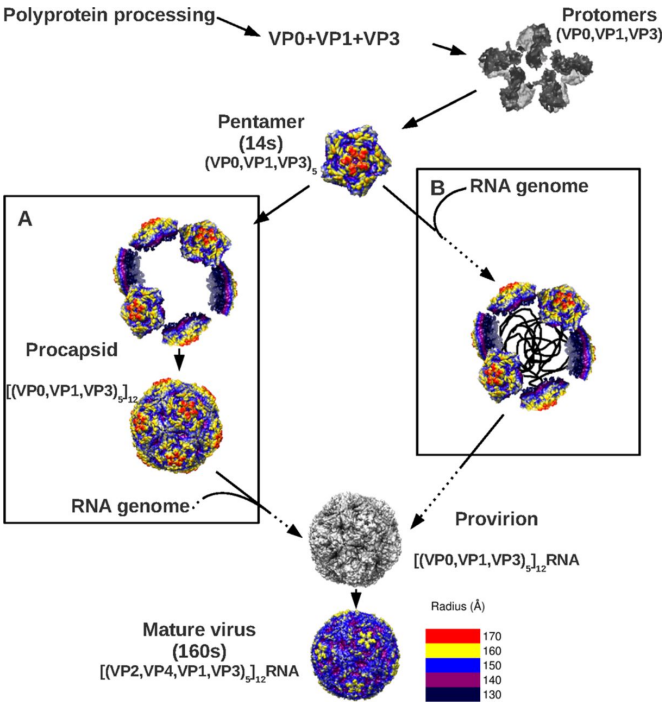
Figure 1.2: Schematic of genome structure for some picornavirus genii, labelled by translation products and polyprotein processing products. Picornaviral RNA exists as a monopartite, linear ssRNA(+) that is composed of a single open reading frame (ORF) encoding a polyprotein flanked by untranslated regions (UTRs) at the N and C termini. This is translated into a single polyprotein prior to cleavage by cis- and trans-acting viral proteases. At the N-terminus, after a covalent linkage to viral protein (VPg) followed by a poly-adenylated region, the relatively long 5′ UTR contains an internal ribosome entry site (IRES) that enables direct translation of the polyprotein (Jang et al., 1988). The P1 capsid protein precursor encodes capsid proteins, while P2 and P3 encode NSPs that are essential for genomic replication and protein processing. The 3′ UTR plays a crucial role in (-)strand synthesis. For FMDV, an 'L' leader protein is present that is, in the case of aphthoviruses and erboviruses, a papain-like cysteine protease, or may have a different function, as is the case of kobuvirus, cardiovirus, teschovirus, sapelovirus and senecavirus (Palmenberg et al., 2010).

proteolysis (Cifuente and Moratorio, 2019). Due to the uncleaved VP0 being less well packed inside empty particles, and absent stabilising interactions with the packaged ssRNA that reinforce inter-pentameric attraction in mature virions, empty particles are invariably less thermostable than mature virions (Basavappa et al., 1994).

At the final stages of assembly, genomic RNA must somehow be packaged into the nascent capsid. Two possible models exist; first, genome is packed into a pre-formed empty 'pro-'capsid through a single packaging portal, and thus via an asymmetric packaging complex. In the second model, pentameric units condense around progeny genome, as observed in small plant viruses, Figure 1.3. This latter mechanism of assembly is popular amongst most ssRNA viruses (Perlmutter and Hagan, 2015), with many binding RNA via flexible, basic, terminal domains termed 'arginine rich motifs' (Speir et al., 1995). Such flexible domains have been observed in Ljungan virus, for instance, at the VP3 N-terminus, which, unlike most other picornavirus structures, was found to have significant observable ordered RNA in cryo-EM reconstructions (Zhu et al., 2015; Li et al., 2012a).

Picornavirus capsids adopt a simple icosahedral architecture, with the majority of mature capsids comprising of 60 copies each of the structural proteins VP1-4 arranged into an icosahedral shell, with pseudo T = 3 symmetry. Exceptions include the parechoviruses and kobuviruses, which maintain an uncleaved VP0 (Jiang et al., 2014). All consist of a common core comprising the major capsid proteins VP1-3. These proteins each adopt the wedge-shaped viral jelly-roll eight-stranded anti-parallel β-barrel motif (strands labelled B-I) that is common to the majority of RNA and DNA spherical virus capsid structures (Rossmann and Johnson, 1989; Harrison, 1992). Half of the strands (B, I, D, G) form an antiparallel sheet on one side of the barrel, facing the other four (C, H, E, F). The narrow end of the barrel wedge consists of BC, HI, DE and FG loops, which point towards the 5-fold axes in VP1; and alternate around the three-fold axes for VP2 and VP3.

The basic capsid building block contains five of these VP1-4 (VP0-3) protomers, arranged into a pentagon around the icosahedral 5-fold axis. Such pentamer assemblies, often sedimenting at 12-14S, have been isolated for a number of picornaviruses subjected to harsh treatments, such as low ionic buffers, or low pH (Li et al., 2012a; Malik et al., 2017). VP1 tends to be the most variable structural protein across serotypes, whilst VP4 is the most conserved. The external loops connecting β-strands form the antigenic signature of the capsid. Where present, the N-terminus



Javier O. Cifuente et al. J. Virol. 2013; doi:10.1128/JVI.03519-12

Figure 1.3: Putative assembly pathways for picornaviruses. Figure taken from Cifuente et al. (2013) and displayed with the permission of ASM via RightsLink.

of VP4 is myristoylated (Chow et al., 1987; Tuthill et al., 2010; Wang et al., 2012), perhaps aiding interaction with the endosomal membrane during genomic egress (see Section 1.1.3 below).

For the majority of enteroviruses, a hydrophobic pocket exists (for the purposes of this work, this pocket will be referred to as the 'P-site'), nestled within VP1, and sitting below the floor of the canyon-like depression that encircles the five-fold axis. X-ray, and, more recently, cryo-EM structures of enteroviruses, frequently show electron density commensurate with a aliphatic chain occupying this P-site. This so-called 'pocket factor', often modelled as a saturated fatty acid (see Table 1.5), seemingly mediates capsid stability, and, for many enteroviruses, its expulsion is an essential step in priming capsids for genomic release (Filman et al., 1989; Tsang et al., 2000; Smyth et al., 2003).

1.1.3 Capsid dynamics: Uncoating and Genome Release

Following entry into the cytoplasm, enteroviruses undergo structural transformations (termed 'uncoating') to states that have historically been identified according to their sedimentation coefficient. For the most extensively studied enteroviruses, such as PV, the mature virus particle transitions to an altered, so-called A-particle (135S) (Fenwick and Cooper, 1962; de Sena and Mandel, 1977; Belnap et al., 2000; Tsang et al., 2001), leading to exposure of the N-terminus of VP1, and release of VP4, which collectively facilitate penetration of the endosomal membrane (Fricks and Hogle, 1990; Casasnovas and Springer, 1994; Shafren et al., 1997; Arita et al., 1998; Pasch et al., 1999; Huang et al., 2000; Groppelli et al., 2017). An 'empty' (i.e. devoid of genome) B-particle (80S), marks the end product of this uncoating process, by contrast to some cardioviruses and aphthoviruses, where 12S pentamers have been isolated instead (Figure 1.4), presumably due to their more environmentally sensitive inter-pentamer contacts (Cartwright et al., 1980; Baxt and Bachrach, 1980; Curry et al., 1995). Interestingly, 12S particles have been isolated following low ionic buffer treatment of some enterovirus empty particles, such as EV-E1 (Li et al., 2012a).

The point at which enterovirus uncoating is initiated varies greatly between different species; while poliovirus receptor binding primes the capsid for genomic release, CV-A9 and E-1 are in fact stabilised by their receptor (akin to FMDV, *vide infra*), and instead seem to rely on cellular cues for A-particle transition, or a secondary receptor (Powell et al., 1997; Shakeel et al., 2013). Unlike aphthoviruses, most enteroviruses do not rely on endosomal acidification to initiate uncoating,

with the exception of many rhinoviruses (Bayer et al., 1999) and EV-D68 (Liu et al., 2018) (Table 1.1).

A plethora of potential uncoating or uncoating-like intermediates have been reported (Huang et al., 2000; Bubeck et al., 2005b; Levy et al., 2010; Garriga et al., 2012; Wang et al., 2012; Lyu et al., 2014; Ren et al., 2013; Butan et al., 2014; Dong et al., 2017; Liu et al., 2018). Many have been heat-induced, but, recently, some receptor- or antibody-induced intermediate structures have been reported, for instance, Lin et al. (2012); Strauss et al. (2013); Lee et al. (2016); Dong et al. (2017); Shah et al. (2020). Key features of A-particles identified by such studies include density within the capsid interior consistent with RNA, an opening at the icosahedral two-fold lined with patches of negatively charged residues, a collapsed P-site, and externalised VP1 N-termini and VP4 subunits (Lonberg-Holm et al., 1975; Huang et al., 2000; Bubeck et al., 2005b; Levy et al., 2010; Bostina et al., 2011; Garriga et al., 2012; Wang et al., 2012; Shingler et al., 2013; Ren et al., 2013).

Breathing

At physiological temperatures, some picornaviral capsids, for instance, PV, HRV14, HRV16, SVDV, CV-A9, have been shown to be dynamic; undergoing a reversible externalisation of VP4 and N-terminal VP1 in a process termed 'breathing' (Li et al., 1994; Lewis et al., 1998; Lin et al., 2012; Roivainen et al., 1993; Lewis et al., 1998; Pulli et al., 1998; Jimenez-Clavero et al., 2000; Katpally et al., 2009). Similar such behaviour has been seen in other families such as bromoviruses and tombusviruses (Witz and Brown, 2001; Speir et al., 2006). For picornaviruses, this is assumed to be a general phenomenon, but may transiently mirror some processes that seem to occur during the early stages of uncoating; for instance, externalisation of the VP1 shown with a CV-A16 uncoating intermediate seems to occur in a similar fashion to breathing, but is more permanent (Ren et al., 2013).

1.1.4 Receptor binding and Entry

To infect a suitable host cell, picornaviruses must first engage with a specific receptor. Reflective of their varied tissue tropism, different picornaviruses engage with a plethora of receptors with a variety of binding epitopes (see Table 1.1). For some picornaviruses, receptor binding causes structural changes to potentiate signalling for internalisation, and prime for RNA egress (see Fig-

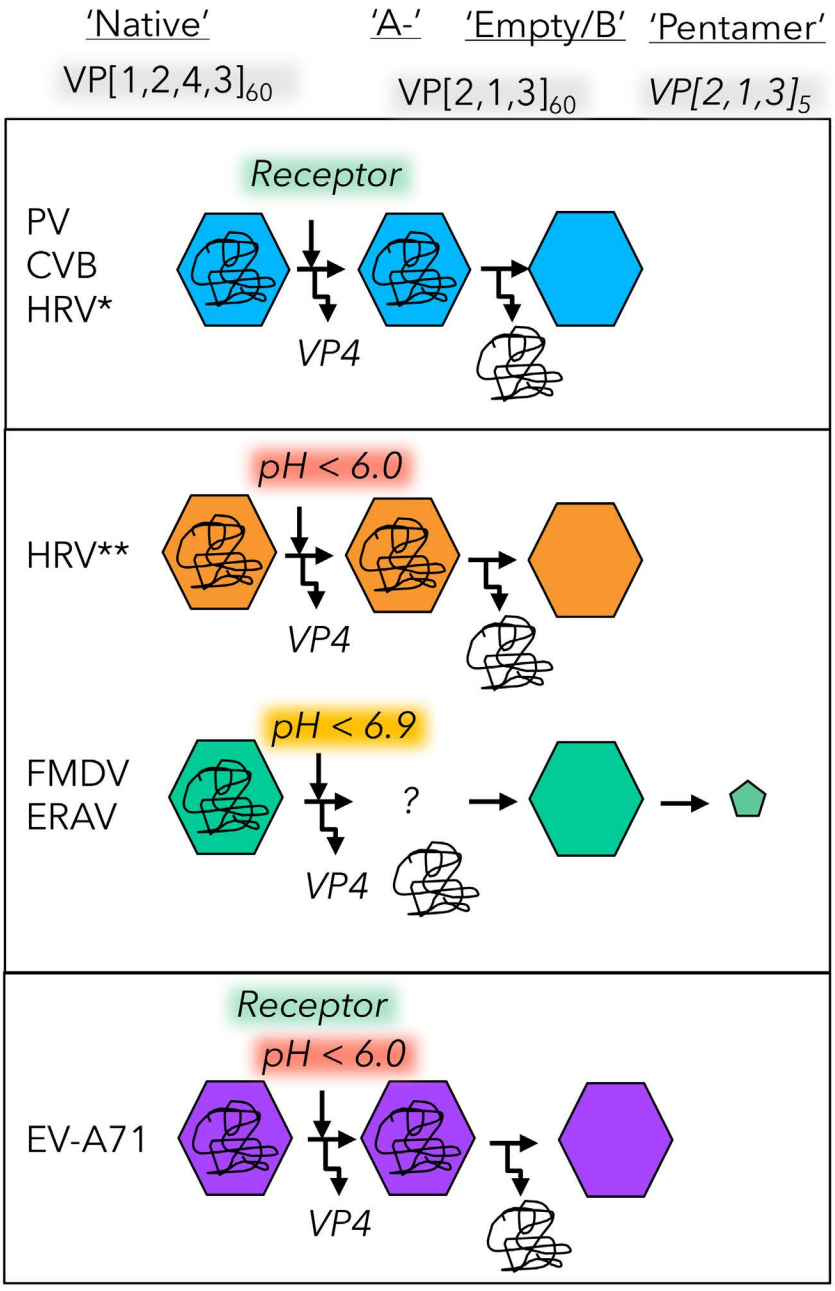


Figure 1.4: Uncoating mechanisms for different picornaviruses. For PV, CV-Bs and some, hence asterisk, HRVs, uncoating is initiated by the receptor binding event, leading to the expulsion of VP4, and often a pocket factor, where present, before genomic release, leading to an empty particle end-product (Levy et al., 2010; Hendry et al., 1999). For other HRVs, pH is instead the trigger, as is also the case for FMDV and ERAV (Giranda et al., 1992; Groppelli et al., 2010; Martín-Acebes et al., 2010). It is unknown whether FMDV forms A- and/or B-particles. EV-A71 uncoating is initiated by receptor binding and pH (Shingler et al., 2013).

ure 1.4). It has been suggested that capsid rigidity may an essential determinant for receptor induced transition to an A-particle; PV, for instance, is relatively rigid; surface loops are well resolved in crystallographic structures and capsids can withstand high pressures, by contrast to that of FMDV (Hogle et al., 1985; Filman et al., 1989; Oliveira et al., 1999). A summary of the entry and assembly processes for different genera is shown in Figure 1.4.

Following attachment, non-enveloped picornaviruses encased in endosomal vesicles enter the cell cytosol and are transported to the replication site. Internalisation has traditionally been thought to occur via the Clathrin-dependent pathway, but, as entry processes have become better understood for a greater number and range of picornaviruses, it has become apparent that entry may instead occur via a number of alternative Clathrin-independent routes, such as caveolar/lipid raft-mediated endocytosis and macropinocytosis (Pietiäinen et al., 2005; Coyne et al., 2007; Brandenburg et al., 2007; De Palma et al., 2008; Heikkilä et al., 2010).

Receptor preferences can switch readily; very minor changes in surface configurations, as little as a single amino acid, can alter receptor binding capabilities (McLeish et al., 2012), while preferred routes of entry are highly dependent on the cell line studied and not the target receptor. Integrin-binding picornaviruses are an excellent example of this latter point. Enteroviruses (CV-A9, E-9) and parechoviruses (human parechovirus 1, HPeV-1), as well as the aphthovirus FMDV, specifically engage with α_v integrins coupled with $\beta_{1,3,6}$ via an RGD motif (Nelsen-Salz et al., 1999; Heikkilä et al., 2009; Seitsonen et al., 2010; Merilahti et al., 2012). E-1 also engages with an integrin, but not via an RGD motif, rather with α_2 associated with a β_1 subunit in an inactive (see below) state, causing a clustering of integrins, which may trigger signalling for endocytosis (Heikkilä et al., 2010). Certain isolates of CV-A9, and tissue culture adapted FMDV can also utilise heparan sulphates for entry (Fry et al., 1999b; McLeish et al., 2012).

Genomic delivery into the cytosol must occur via either direct penetration of a host membrane, or a sub-viral particle vehicle in order to protect the genome. Historically, a five-fold axis channel was speculated as being the portal for genome release, based on low resolution EM maps (Hewat et al., 2002) and comparisons with the bacteriophage phiX174 (McKenna et al., 1992; Smyth and Martin, 2002), but studies have shown that, whilst pentameric sub-units undergo limited rearrangements compatible with particle expansion, an opening at the 2-fold axis channel is often observed and genomic egress seems to occur at a site between the 2- and quasi-3-fold axes (Levy

Table 1.1: Receptor preferences for a small collection of picornaviruses. Trigger refers to the environmental changes that initiate the uncoating process. HS is heparan sulphate; tc stands for tissue culture adapted.

Virus	*Genus*	*ssRNA length / bases*	*Main Receptor(s)*	*Receptor Binding site*	*Uncoating Trigger*	*Example PDB*	*Citation*
FMDV	Aphtho	8300	αvβ3, αvβ6, HS (tc)	VP1 GH loop	pH	4GH4	Curry et al. (1996)
ERAV	Aphtho	7839	Sialic acids	VP1, close to EF loop from adj protomer	pH	2WFF	Fry et al. (2010)
Poliovirus	Entero	7441	CD155	Canyon	Receptor	1HXS	Miller et al. (2001)
HRVs	Entero	7200	ICAM-1 (HRV 14,16); LDL-R (HRV2)	Canyon; shallow pits near 5-fold	Receptor or pH	4PDW (1AYM)	Lacroix et al. (2014), (Oliveira et al., 1993)
CVA9	Entero	7452	αvβ3, αvβ6	CT VP1	Not pH, not receptor	1D4M	Hendry et al. (1999)
EMCV	Cardio	7800	VCAM-1		pH	2MEV	Krishnaswamy and Rossmann (1990)
EV-A71	Entero	7400	SCARB2, SiA	Canyon	pH	3VBS	Wang et al. (2012)
EV-F3	Entero	7430	SiA	see EV-F3 Chapter	Unknown, not pH	5OSN	Roedig et al. (2017)

et al., 2010; Bostina et al., 2011; Wang et al., 2012; Shingler et al., 2013; Lyu et al., 2014; Mullapudi et al., 2016; Lee et al., 2016).

Oligosaccharides as Receptors

Oligosaccharide chains, often sulphated or sialylated glycans, displayed on cell surface glycoproteins, act as receptors for many viruses, including influenza, rotavirus and encephalomyocarditis virus (Tavakkol and Burness, 1990).

Sialic acids comprise derivatives of neuraminic acid, often found as the outermost nine carbon sugar of cell-surface glycoproteins, attached, via an α-2,3 or α-2,6 linkage, to an underlying galactose or N-acetyl galactosamine. Tissue culture adapted strains of EMCV, a cardiovirus, bind glycophorin A via a sialic acid moiety, and enter cells via a sialic acid-mediated pathway. ERAV, an aphthovirus, also binds sialic acid (PDB: 2XBO) ((Warner et al., 2001; Stevenson et al., 2004; Fry et al., 2010), while enteroviruses CV-A24v and EV-70 utilise glycan-containing receptors for attachment (Zocher et al., 2014). However, of the picornaviruses that have been shown to bind oligosaccharide chains (see Table 1.2), few have functional evidence indicating their use of sialic acids as a receptor (Zhou et al., 2000; Lipton et al., 2006; Tsunoda et al., 2009).

FMDV, an aphthovirus, TMEV, a cardiovirus, and some echoviruses (enteroviruses) (Goodfellow et al., 2001) recognise heparan sulphate as a receptor, whilst some EV-A71s can utilise sialylated glycoproteins (Yang et al., 2009; Tan et al., 2013). Attachment to sugar moieties often occurs following tissue culture adaptation, e.g. for FMDV and EMCV (Jackson et al., 1996; Guy et al., 2009).

Integrins as Receptors

Integrins are a homologous family of divalent cation dependent receptors that are fundamental mediators of cell adhesion to the extracellular matrix. In fact, the name 'integrin' derives from the their role in linking the extracellular matrix with the cytoskeleton (Tamkun et al., 1986). They are non-covalently associated heterodimers of type I transmembrane glycoprotein α and β subunits found on the surface of most cells (Hynes, 2002). There are 18 α and 8 β subunits in mammals, which together form 24 canonical integrins that act as bi-directional messengers, contributing to the regulation of numerous cellular processes (Hynes, 2002; Barczyk et al., 2010).

Table 1.2: Picornaviruses that utilise sialic acids for cellular entry.

** 2,3 possible, but less preferable ; ** low neurovirulent strains only.*

Virus	Linkage	Publication	PDB
ERAV	2,3	Fry et al. (2010); Stevenson et al. (2004); Warner et al. (2001)	2XBO
Mengovirus	2,3	Anderson and Bond (1987)	None
CV-A24v	2,6 *	Zocher et al. (2014)	4Q4X
EV-A70	2,3	Nokhbeh et al. (2005)	None.
TMEV	2,3 **	Zhou et al. (2000)	None.
EV-F1		Stoner et al. (1973)	None.
EMCV		Tavakkol and Burness (1990)	None.
EV-D68	2,3; 2,6	Liu et al. (2015); Nilsson et al. (2008)	5BNN;5BNO;5BNP
Porcine Sapelovirus	2,3	Kim et al. (2016)	None.

Integrins generally appear as a large 'head' that contains sites for ligand binding and subunit association, propped on two 'legs'. Each α and β subunit contains a large multi-domain N-terminal extracellular ectodomain, a single-helical transmembrane domain, and a relatively compact cytoplasmic domain. Ectodomains contain three β sandwich domains; an Ig-like thigh, and two 'calf' domains, whilst 200 amino acid α I-domains and β I-like domains associate to form the ligand-binding head (see Figure 1.5) (Dickeson and Santoro, 1998).

A highly promiscuous subset comprise α_v, which can associate with several different β subunits, including β_3 and β_6. $\alpha_v\beta_6$ is commonly found in epithelial cell lines, and recognises fibronectin as its natural ligand (Busk et al., 1992), while $\alpha_v\beta_3$ typically recognises vibronectin (Pytela et al., 1985; Smith and Cheresh, 1990).

Each heterodimer can undergo substantial conformational switching, a process that is mediated by metal ions. Structural and binding studies have shown the so-called 'extended' conformation occurs with Mn^{2+}, and has a high affinity for ligands, in contrast to the 'bent' conformation (Takagi et al., 2002). Structural biological investigations into integrins, even of truncated constructs, has been challenging owing to their substantial structural heterogeneity. Indeed, although metal ions can serve to shift the conformational equilibrium in preference of one pose over another, this shift is never total. For instance, for $\alpha_v\beta_3$, it was found that with Ca^{2+} treatment, which reportedly shifts the conformation towards the inactive, compact form, circa 14 % of molecules remained extended (Ohi et al., 2004). Similarly, for Mn^{2+}, the extended:compact ratio was only 80:20 (Ohi et al., 2004). High affinity ligands can, however, force the equilibrium to ca 98

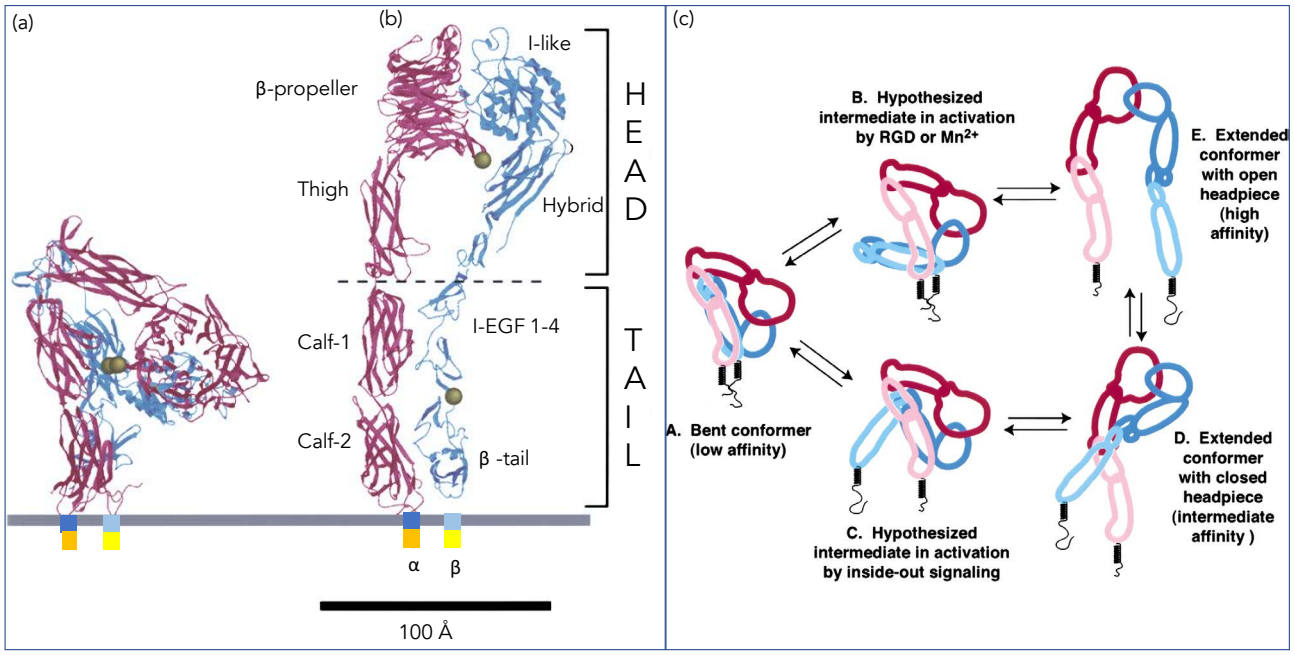


Figure 1.5: (a) Model of ectodomain derived from crystallographic structure of $\alpha_v\beta_3$ (PDB: 1JV2) showing I-EGF domains 1,2 and part of I-EGF3 for the beta moiety, (Xiong et al., 2001) and (b) a model proposed by Takagi et al. (2002) of the extended conformation. The grey line shows the approximate position of the cell membrane, and boxes show, schematically, the likely positioning of the trans and cytoplasmic portions for these conformations, not to scale. The black scale bar refers to the scale for the crystallographic structure on the left. (c) Proposed quaternary rearrangements as proposed by Takagi et al. (2002) based on structures of EM studies, excepting the hypothetical forms that are purely author conjecture. Both figures have been taken and adapted from Takagi et al. (2002), with permission from Elsevier via RightsLink.

Table 1.3: Locations of RGD motif(s) for picornavirus capsid proteins in the P1 polyprotein, with associated UniProt accession numbers. Produced using Uniprot (Consortium, 2018) and Merilahti et al. (2012).

Type	Accession Number	Location of RGD motif in viral polyprotein	
		VP3	VP1
CV-A9	D00627		858-860
E-5	AF083069	410-412	
E-9	X84981		651-653/860-862
EV-B83	AY843301		832-834
EV-B99	JF260926		860-862
FMDV-O	Q6PMW3-1		869-863
HPeV-1	L02971		764-766
HPeV-2	AJ005695		763-765
HPeV-4	DQ315670		767-769
HPeV-5	AF055846		772-774
HPeV-6	AB252582		766-768

% (Takagi et al., 2002). Crystallographic structures of integrin constructs exclusively show integrins in the 'bent', inactive, form, such as in Xiong et al. (2001); Dong et al. (2012).

Integrins have been identified as key cellular receptors for adeno-, hanta-, herpes- and rotaviruses, a likely reflection of their abundance in most cell types (http://www.signaling-gateway.org/, search 'integrins'), and crucial role in cellular signalling (Pellinen and Ivaska, 2006; Merilahti et al., 2012). Picornaviruses engaging with heterodimers containing α_v monomers generally do so via an RGD motif, see Table 1.3.

1.1.5 Metal Ions and Capsid Stability

Metal ions play an essential role in mediating capsid integrity across the viral kingdom. For instance, turnip rosette virus, sobemoviruses and sesbania mosaic virus capsids are stabilised by metal ions, and their removal likely pre-empts uncoating (Hull, 1977; Savithri and Erickson, 1983; Satheshkumar et al., 2004). Metal ion additives, such as Na^{+}, have also been shown to help maintain vaccinia virus infectivity following heat treatment to 50 °C, probably due to stabilisation of the capsid against denaturation and not encapsidated DNA, since heavier ions, such as Co^{2+} and Fe^{2+}, greatly reduced infectivity as did lower pHs of 6.0 (Kaplan, 1963).

The role of metal divalent ions in picornavirus stability, including FMDV, has long been established. Dialysis of FMDV against EDTA has been shown to trigger genome release (Rowlands et al., 1975), a process that does not occur when in complex with its integrin receptor, whilst incubation with Mg^{2+} and Ca^{2+} ions appears to enhance capsid thermostability (Kotecha et al., 2016). Some years ago, crystallographic data from FMDV crystals soaked with terbium, sometimes used as a Ca^{2+} surrogate (Brittain et al., 1976), included electron density at the three-fold axes that was interpreted as 20 % occupied terbium ions (Logan, 1991).

One putative metal ion binding site for some picornaviruses, especially human rhinoviruses, is the 5-fold axis channel that spans the capsid shell (Hadfield et al., 1997). For rhinoviruses, coxsackieviruses and phiX174, metal ions have been modelled towards the periphery of the channel, but, for CV-A9, ions have been modelled deep within the channel instead (Smyth and Martin, 2002).

Of the 83 deposited PDB structures that are flagged when searching for picornavirus capsid structures [Text Search for: picornavirus and TAXONOMY is Viruses and TAXONOMY is only Viruses and SYMMETRY is Icosahedral], only 12 have metal ions modelled into the structure. Four

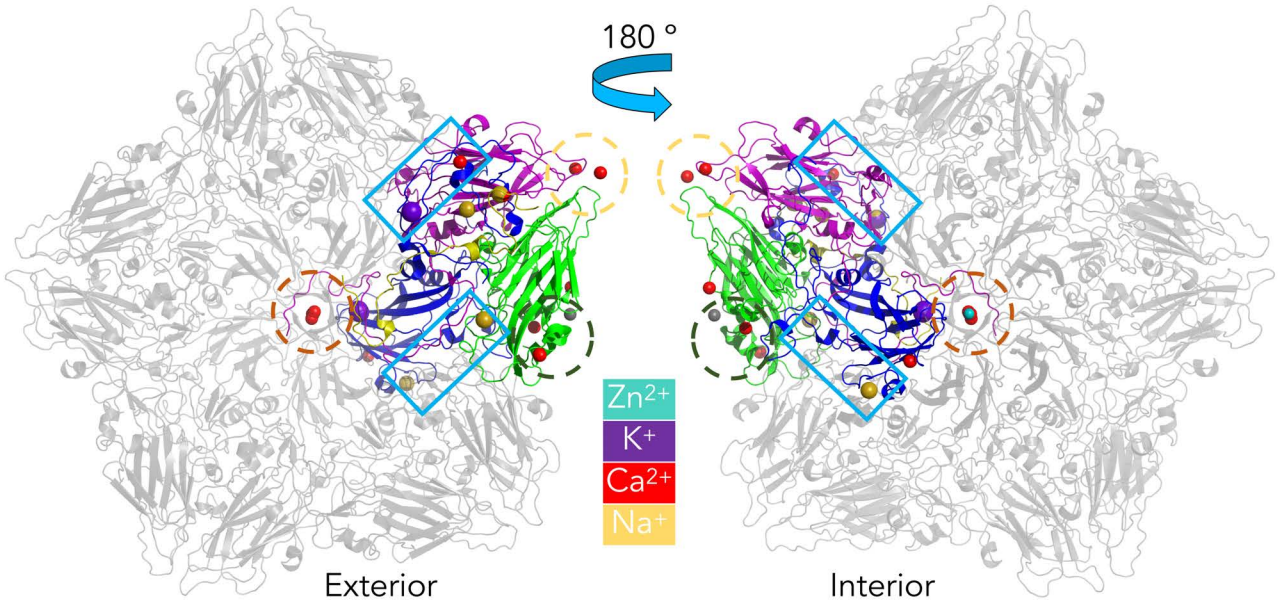


Figure 1.6: Metal ions for picornaviruses modelled in PDB structures mined as described in the main text, mapped onto an EV-F3 pentamer (model from Chapter 4, section 1.0.2). All metal ions shown are confined to a protomer, which is coloured by structural protein. Typically, metal ions are modelled in one of four different 'types' of sites, shown as shapes; blue box: inter-protomer regions, yellow dashed circle: 3-fold axis plug; red dashed circle: 5-fold axis plug, and, finally, green dashed circle: 2-fold bridge. See the key at the centre of the figure for metal ion colours.

of these deposited structures contained K^+ ions and one (HRV16) with Zn^{2+} (Hadfield et al., 1999; Oliveira et al., 1993), while Ca^{2+} ions have been modelled for human rhinoviruses HRV14, HRV1A, and HRV3 at the 5-fold axis (Rossmann et al., 1985; Kim et al., 1989; Zhao et al., 1996). An additional Ca^{2+} ion has been proposed in HRV14 at the icosahedral 3-fold axis (Rossmann et al., 1985; Zhao et al., 1997), and electron density in CV-B3 at the 3-fold and 5-fold axes has been proposed to be from Ca^{2+} ions (Muckelbauer et al., 1995). Some evidence of a Ca^{2+} ion at the N-terminal region of VP1 (close to S1021) in CV-A21 has also been presented (Xiao et al., 2005a). Collectively, modelled metal ions in picornavirus structures seem to adopt three distinct regions, outlined in Figure 1.6.

1.1.6 Enteroviruses: Poliovirus and Bovine Enterovirus

Enteroviruses are the causative agents of a spectrum of diseases; ranging from mild rashes and the common cold, to viral meningitis and paralysis (Tuthill et al., 2010), of which fifteen species, termed EV-(A-L), RV-(A-C), have been identified (Racaniello, 2007; Zell et al., 2017). For picornaviruses, species members are distinguished by high amino acid identity for P1, 2C, 3C and 3D, almost identical genome maps and belong to the same clade. Members of a given genus tend to fulfil three criteria: (a) homologous 2B, 3A, 3B and, if present, L, non-structural proteins; (b) Over 36 % amino acid identity for the non-structural proteins 2C and 3CD as well as 33 % P1 amino acid identity; (c) be of the same monophylletic group in phylogenic trees (Zell et al., 2017). Virions are generally transmitted via enteric and respiratory routes, but can spread to other organs, such as the central nervous system or the heart (Zell et al., 2017).

Polioviruses (PV), the best known members of the human enterovirus species C, are the causative agent of poliomyelitis disease. PV gained prominence in the 1950s, causing numerous cases of paralysis amongst children, perhaps as a consequence of improved hygiene preventing early PV exposure in passively immune infants (Landsteiner, 1909; Skern, 2010). Global immunisation initiatives with live attenuated oral and/or inactivated vaccines have made a substantial impact (see Figure 1.7); diminishing global cases by 99 % since 1988. Cases, however, persist in disadvantaged regions, such as Papua New Guinea and Somalia, and PV-1 is still endemic in Afghanistan, Nigeria and Pakistan (Zambon and Martin, 2018) (Figure 1.8).

Bovine enteroviruses (BEV) span species EV-E and -F and are endemic in cattle in many regions

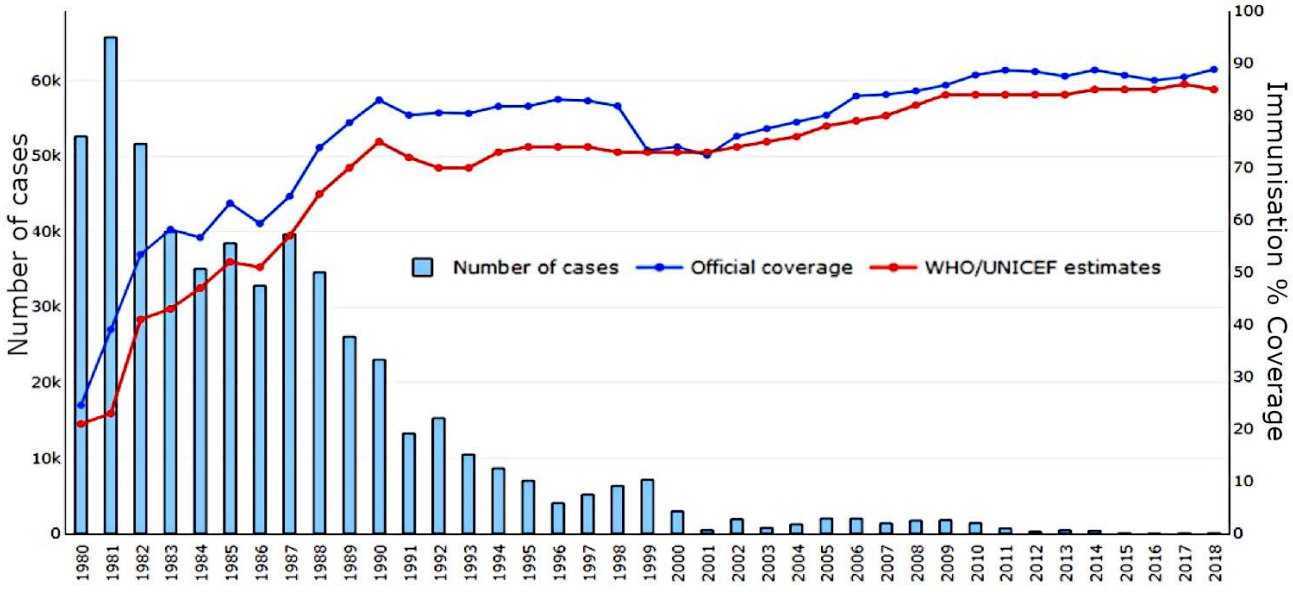


Figure 1.7: Polioylitis global annual reported cases (bar graph) and immunisation coverage (1980-2018) with oral polio vaccine or inactivated polio vaccine (Pol3), taken from www.who.int/immunization/monitoring_surveillance/burden/vpd/surveillance_type/active/poliomyelitis .

(Jiménez-Clavero et al., 2005). Infections are often asymptomatic, although they have been associated with diarrhoea and abortions, and clinical symptoms are hard to reproduce in experimental infections of animals (Jiménez-Clavero et al., 2005). Thus, BEVs are excellent surrogates for more obviously pathogenic enteroviruses (Ley et al., 2002; McCarthy et al., 1999; Zhang et al., 1990).

Classification and Structure

Poliovirus

PV exists as three serotypes; PV-1, PV-2 and PV-3 (Ansardi et al., 1996).

Two distinct antigenic forms have been identified for PV, termed 'D' (or 'N', native) (Mayer et al., 1957; Le Bouvier, 1959) and 'C' (or 'H', heated) (Le Bouvier et al., 1955; Hummeler et al., 1958), with the latter being non-infectious. Typically, in wild-type preparations, D is predominant, albeit to varying degrees, and switching to C antigenicity is observed upon heating (hence 'H'). Only D provides a significant neutralising antibody response and thus is essential for vaccine candidates (Schwerdt and Fogh, 1957; Le Bouvier et al., 1955).

PV capsids are broadly similar in structure to other picornaviruses, especially enteroviruses, but

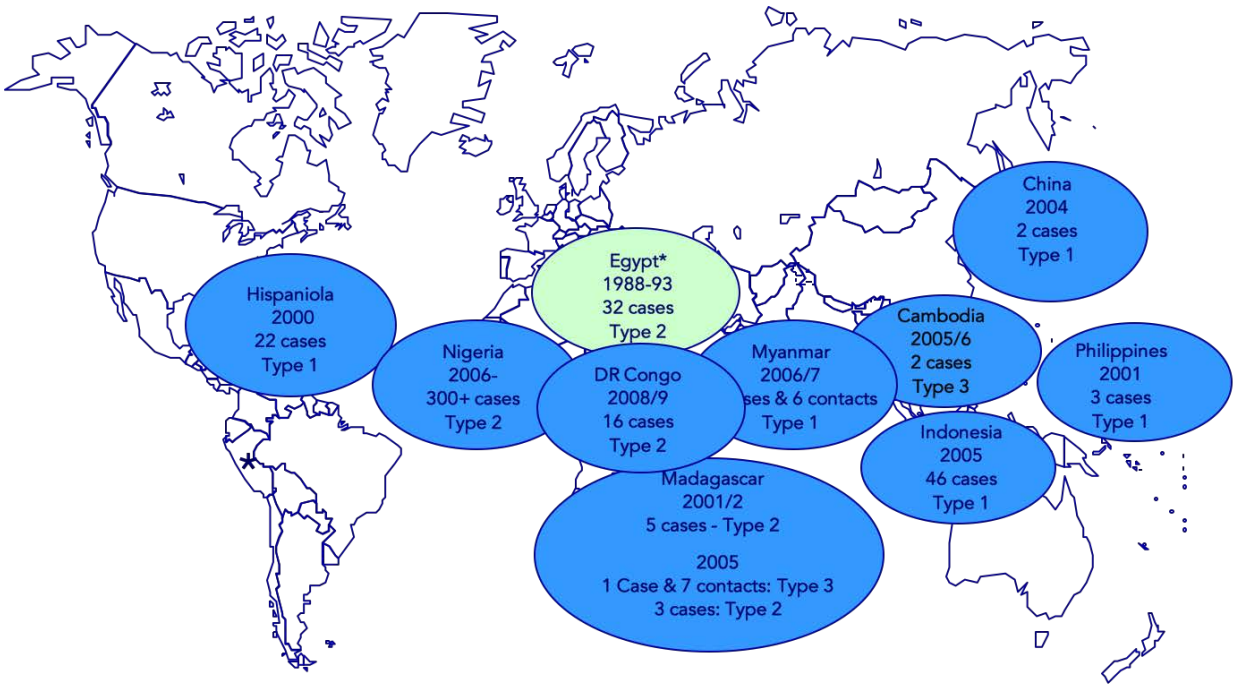


Figure 1.8: Summary of PV cases reported by WHO in 2009. Data from www.who.int.

with prominent, star-shaped mesas at the 5-fold axes, shaped by the VP1 BC, HI and DE loops, and three-bladed propeller-like structures at the 3-fold axes (Figure 1.9(a)).

EV-F3

Until relatively recently, BEV was segregated into two serotypes; '1' (BEV-1) and '2' (BEV-2), updating an earlier definition of eight serotypes (Dunne et al., 1974; Knowles and Barnett, 1985). However, just over a decade ago, a review of a body of BEV sequences revealed significant divergences between sequences in all viral proteins. This prompted calls for a reclassification into two different clusters in line with Knowles and Barnett's previous classification; BEV-A (née BEV-1) and BEV-B (née BEV-2), but with each containing two and three geno-/serotypes, respectively (Zell et al., 2006). All studies in this thesis focus on the sole member of the third group of BEV-B, of enterovirus F species, and thus termed EV-F3. Crystallographic structures for serotypes from each of these two species (EV-E1 and EV-F3) have been deposited (PDB: 1BEV, 5OSN) (Smyth et al., 1995; Roedig et al., 2017), but no published structural description exists for EV-F3.

EV-E1 has lower buoyant density than rhinoviruses and aphthoviruses and higher physical sta-

bility than some other enteroviruses (Moll and Davis, 1959; Rowlands et al., 1971; Smyth et al., 1995).

Emblematic of the Enterovirus genus to which they belong, both BEV species contain a P-site within VP1, which, in the case of EV-E1, has been shown experimentally (in the only published robust analysis of enterovirus pocket factors) to harbour an array of short chain fatty acid molecules, including palmitic acid (C16:0) and myristic acid (C14:0) (Smyth et al., 2003). For EV-F3, sphingosine, a popular choice for enteroviruses, including PV (see Table 1.5, section 14.1), has been previously modelled (Roedig et al., 2017).

Receptor Binding

Enteroviruses frequently engage with receptors via the so-called 'canyon' that encircles the five-fold axis (see Table 1.1 in Section 1.1.4) (Rossmann et al., 1985). For poliovirus, binding to CD155 (or 'PVr', Poliovirus receptor) initiates the transition to an A-particle intermediate primed for genomic release, whereas for both CV-A9 and E-1, receptor binding stabilises the virus structure, and virions instead rely on cellular cues to trigger conformational change (de Sena and Mandel, 1977; Shakeel et al., 2013). Efforts to structurally characterise the receptor for EV-F3 are described in Chapter 4.

Two key uncoating intermediates, or more accurately, structures pertaining to the canonical 135S intermediate (A-particle) and the end-product empty 80S (B-particle), have been structurally characterised for a number of enteroviruses, for instance, see Figure 1.9 (Lonberg-Holm et al., 1975; Huang et al., 2000; Bubeck et al., 2005b; Levy et al., 2010; Bostina et al., 2011; Garriga et al., 2012; Wang et al., 2012; Shingler et al., 2013; Ren et al., 2013). A and B particles share similar antigenicity that is distinct from native virus (Fenwick and Cooper, 1962; Ren et al., 2013). For EV-A71, B-particles are expanded, and almost identical to 'natural' empty particles from immature capsids (Wang et al., 2012). Contrastingly, for CV-A16, with 80 % sequence identity with EV-A71, B-particles may be less stable than natural empties that are non-expanded and almost identical to 135S-like species (Ren et al., 2015).

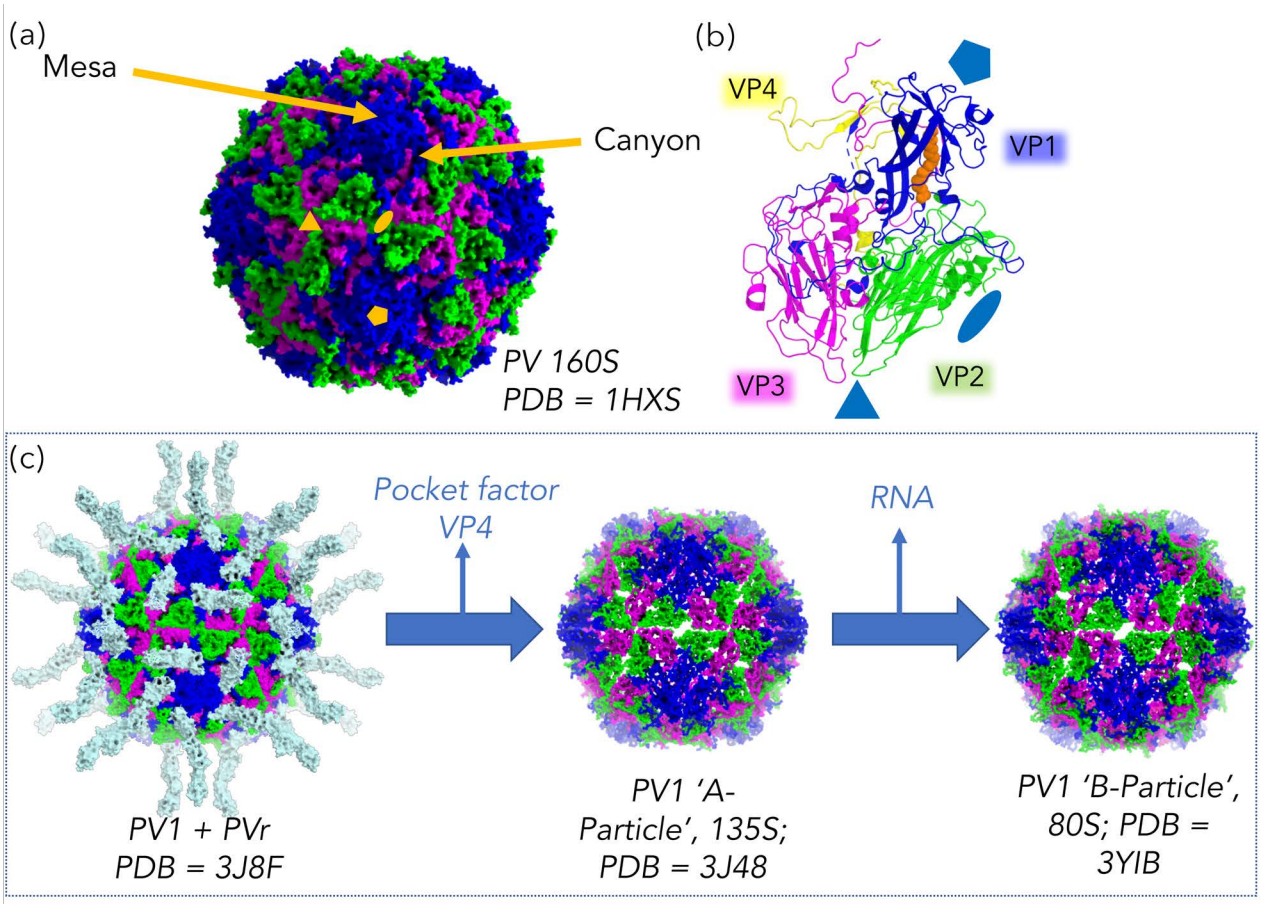


Figure 1.9: Overview of PV structural components. (a) 160S mature virion capsid surface, with 2-, 3-, and 5-fold axes shown as orange shapes. (b) Protomer coloured by structural proteins: VP1 blue; VP2 green; VP3 magenta and VP4 yellow. Axes to show context within the mature virion are shown as blue shapes. The pocket factor, that sits within the VP1 is shown as orange spheres. (c) Structures of PV, following the entry pathway. Note that alternative PV structures are available on the PDB. First, PV engages with its cellular receptor, PVr (pale cyan) via a capsid canyon (in reality, particles would not be completely decorated with receptor). This receptor binding event triggers the formation of an A-particle, with the release of VP4, expulsion of the pocket factor and externalisation of the N-terminal region of VP1, which serves to form a pore across the endosomal membrane through which the genome may be released. An empty, 'B-particle' has been isolated which is believed to be the favoured particle structural form following genomic egress.

1.2 Aphthoviruses: Foot and Mouth Disease Virus

1.2.1 Overview

Foot-and-mouth disease (FMD) is one of the most infectious animal pathogens, causing vesicular disease of cloven-hooved animals and remains endemic in many parts of the world, especially third-world countries where the disease exacerbates the 'Hunger Gap' (Grubman and Baxt, 2004). Although rarely lethal, outbreaks in countries where the disease is not endemic have a substantial economic impact due to productivity losses and early mortality of young animals (Jamal and Belsham, 2013). A significant proportion of this thesis will be dedicated to studies pertaining to FMDV (Chapters 5, 7).

1.2.2 Genome and Structure

FMDV comprises an 8.3 kb, single-stranded positive-sense RNA genome, encased within a 25-30 nm diameter protein capsid (Sangar et al., 1987; Clarke et al., 1987). There are seven antigenically distinct FMDV serotypes: O, A, C, Asia 1, Southern African Territory (SAT)-1, SAT-2 and SAT-3, each of which comprise many subtypes (Pereira, 1981; Robson et al., 1977; Domingo et al., 2005). Structural phylogenetic relationships between published structures from some of these serotypes may be found in Figure 1.10.

FMDV's P1 region is preceded by a leader protein (L) (Sangar et al., 1987), a feature present in cardio- and hepatoviruses, but absent in entero- and rhinoviruses (Dorner et al., 1982), which is highly variable between serotypes (George et al., 2001) (Figure 1.2, Secion 1.1.2). Aphthoviruses also contain, within the same reading frame, two AUG codons, 84 nucleotides apart, that initiate translation *in vivo* and *in vitro*.

The first atomic resolution structure of a FMDV (type O1BFS) was published in 1989, and, since, then, near-atomic resolution structures of several serotypes (Acharya et al., 1989; Lea et al., 1994; Curry et al., 1996; Fry et al., 1999b, 2005; Reeve et al., 2010) have been published. Despite this, our understanding of FMDV at a basic structural level remains limited. Mature FMDV capsids comprise VP1-4 and adopt classic picornavirus architecture (Figure 1.11). VP4, the most conserved capsid protein, shows 81% sequence conservation across serotypes, whilst VP1, crucial for serotype specificity and receptor attachment, is highly variable in terms of length (209-213 amino acids) and

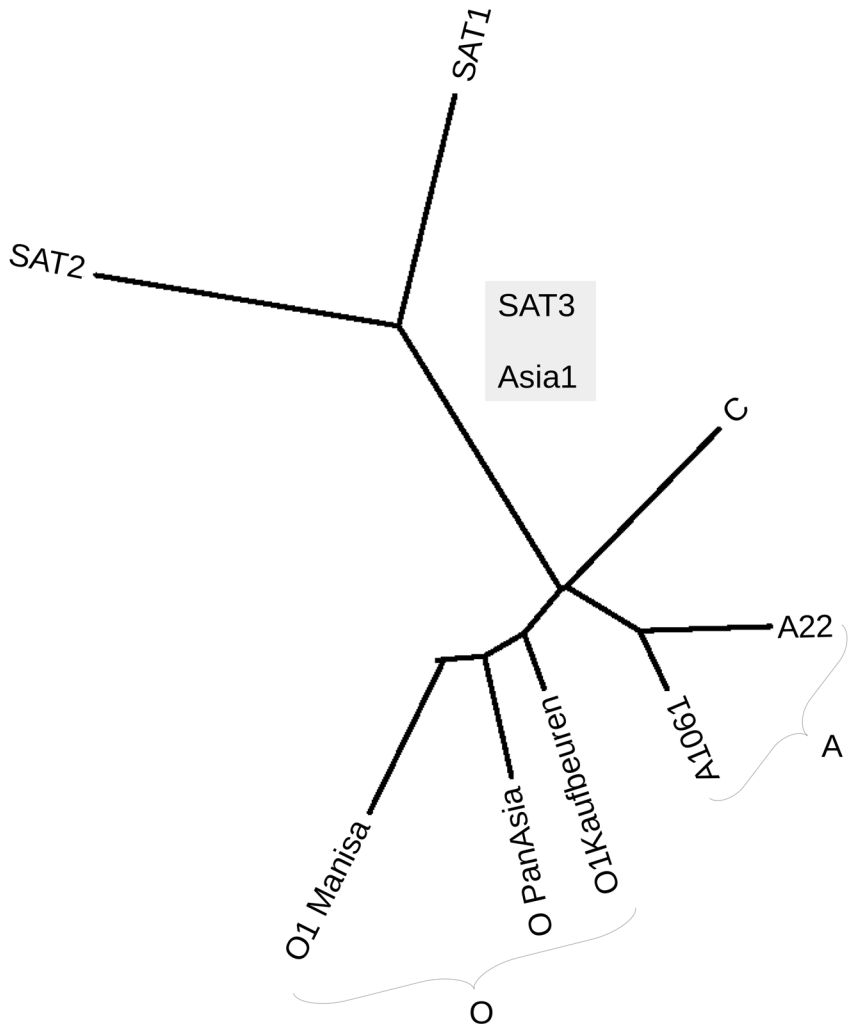


Figure 1.10: Structural phylogenetic tree generated using SHP (Stuart et al., 1979) with published FMDV structures, PDB codes clockwise from top left: 5ACA, 2WZR, 1FMD, 4GH4, 1ZBE, 1BBT, 5NED, 5NEJ, 5AC9 (Fry et al., 1993; Lea et al., 1994; Curry et al., 1996; Fry et al., 2005; Kotecha et al., 2015, 2017). No structures have been deposited on the PDB for Asia1, nor SAT3 serotypes.

content across serotypes (Feng et al., 2003). VP2 and VP3 are important in virion stability and VP2 and VP4 for maturation (Carrillo et al., 2005).

Compared to enteroviruses, FMDV's capsid exterior is mostly smooth, lacking the canyon around the 5-fold axes. Rather than being hidden from immune surveillance, the RGD host receptor recognition motif is on the exposed ca 30 residue flexible VP1 'G-H' loop (residues 1145-1147 in type O), a neutralisation site for the host humoral immune response (Logan et al., 1993; Neff et al., 1998; Jackson et al., 2000b; Berryman et al., 2005; Burman et al., 2006).

FMDV capsids lack a 'P-site' since, unlike enteroviruses, the VP1 β-barrel is filled with bulky aromatic side chains. Capsids also lack the additional VP1 strand of the stabilising extended β-sheet across the pentamer interface compared to enteroviruses, and this, combined with the high concentration of histidine residues at this interface, seems to impart extreme thermo- and acid-lability (Newman et al., 1973, 1977; Curry et al., 1995; Wachsman et al., 1998). FMDV degrades into pentameric units upon short exposures to pHs of 6.5 and below (Cartwright et al., 1980; Curry et al., 1995; van Vlijmen et al., 1998; Martín-Acebes et al., 2010, 2011).

1.2.3 Antigenic Sites

A number of antigenic sites have been identified for different FMDV serotypes, mainly through sequencing mutants that escape neutralising monoclonal antibodies for serotypes: A (Thomas et al., 1988; Baxt et al., 1989), C (Mateu et al., 1990) and O (Parry et al., 1989; Kitson et al., 1990; Crowther et al., 1993). Site 1 or A, formed from part of the VP1 GH-loop, has historically been viewed as immunodominant, but recent studies have shown no such immunodominance, at least not for O_1 serotypes (Aggarwal and Barnett, 2002). Note that a fifth site has been identified that seems to be unique to type O (Kitson et al., 1990). Major sites are summarised in Table 1.4.

Cross-protection by eliciting neutralising antibodies to pan-serotype conserved epitopes offers an alternative vaccine strategy to multivalent vaccines. The relatively lower cross-serotypic variability of VP2-3 (when compared to VP1) means that the antigenic sites at VP2-3 are more likely to be cross-reactive than those on VP1 (Logan et al., 1993; Marquardt et al., 2000; Mateu et al., 1990; Salem et al., 2019).

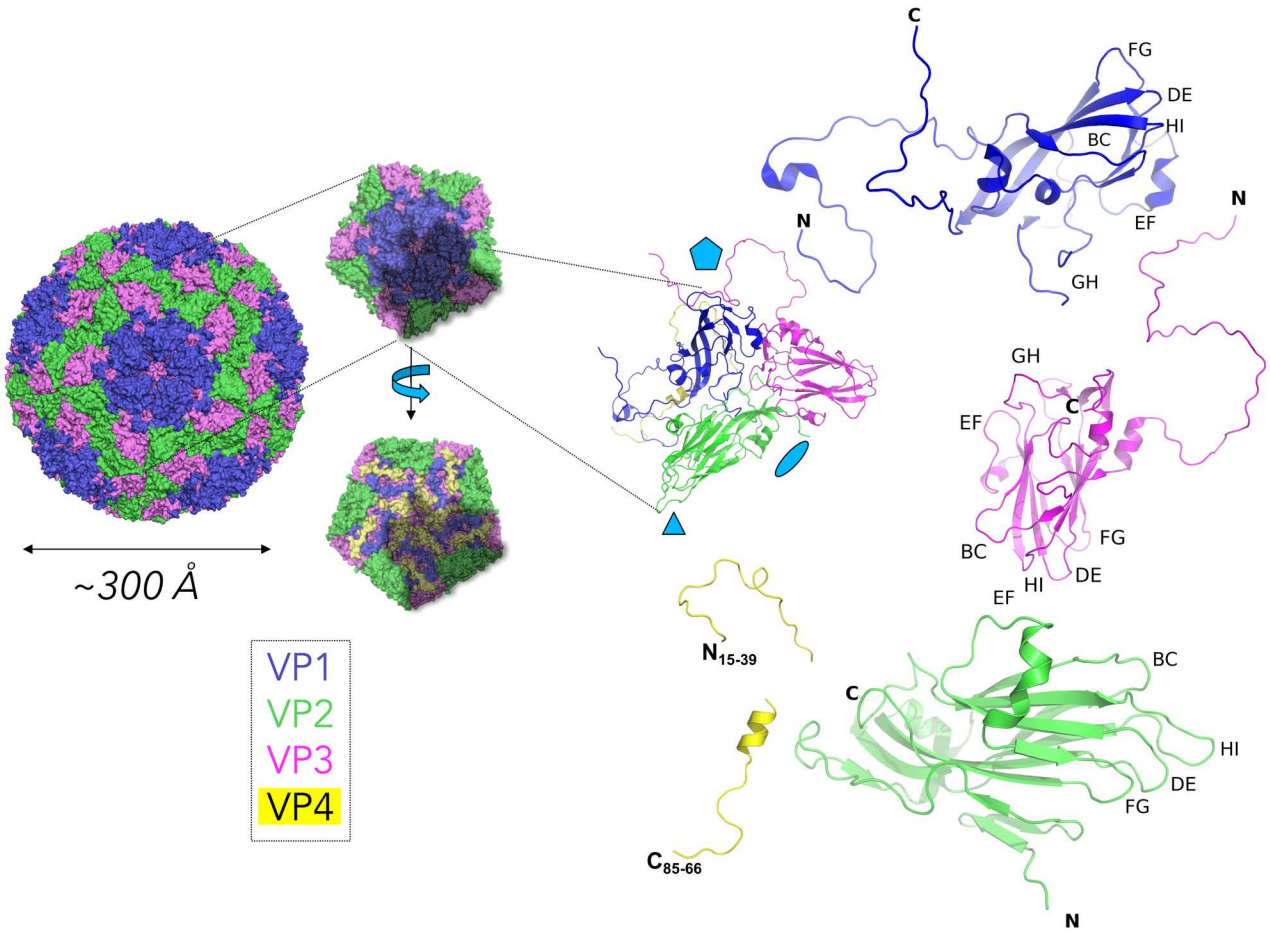


Figure 1.11: Overview of FMDV structure. Far left, capsids are of icosahedral symmetry, containing structural proteins VP1 (blue), VP2 (green), VP3 (magenta) and VP4 (yellow). The pentamer is the stable building block of capsids, comprising of five protomers, with one protomer shaded in grey. The inner surface of the pentamer (below twisted arrow) shows VP4 (yellow), which traverses the capsid interior. Surrounding the protomer, zoomed in views of the structural proteins VP1-4 are shown, with loops labelled based on the strands with which they are associated, and N and C termini labelled in bold typeface.

Table 1.4: Antigenic sites or loops identified for the majority of FMDV serotypes, based on sites listed in Reeve et al. (2016), but supplemented with material from Grazioli et al. (2013), Thomas et al. (1988) and Baxt et al. (1989). Note that, for a number of cases, studies have been conducted on a limited number of subtypes; 'A' serotype sites, for instance, have been characterised for A10 and A12 subtypes (Baxt et al., 1989; Thomas et al., 1988). 'CT' means C-terminus, and the 'Site Number' is based on sites defined by Grazioli et al. (2013).

Site Number	Approximate residue position	Serotypes where residue is known to be antigenic	Serotypes where loop known to be antigenic	Loop Name
2	2072	O A C SAT1 Asia1	SAT2	BC
2	2074	O Asia1 C	A SAT1 SAT2	BC
	2193	O	A	HI
3	3056	O	A C Asia1	BB knob
	3072	SAT1	A	BC
	3077	SAT1	A	
	3138	SAT1	A	EF
4	3223	Asia1		
	1047	O		
1	1142		O A SAT1 SAT2 Asia1	GH
1	1144	SAT1 C Asia1	O A SAT2	GH
1	1149	SAT1 A Asia1	O C SAT2	GH
1	1164	SAT2	O SAT1 A C Asia1	GH
1	1169		O SAT1 SAT2 A C Asia1	HI
	1209		O A C SAT2	*CT*
	1211	O A C	SAT2	*CT*

1.2.4 Receptor binding

For FMDV, receptor binding events are mediated by the long VP1 GH-loop that attaches to an integrin receptor (Logan et al., 1993; Mason et al., 1994; Jackson et al., 1997, 2000a,b; Burman et al., 2006; Kotecha et al., 2017). Amongst the four integrin species that reportedly mediate FMDV infection *in vitro*, $\alpha_v\beta_6$ has the highest affinity for FMDV in cell culture (Berryman et al., 2005) and is expressed in epithelial cells, thereby correlating with FMDV's tissue tropism (Jackson et al., 2000a), analogous to CV-A9 of enterovirus B species (Shakeel et al., 2013). Structures of FMDV (PanAsia and O1M serotypes) bound to $\alpha_v\beta_6$ ectodomain show this binding event to cause no distortion of the capsid exterior, except for the GH-loop of VP1. Incubation with $EDTA^{2+}$ has been found to destabilise non-receptor associated FMDV capsids, but not when associated with $\alpha_v\beta_6$, suggesting that receptor interactions somehow stabilise the capsid (Rowlands et al., 1975; DiCara et al., 2008), but whether this stabilisation effect extends to mildly acidic environs remains unknown.

Uncoating is triggered following endosomal transport into the cell by acidification of the early endosome (Groppelli et al., 2010). It was long thought that rhinoviruses, cardioviruses, and aphthoviruses, dissociated directly to pentamers (Brown and Cartwright, 1961; Tuthill et al., 2009), rather than via an A-particle intermediate (Baxt and Bachrach, 1980; Tuthill et al., 2009). However, isolation of empty ERAV particles at low pH in the presence of BSA and 2 mM $CaCl_2$ (known to stabilise PV A-particles (Tuthill et al., 2006)), has brought this theory into question (Tuthill et al., 2009).

1.3 Therapeutic Applications

1.3.1 The Quest for Broad-Spectrum Antivirals

Much work has focussed on developing treatments for picornavirus infected individuals that cater for a broad-range of sub-and sero-types so as to alleviate symptoms, but also to control the spread of infection by reducing virus excretion and shedding to the environs. Five key stages have been targeted (see Figure 1.12): (a) Entry; (b) Translation; (c) Proteolytic processing; (d) Replication; (e) Assembly (Kim et al., 2012; Benschop et al., 2015).

Candidate antivirals must not disrupt vaccine responses, and should be robust against resistance. ssRNA virus RNA-dependent RNA polymerases (RdRps) lack proofreading abilities, meaning that mutation rates are high and resistance can occur rapidly (Ferrer-Orta et al., 2015; Benschop et al., 2015). For instance, a P-site inhibitor, V-073, that was on trial for PV eradication, was no longer effective after two mutations were identified *in vitro* (I1194F, 1192 for PV-3 and A3024V) (Kouiavskaia et al., 2011; Liu et al., 2012d).

Many groups are exploring the use of anti-enteroviral preparations to protect against outbreaks of PV, while EV-A71 anti-virals are being investigated in conjunction with vaccine development, as reviewed in Tan et al. (2014). Many existing drugs, such as ribavirin (treatment for HCV, Lee et al. (1998)), appear to have inhibitory effects on PV replication in addition to EV-A71, E-5 and E-18 (Crotty et al., 2000; Park et al., 2011a,b; Zhang et al., 2012), but their already established use may accelerate resistance in cases where individuals are unknowingly co-infected with EVs (Benschop et al., 2015).

Entry: Pocket Factors

As discussed above, the P-site is emptied prior to uncoating and so, in the quest for broad-spectrum anti-enteroviral, much work has focussed on finding a more potent P-site binder so as to inhibit this initial step (Figure 1.12(a)), with some success (Rombaut et al., 1990; Lewis et al., 1998; Martikainen et al., 2015). WIN compounds (originally developed by the Sterling-Winthrop Research Institute, hence 'WIN') are perhaps the most widely studied of P-site inhibitors, whilst Pleconaril may be the most extensively structurally reported (Lewis et al., 1998; Abdel-Rahman and Kearns, 1998; Pevear et al., 1999). For some enteroviruses, conformational changes imposed by such P-site binders abrogate engagement with cellular receptors (Pevear et al., 1989).

Replication: Arbidol

One recent example of a very broad-spectrum replication inhibitor is Arbidol (ARB, umifenovir, ethyl-6-bromo-4-[(dimethylamino)methyl]-5-hydroxy-1-methyl-2-[(phenylthio)methyl]-indole-3-carboxylate hydrochloride monohydrate), an anti-influenza drug that also shows antiviral activity against numerous enveloped and non-enveloped viruses, including some picornaviruses such as PV-1, PV-3, CVA-B5, FMDV and ERAV (Pécheur et al., 2016). Its mode of action appears to be via interference

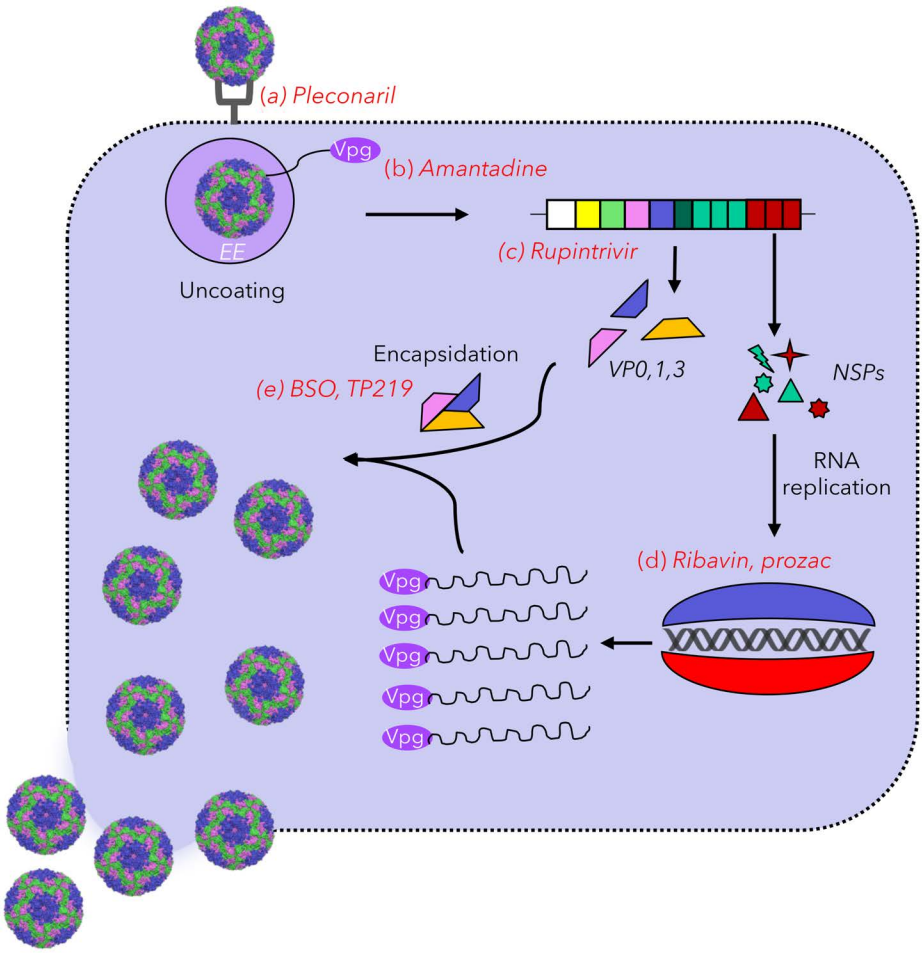


Figure 1.12: Schematic showing the major stages of the picornavirus infection process. The red text gives examples of antiviral candidates that have been shown inhibit picornavirus infection at a given stage. (a) Entry has been blocked for many enteroviruses P-site binders, such as Pleconaril (Pevear et al., 1999). Following receptor binding and entry via an often cell-type-dependent mechanism, the virion is trafficked into the cytoplasm encased in an early endosome (EE). At some point, either in the late or early endosome, the virion releases its precious genome into the cytoplasm. This is then (b) translated, forming the P1-P2-P3 polyproteins. Translation inhibitors include Amantadine which inhibits IRES-dependent translation and has been FDA approved for influenza (not recommended) and Parkinson's disease only and has been shown to be effective against E5, E18 and EV-A71, but not PV (Benschop et al., 2015). Polyproteins then undergo (c) proteolytic processing. 3C-protease inhibitors include Rupintrivir, which failed at Phase II clinical trials (Hayden et al., 2003). (d) RNA is then replicated ready for encapsidation into progeny virions. Replication ihibitors include Ribavirin and Prozac. Ribavirin targets 3Dpol and its incorporation into the RNA chain disrupts nucleoside pairing and has show effectivity against PV and EV-A71 infection as well as some echoviruses (Zhang et al., 2012; Crotty et al., 2000; Lee et al., 1998). (e) Assembly of virions has been shown to be inhibited by GSH inhibitors such as BSO and TP219 (see main text) (Smith and Dawson, 2006; Thibaut et al., 2011, 2014). Finally, newly formed mature virions exit, possibly by cell lysis.

Table 1.5: Natural pocket factors modelled for enteroviruses on the PDB. Methods listed are either X-ray crystallography (xt), or Cryo-electron microscopy (em). The 'FA chain length' refers to the number of carbons in the main chain of the modelled fatty acid. Metal ions 'M ions' modelled are listed in their atomic form. For clarity, abbreviated enterovirus names are used (see abbreviations section).

Virus	PDB	Reference	M ions	Method	Pocket factor	Length
BEV-1	1BEV	Smyth et al. (1995)	-	xt	Myristic acid	14
EV-F3	5OSN	Roedig et al. (2017)	K^+	xt	Sphingosine	18
CV-A9	1D4M	Hendry et al. (1999)	-	xt	Sphingosine	18
CV-A10	6ACU	Zhu et al. (2018)	-	em	Sphingosine	18
CV-A16	5C8C	Ren et al. (2015)	K^+	xt	Stearic acid	18
CV-A16	5C4W	Ren et al. (2015)	K^+, Na^+	xt	Sphingosine	18
CV-A16	5C9A	Ren et al. (2015)	K^+	xt	Sphingosine	18
CV-A21	1Z7S	Xiao et al. (2005a)	Ca^{2+}	xt	Myristic acid	14
CV-A24v	4Q4Y	Zocher et al. (2014)	Ca^{2+}, Mg^{2+}	xt	Myristic acid	14
CV-B3	3JD7	Lee et al. (2016)	-	em	Palmitic acid	16
EV-A71	4N53	Lyu et al. (2014)	-	xt	Sphingosine	18
EV-A71	3ZFE	Plevka et al. (2013)	Na^+	xt	Sphingosine	18
EV-A71	4AED	Plevka et al. (2012)	Ca^{2+}	xt	Lauric acid	12
EV-A71	3VBH	Wang et al. (2012)	K^+, Na^+	xt	Sphingosine	18
SVDV	1OOP	Fry et al. (2003)	-	xt	Sphingosine	18
E-1	1EV1	Filman et al. (1998)	-	xt	Palmitic acid	16
PV-1	1HXS	Miller et al. (2001)	-	xt	Palmitic acid	16
PV-1	1AR6	Wien et al. (1997)	-	xt	Sphingosine	18
PV-1	1AR7	Wien et al. (1997)	-	xt	Sphingosine	18
PV-1	1AR8	Wien et al. (1997)	-	xt	Sphingosine	18
PV-1	1AR9	Wien et al. (1997)	-	xt	Sphingosine	18
PV-1	1ASJ	Wien et al. (1997)	-	xt	Sphingosine	18
PV-1	1AL2	Wien et al. (1997)	-	xt	Sphingosine	18
PV-3	1PVC	Filman et al.1995	-	xt	Sphingosine	18
HRV2	3DPR	Querol-Audí et al. (2009)	Ca^{2+}	xt	Lauric acid	12
HRV2	1V9U	Verdaguer et al. (2004)	Ca^{2+}	xt	Lauric acid	12
HRV2	1FPN	Verdaguer et al. (2000)	-	xt	Lauric acid	12
HRV16	1ND2	Zhang et al. (2004)	Zn^{2+}	xt	Myristic acid	14
HRV16	1AYN	Oliveira et al. (1993)	Zn^{2+}	xt	Lauric acid	12
HRV16	1AYM	Oliveira et al. (1993)	Zn^{2+}	xt	Lauric acid	12

of intracellular membrane sites used for replication (Herod et al., 2019).

Entry and Assembly: Glutathione and Viruses

The tripeptide glutathione (GSH) is the most abundant thiol peptide found in mammalian cells; occurring in upwards of 1 mM quantities. Through scavenging reactive oxygen species (ROS), GSH plays a crucial role in protecting cellular components against oxidative stress, signal transduction, gene expression and apoptosis. Upset of the delicate balance between GSH and its oxidised form, GSSG, has been associated with multiple diseases, including cancer, and some viral infections, such as those caused by HIV-1 and influenza (Herzenberg et al., 1997; Cai et al., 2003; Bansal and Simon, 2018).

GSH synthesis is a two-step process; during the first and rate limiting step, γ-glutamylcysteine synthase catalyses the formation of Cys-Glu. Glycine appendage then occurs, catalysed by glutathione synthase. L-butathione sulfoximate (BSO) is known to inhibit the initial formation of Cys-Glu (Figure 1.13), and, as-such, has been used as a tool to probe the role of glutathione in viral entry and replication. Studies utilising BSO, and a related inhibitor, TP219 (which forms covalent linkages with free GSH, thereby depleting available intracellular GSH, Figure 1.13), have shown that GSH depletion leads to a blockage of protomer assembly into pentamers for CV-B3, PV and E-9 (Thibaut et al., 2014).

Such observations hint at GSH's potential as the basis of a broad-spectrum enteroviral formulation. Thibaut et al. (2014) identified a number of GSH-resistant mutations that reside at the protomer interface, and, based on this, proposed that GSH binds within a partly-solvent exposed cleft at the inter-protomer interface, where it may confer its stabilising effect against capsid dissociation (Thibaut et al., 2014).

In addition to playing a structural role, GSH's mediation of the intracellular oxidative environment has also been shown to impact picornavirus assembly. The herbal plant derivative isochlorogenic acid C ICAC is a di-O-caffeoyl derivative of chlorogenic acid (CHA) that can be used to modulate glutathione (reduced/oxidised glutathione, GSH/GSSG) redox homeostasis by interacting with several antioxidant enzymes that impact the GSH/GSSG balance (Cheng et al., 2014; Ho et al., 2008; Cao et al., 2017) and treatment of EV-A71 infected mice with ICAC was shown to relieve symptoms (Cao et al., 2017). It has been shown that EV-A71 infection triggers an accumu-

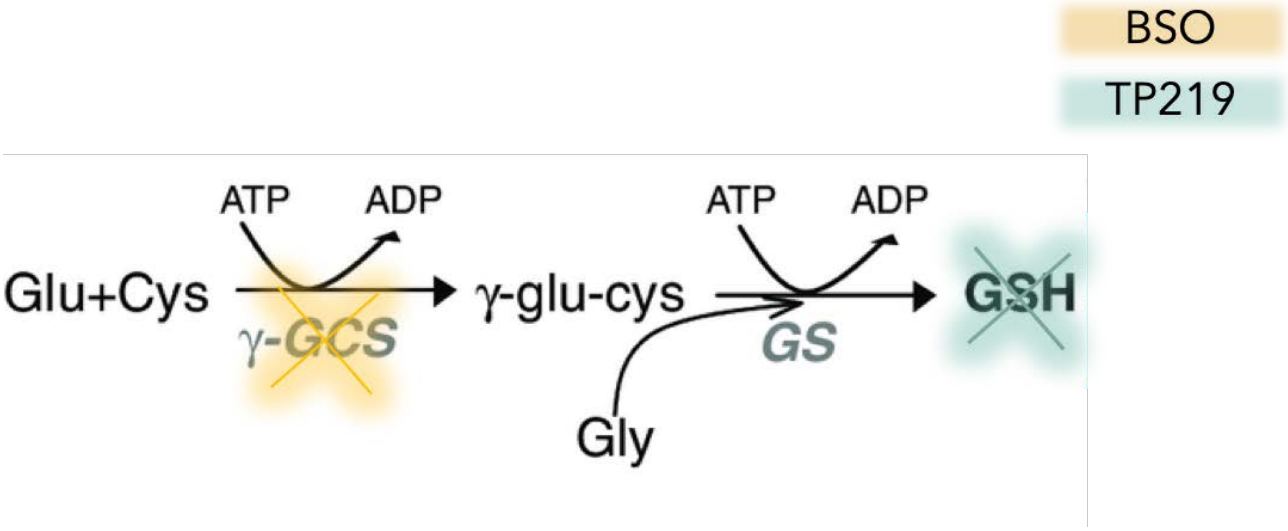


Figure 1.13: Glutathione biosynthesis is conducted via two ATP-dependent stages. The first step involves the synthesis of γ-glutamylcysteine from L-glutamate and cysteine, catalysed by the enzyme γ-glutamylcysteine synthetase (γ-GCS), and is the rate-limiting step. Glycine is then appended to the C-terminus of the γ-glutamylcysteine dipeptide by glutathione synthetase (GS). The stages that glutathione inhibitors TP219 and BSO work to reduce intracellular gluathione levels are shown as coloured crosses.

lation of ROS (Abdel-Moneim et al., 2018), and similar effects have been observed for CV-B3 (Yu et al., 2013) and HIV (Robinson et al., 1996).

1.4 Vaccination

RNA viruses tend to have a significantly higher number of quasi-species, relative to their DNA virus counterparts, due to the lack of proof reading functionality of RNA-dependent RNA polymerase enzyme (RdRp). As such, substantial antigenic variation is observed between even different serotypes and subtypes of the same virus (Ferrer-Orta et al., 2015). This greatly impacts on vaccine efficacy and thus preparations with multiple epitopes is often necessary.

Traditionally, vaccines have been prepared using live, or inactivated virus, requiring high containment facilities and extensive testing (Cottam et al., 2008). A raft of alternative methods have been explored as potential vaccination preparations, but much of the vaccine-related studies in this thesis have focussed on the use of 'Virus-Like Particles'.

1.4.1 Virus Like Particles

The phrase 'Virus like Particles', or VLPs, was first coined in the early era of electron microscopy to describe particulate bodies resembling virus size and shape in papillomas, hepatomas and tumours. Today, the term VLPs refers to supramolecular complexes of self-assembled viral proteins (usually produced by recombinant expression) into virus-like capsid structures virtually identical to their native counterparts, but devoid of genome and incapable of propagation, although they have been known to package host cell genome during expression.

VLPs are non-infectious, but typically retain the unique conformational epitopes displayed by their native counterparts, and thus are excellent candidates for some structural and functional studies (D'Argenio and Wilson, 2010). Since they are recombinantely expressed, careful sequence manipulations enable major modifications to the external VLP surface to probe native functions or within the bulk capsid to enhance stability (Schwarz et al., 2005; Brune et al., 2016).

VLP utility in the vaccine field has been widely recognised, with seven vaccines currently approved for human use, and multiple others in pre-clinical phase or clinical trials (Bachmann and Jennings, 2010; Pascual et al., 2017). Despite numerous hopeful vaccine candidates, only few have been licensed as vaccine products to date, notably VLPs of human papillomavirus (Paavonen et al., 2007) and hepatitis B (Liu et al., 2012a,c).

A vast array of expression systems have been tested for picornavirus VLP production, each with their own advantages and disadvantages. Regardless of the expression system, the majority of VLP expression strategies involve either (i) the co-expression of both the P1 and 3C or (ii) 3CD proteins, or (iii) expression of P1 and 3C separately, or the co-expression of VP0, VP1 and VP3 separately (Chung et al., 2006; Wang et al., 2016; Fox et al., 2017; Marsian et al., 2017; Yang et al., 2019). Recently, a novel system was reported expressing VP1, VP0, and VP3-2A to overcome cytotoxicity associated with 3C protease (Xu et al., 2019).

Chapters 5 and 6 present further details on FMDV and PV VLP vaccine development.

1.5 Aims

The aims of the work described in this thesis are as follows; to properly structurally characterise a model enterovirus, EV-F3, to try to further understand the antigenic stability of FMDV and po-

liovirus VLPs and, finally, set out a framework for understanding the uncoating process of aphthoviruses.

Following a description of methods used in this work, efforts to understand the mystery of the stability of EV-F3 are described, including the identification of a putative receptor, and of a druggable site. Structures of FMDV vaccine candidates with an engineered disulphide bridge will then be presented, in addition to an un-anticipated medium-resolution pentamer structure, and a brief description of provisional studies into the rapid characterisation of VLP stability.

PV VLPs have also been a major research interest in our group, and some provisional cryo-EM analysis of PV-1 and PV-3 candidates are shown in Chapter 6.

Finally, aphthovirus uncoating remains mysterious, and, in the final Chapter, I describe some foundations laid for future studies into the structural basis of uncoating, specifically through expressing integrin receptors that may be used for an *in vitro* system, along with controlled acidification and intermediate stabilisation. I show, with a low-resolution structure of ERAV, the potential for aphthoviruses to adopt an expanded empty particle state that is distinct from previously characterised structures.

Much of the presented in this thesis reflects my own journey to getting to grips with structural biology and virology in general, but I hope that it also has some value for others.

Chapter 2

Materials and Methods

2.1 Molecular Biology

2.1.1 Solutions and Media

All glassware for solution preparations were autoclaved prior to use.

Buffers were sterilised either by autoclaving to over 100 °C, or by filtering through a 0.22 μm filter (Merck-Millipore). LB for cloning was made from pre-prepared powder (Thermo Fisher), diluted to 25 g/L. LB-Agar plates were made using pre-prepared LB powder and agar mixed in a 2:1 ratio with MQ water prior to autoclaving at > 100 °C. For all cloning experiments, ampicillin or carbampicillin antibiotics, stored at -20 °C as 1000 x stocks and working concentrations (100 μg/mL) were added to LB or LB-agar plates immediately prior to use.

Mammalian tissue culture media and Dulbecco's phosphate-buffered saline (DPBS) were purchased from suppliers stated in Sections 2.1.3-2.1.5 below.

2.1.2 Cloning

Polymerase Chain Reaction and Restriction Digestion

Primer oligonucleotides were ordered from either Integrated DNA Technologies (IDT), or Sigma Aldrich. In each case, lyophilised DNA was resuspended in Tris-EDTA (TE) buffer to a concentration of 100 μM, from which aliquots of 10 μM working stocks were prepared with sterilised distilled water.

Figure 2.1: Typical PCR workflow used with thermocycler. Sometimes a 'gradient' of temperatures was used at the annealing step for a given fragment. The central three stages were cycled 20-25 times, often depending on the fragment length.

Step	Temperature / °C	Time / s
Initial denaturation	94	60
Denaturation	98	10
Annealing	55*	60
Elongation	72	30 * kB
Hold	4	∞

20-25 x

Amplification of target fragments was performed by PCR (Saiki et al., 1985; Mullis et al., 1986). For the majority of PCR reactions presented in this work, Pyrobest^{TM} DNA polymerase was used (Takara). Some early experiments utilised Phusion High-Fidelity DNA polymerase 5 x mastermix (New England Biolabs, NEB), but this was found to be less specific. For almost all cases, 50 μL of PCR reactions were prepared on ice to minimise non-specific amplification, with sterilised distilled water, 10 X Pyrobest Buffer II, dNTP Mixture (2.5 mM of each dNTP), template DNA (typically around 500 ng) and, usually, 1 μM of each primer (from a 10 μM stock solution). DNA polymerase was always added last, after dNTP addition, to avoid primer digestion by 3' 5' exonuclease activity of the DNA polymerase, before placement in a pre-heated PCR machine.

The typical thermal cycling protocol is shown in Table 2.1. The elongation step was adjusted to the construct length, and, in some cases, a range of temperatures were screened based on the recommended annealing temperature for a given primer pair (provided by IDT).

The majority of cloning performed was ligation dependent, and, in every case, restriction enzymes were obtained from NEB, and were, where possible, 'high-fidelity', meaning that they allegedly maintain activity over a greater range of conditions so as to reduce off-target products.

Agarose Gel Electrophoresis

Agarose gel electrophoresis was used to monitor the success of cloning, and to separate PCR and/or restriction digest products. Gels were typically prepared with 1 % w/v agarose powder in Tris-acetate-EDTA (TAE) buffer, but, in cases where much shorter DNA fragments needed to be resolved, this was increased to 2 % w/v agarose, which was then heated until dissolved. After

cooling for 1-2 min, SYBR Safe DNA gel stain (Thermo Fisher) was added, and the gel poured into the desired mould. Once set, analyte DNA was loaded using a DNA loading dye (NEB) and separated by electrophoresis at 100 V until the lower dye band was within a third of the end of the gel. Gels were then visualised on a UV transilluminator at 302 nm, before transfer to a UV box and desired bands excised for purification.

Gel extracts were then purified using a Qiaquick purification kit (Qiagen).

Transformation of Competent *E coli* DH5α cells

Transformations were performed using sub-cloning efficiency DH5α cells (Thermo Fisher) which were divided into 50 μL aliquots upon receipt and stored at -80 °C. Competent cells were thawed on ice before, typically, 1-5 μL of plasmid DNA was added (typically a range of DNA concentrations were trialed for ligations, and only a small quantity was used for re-transformation) and incubated on ice for 30 min. Cells were then heat shocked for 40 s in a water bath set to 42 °C, before recovery on ice for 5 min. 200 μL Super Optimal broth with catabolite repression medium (SOC) was added to cells, and then transferred to a shaking incubator at 180 rpm, 37 °C for 1 hr. Cells were then plated onto LB-agar plates with ampicillin.

Preparation of Plasmid DNA

A single colony from a transformed LB-agar plate, prepared as above, was extracted with a pipette tip and used to inoculate 5 mL LB medium with ampicillin. This was incubated overnight at 180 rpm, 37 °C. The following day, cultures were spun down for 10 min at 4 °C, 3500 g, and the supernatant discarded. Pellets were then purified as described below, or frozen at -20 °C.

Purification of Plasmid DNA

All plasmid purifications were performed using either a QIAprep Spin Miniprep Kit (Qiagen), or, in the case of the transient $\alpha_v\beta_6$ and $\alpha_v\beta_3$ systems, a QIAgen Gigaprep Kit (Qiagen).

Inserts were checked by sequencing (GATC).

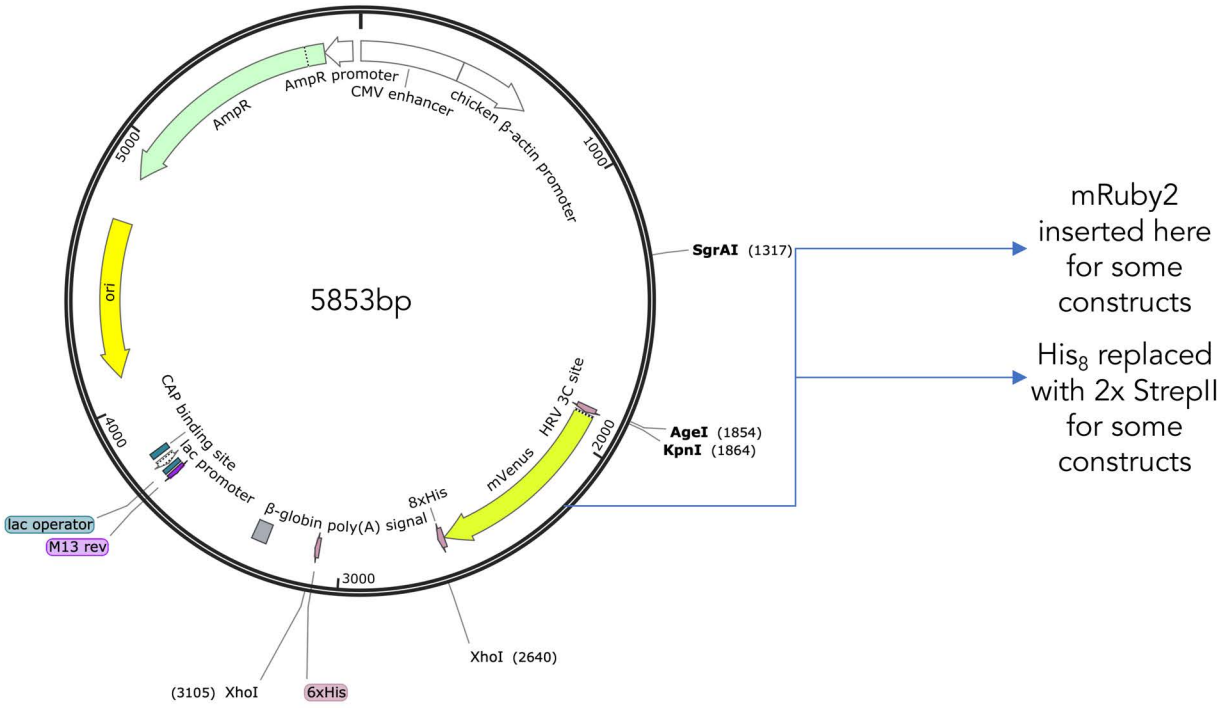


Figure 2.2: Schematic of pHLSec vector used for transient expression. Note that some variants were used, where the MV tag was replaced with mRuby2 fluorescent tag, and the 8xHis site was replaced with a twin strepII, or 1D4 tags. In some early cases, the fluorescent tag was removed altogether.

Vectors

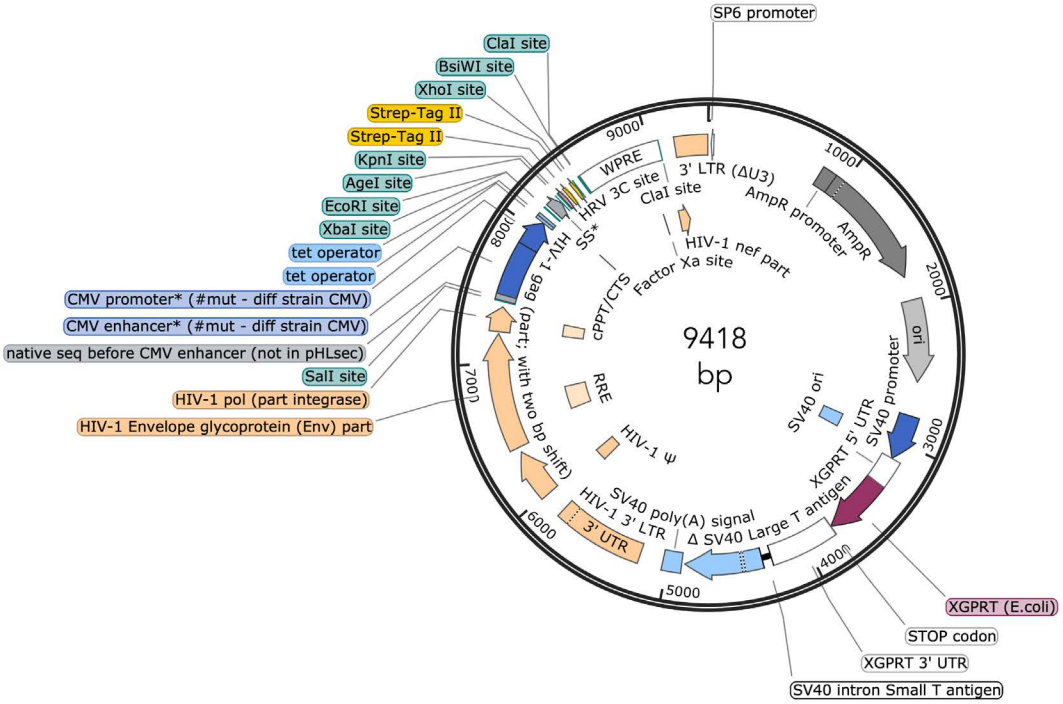
The high copy number pHLSec plasmid vector (Aricescu et al., 2006) with MV tag was gifted by Nicola Abrescia, BIOGUNE, Bilbao, while mRuby2 tagged vectors were donated by other members in Strubi (Figure 2.2).

The pHR-CMV/SFFV-TetO2_3C-Twin-Strep/1D4 vectors for lentiviral expression were developed within Strubi and derived from a second-genertion pHR-SIN-CSGW lentiviral transfer plasmid (Elegheert et al., 2018), see Figure 2.3 and below.

Integrin Constructs

Ectodomains of bovine α_v and β_6 integrins were donated by Alison Burman at Pirbright. The remaining transmembrane and cytoplasmic domains (GenBank: NM_174367.1, GenBank: NM_174698.2) were synthesised by IDT technologies, and amplified by PCR. Full length constructs (amino acid sequences of which are displayed in Figure 2.4) were initially cloned into pHLSec, with a C-terminal

(a)



(b)

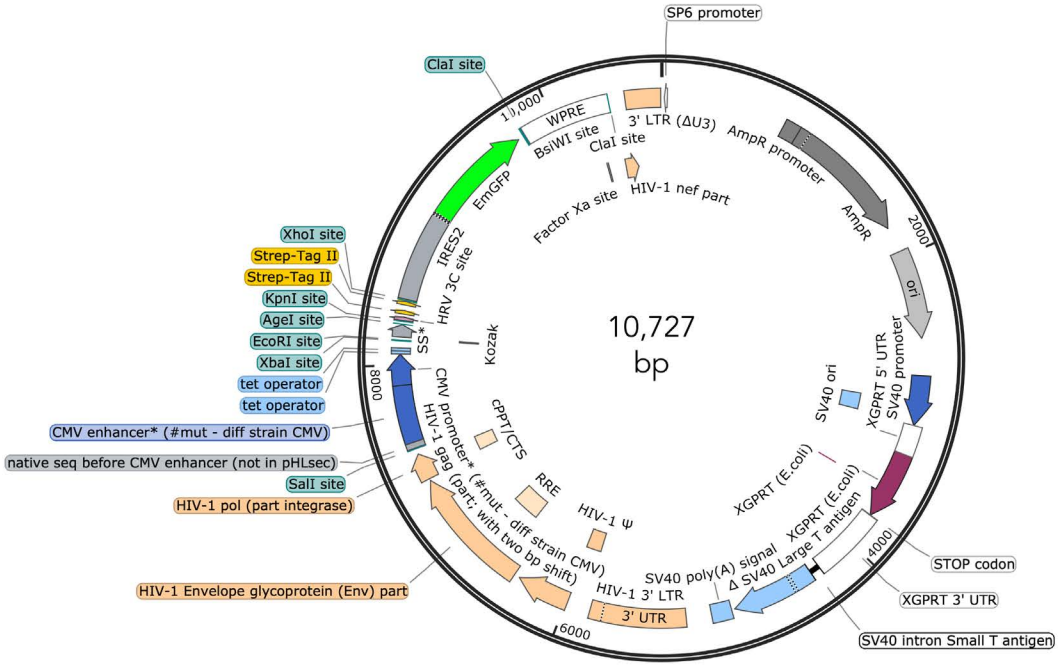


Figure 2.3: The two main vectors used for lentiviral expression. All vectors were kindly donated by members of Strubi. A flurosecent tag was inserted between the KpnI site and StrepII tag shown in (a), as prepared by Dr. Jonathan Elgeheert. (b) Vector map showing the IRES fluorescent tag used for later stable cell lines with β_6, except, in some cases, the CMV promoter was replaced by SFFV in the same region, again by Dr. Jonathan Elgeheert.

8xHis-tag, secretion signal for mammalian expression, and a C-terminal fluorescent tag (either MV, or mRuby2). A ligation-dependent approach was used, with full length α_v inserted between the AgeI and KpnI sites and β_6 was inserted between the EcoRI and KpnI restriction sites (the presence of a AgeI site within β_6's ectodomain necessitated use of the EcoRI site instead).

For head-group only constructs, a similar protocol to that described previously was used (Dong et al., 2014; Kotecha et al., 2017). 'α_v-head' included residues 1-594 of the ectodomain, with a M400C mutation, C-terminal 3C protease site, ACID coiled coil (see Figure 2.5), strepII tag and 6 x His tag. β_3 headpiece comprised residues 1-472 with a Q267C mutation, 3C site, BASE coiled coil (see Figure 2.5) and 6 x His tag, also at the C-terminus. For both constructs, the required region was amplified by PCR, inclusive of the N-terminal AgeI tag already present for α_v and β_3, but, with a C-terminal KpnI site for α_v and a C-terminal NotI site for β_3 due to a KpnI site in the centre of the construct. Additional cysteine mutations were included to generate a disulfide bond to link the $\alpha_v\beta_3$ heterodimer. The ACID/BASE peptide regions replace the cytoplasmic domain, linking α_v / β_3, together to form a heterodimeric α-helical coiled-coil. The β_3 ectodomain fragment (GenBank: NM_000212.2) was kindly donated by Thomas Bowden. This was based on a human-derived sequence, but this has > 90 % sequence identity with the bovine sequence and thus likely not a cause for concern. Ligations were performed using Quick ligase (NEB), long ligase (Takera) or T4 DNA ligase (Promega) and applied according to advised protocols. Typically, vector:DNA were in a 3:1 ratio and, as required, in a 1:1 ratio to maximise ligation success.

Ligations were performed using Quick ligase (NEB), long ligase (Takera) or T4 DNA ligase (Promega) and applied according to advised protocols. For the majority of cases, vector:DNA were incubated in a 3:1 ratio and, separately 1:1 ratio in an attempt to maximise ligation success.

2.1.3 Transient Expression

For transient transfections in adherent mammalian cells (HEK293T, HEK293S, BHK21, COS7, WII7, L and HUH7), 2-4 x 10^5 cells/well were plated one day prior to transfection. For transfection, 1-2 μg of each construct (10 μg DNA was used for 175 mL transfections) was used per well of a 6-well plate (Corning). Briefly, cells were washed with DPBS (Gibco), before addition of 1.5 mL serum free DMEM (BHK21-GMEM in the case of BHK21 cells). Some initial experiments were conducted to determine the required quantity of construct for transfection, since the ratio of plasmid DNA

αv

FNLDVESPAEYSGPEGSYFGFAVDFFVPSASSMFLLVGAPKANTTQPGIVEGGQVLKCDWSSHRR
CQPIEFDATGNRDYAKDDPLEFKSHQWFGASVRSKQDKILACAPLYHWRTEMKQEREPVGTCFLQ
DGTKTVEYAPCRSKNIDADGQGFCQGGFSIDFTKADRVLLGGPGSFYWQGQLISDQVAEIVSKYD
PKVYSIKYNNQLATRTAQAIFDDSYLGYSVAVGDFNGDGIDDFVSGVPRAARTLGMVYIYDGKNM
SSLHNFTGEQMAAYFGFSVAATDINGDDYADVFIGAPLFMDRGSDGKLQEVGQVSVSLQKASRDF
QTIKLNGFEVFARFGSAIAPLGDLDQDGFNDIAIAAPYGGEDKKGIVYIFNGRPTGLNTVPSQIL
EGKWAARSMPPSFGYSMKGATDIDKNGYPDLIVGAFGVDRAVLYRARPVITVNAGLEVYPSILNQ
ENKTCPLPGTDLKVSCFNVRFCLKADGKGALPTKLDFQVELLLDKLKQKGAIRRALFLHNRSPGH
SKNMTISRGGQMQCEELIAYLRDESEFRDKLTPITIFMEYWLDYRTAADATGLQPILNQFTPANV
SRQARILLDCGEGNVCKPKLEVSVDSDQKKIYIGDDNPLTLIVKAQNQGEGAYEAELIVSIPLQA
DFIGVVRNSEALARLSCAFKTENQTRQVVCDLGNPMKAGTQLLAGLRFSVHQQSEMDTSVKFDLQ
IQSSNLFDKVSPVVSYKVDLAVLAAVEIRGVSSPDHIFLPIPNWKYKENPETEEDVGPVVQHIYE
LRNNGPSSFSKAMLHLQWPYKYNNNTLLYILQYDIDGPMNCTSDMEINPLRIKISNSQTSERNDT
VGGQGDRNHLITKRDLTLNEGDVHTLGCGIAECLKIVCQVGRLDRGKSAILYVRSLLWTETFMNK
ENQNHSYSLKSSASFNVIEFPYKNLPIEDIFNSTLVTTNVTWGIQPAPMPVPVWVIILAVLAGLL
LLAVLVFVMYRMGFFKRVRPPQEEQEREQLQPHENGEGNSET

β6

GCAVGGAETCEDCLLIGPQCAWCSQENFTHLSGVGERCDTPANLLAKGCQLTFIENPVSQVEILT
NKPLSIGRQKNSSDIVQISPQSLALKLRPGLEQTLQVQVRQTEDYPVDLYYLMDLSASMDDDLNT
IKELGSLLSKEMSKLTSNFRLGFGSFVEKPISPFMKTTPEEIANPCSSIPYFCLPTFGFKHILPL
TNDAERFNEIVKNQKISANIDTPEGGFDAIMQAAVCKEKIGWRNDSLHLLVFVSDADSHFGMDSK
LAGIVIPNDGLCHLDSKNEYSMSTILEYPTIGQLIDKLVQNNVLLIFAVTQEQVHLYENYAKLIP
GATVGVLQKDSGNILQLIISAYEELRSEVELEVLGDTEGLNLSFTAICNTGIPVPHQKKCSHMKV
GDTASFNVTVSLPNCERRSRHIILKPVGLGDALEILVSPECSCDCQKEVEVNSSKCNNGNGSFQC
GVCACHPGHMGHHCECGEDTLSTESCKEAPGRPSCSGRGDCYCGQCVCHLSPYGNIYGPYCQCDN
FSCVRHKGLLCGDNGDCDCGECVCRSGWTGEYCNCTTSTDPCVSEDGILCSGRGDCVCGKCICTN
PGASGPTCERCPTCGDPCNSKRSCIECYLSADGQAQEECVDKCKLAGATINEEEDFSKDSFVSCS
LQGENECLITFLLTTDNEGKTVIHSINEKDCPKPPNIPMIMLGVSLAILLIGVVLLCIWKLLVSF
HDRKEVAKFEAERSKAKWQTGTNPLYRGSTSTFKNVTYKHKEKQKVDLSTDG

β3

GPNICTTRGVSSCQQCLAVSPMCAWCSDEALPLGSPRCDLKENLLKDNCAPESIEFPVSEARVLE
DRPLSDKGSGDSSQVTQVSPQRIALRLRPDDSKNFSIQVRQVEDYPVDIYYLMDLSYSMKDDLWS
IQNLGTKLATQMRKLTSNLRIGFGAFVDKPVSPYMYISPPEALENPCYDMKTTCLPMFGYKHVLT
LTDQVTRFNEEVKKQSVSRNRDAPEGGFDAIMQATVCDEKIGWRNDASHLLVFTTDAKTHIALDG
RLAGIVQPNDGQCHVGSDNHYSASTTMDYPSLGLMTEKLSQKNINLIFAVTENVVNLYQNYSELI
PGTTVGVLSMDSSNVLQLIVDAYGKIRSKVELEVRDLPEELSLSFNATCLNNEVIPGLKSCMGLK
IGDTVSFSIEAKVRGCPQEKEKSFTIKPVGFKDSLIVQVTFDCDCACQAQAEPNSHRCNNGNGTF
ECGVCRCGPGWLGSQCEC

Figure 2.4: Amino acid sequences for full length α_v and β_6 as labelled. Shading corresponds to: ectodomain, dark green; transmembrane domain, gold; cytoplasmic domain, blue. The head group used for later transient expressions is shaded in bright green for β_3, and, in the case of α_v, boxed in bright green.

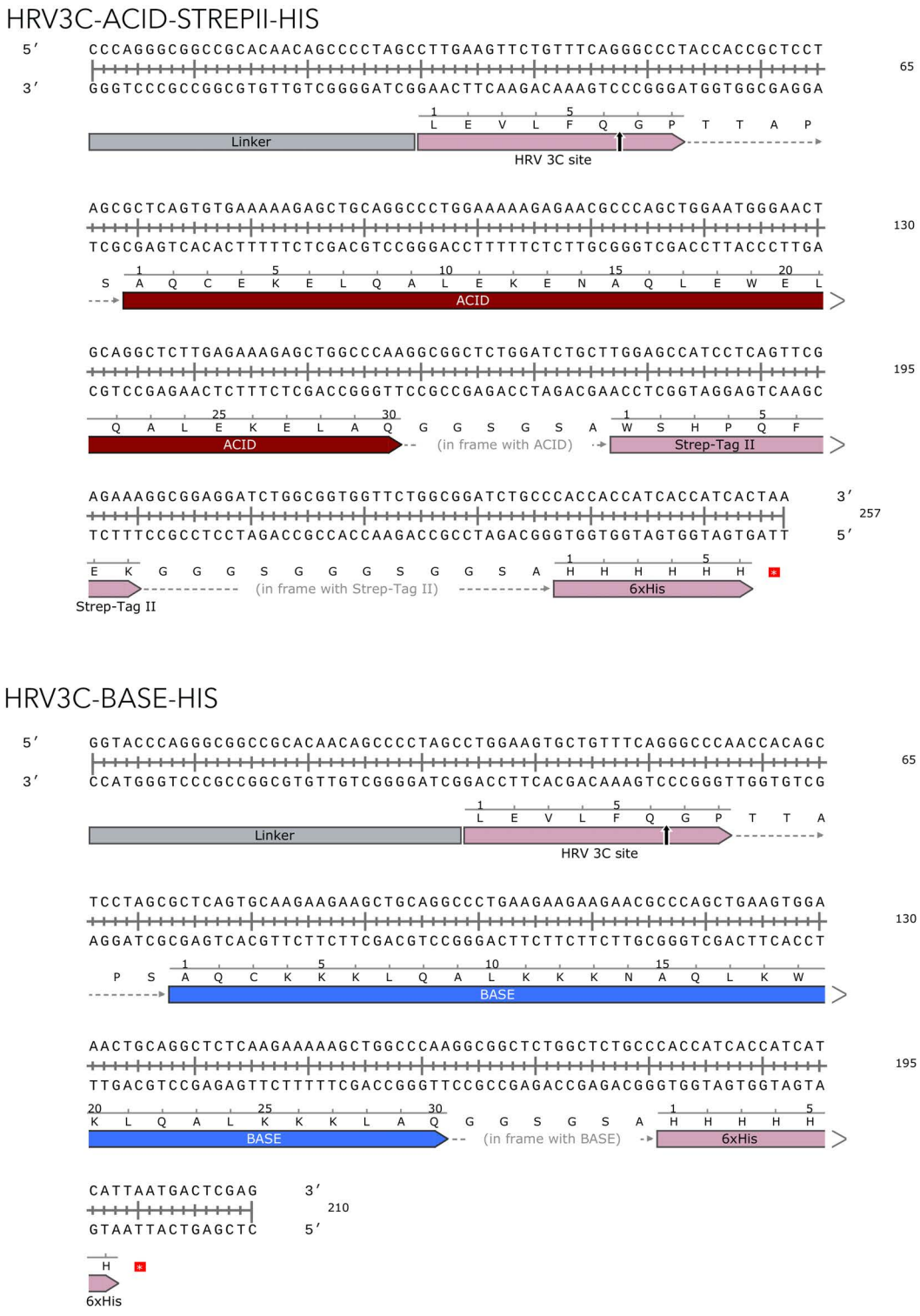


Figure 2.5: DNA and corresponding amino acid sequences of ACID and BASE portions of constructs. Pertinent features, such as purification sites, are shown in coloured boxes. The α_v and β_3 sequences were cloned into the respective N-termini using ligation dependent cloning, and should be identical to those used in Kotecha et al. (2017).

added per reaction can determine successful functional expression of a membrane protein. Too much plasmid can cause expressed protein to remain intracellularly and mis-folded (this could also be rectified with a weaker promoter) (Andréll and Tate, 2013).

A number of transfection reagents were trialed for transient integrin expression. Cationic reagents work by forming electrostatic interactions with the negatively charged plasmid DNA so as to enable binding to the plasma membrane and subsequent internalisation. Excess positive charges can cause cellular toxicity, and thus the ratio of DNA to transfection reagent is quite important. Lipofectamine2000 and PEI were added at 2.5 x the quantity of DNA used, while XtremeGENE (Roche) was used at 3 x the quantity of DNA used.

For calcium phosphate transient transfection trials, prior to transfection, two solutions were prepared in separate sterile tubes: the first, solution A ,contained 3 μg total plasmid DNA and 12.4 μL 2 M calcium phosphate solution, topped up to 100 μL with sterile MQ H_2O. Solution B contained 100 μL of 2 X HBS (HEPES buffered saline). Solution B was then added drop-wise to A while vortexing before incubation at room temperature for 30 min. The solution was then vortexed again for 1 min before drop-wise addition to washed cells. After 6 hr, media was replaced with DMEM, supplemented with 2 % FBS (note that a 12 hr incubation was found to disrupt cells).

For both 175 mL and 6-well transfections, cells were harvested three to four days post-transfection, cell lysate was collected by centrifugation at 4 °C at 10,300 g for 10 min. The resulting pellet was then rinsed with wash buffer (20 mM Tris 250 mM NaCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, pH 8.0) before 5 min centrifugation at 4 °C at 10,300 g. The pellet was resuspended in 1 % DDM in wash buffer with 1 % (v/v) mammalian protease inhibitor cocktail (Sigma-Aldrich, cat. P8340) before being spun for 25 min at 10,300 g at 4 °C. The resulting supernatant was then used for western blot analysis.

Fluorescence Microscopy

Transfection success was partially monitored using an Axiovert fluorescent microscope (Zeiss) with the appropriate filters. Exposure times were varied between 0.5-5 s, and compared to positive and negative transfection controls.

Small Scale Nickel Affinity Purification

Harvested pellet, or, in the case of head domain constructs, supernatant, was harvested and solubilised as described in Section 2.1.3 and mixed with Ni-NTA agarose beads (Qiagen). Samples were washed with 20 and 50 mM imidazole, and eluted with 500 mM imidazole. Once checked by western blot, promising purified fractions were concentrated, and run on a reducing SDS page gel. For trial studies, bands suspected to correspond to β_6 integrin were excised and initially sent for fingerprinting mass spectrometric analysis.

For mass spectrometric analysis, protein bands believed to be associated with the integrin construct were excised from a 8 % polyacrylamide gel and digested with trypsin before analysis by a nano-flow liquid chromatography system, coupled to a tandem mass spectrometer (nLC MS/MS) at the Proteomics Core Facility at CIC bioGUNE. All analyses were performed by the facility.

2.1.4 Stable Cell line generation

Lentiviral expression offers a route to stable cell line generation that is more rapid than traditional selection approaches, which typically take 8-10 weeks per construct (Chaudhary et al., 2012). By contrast, lentiviral systems enable the establishment of a stable cell line expressing a given construct within a week. Lentiviral methodology for stable cell line generation has been an established method for some years, with some commercially available kits in existence. However, such kits are exceptionally expensive and little systematic research into optimisation has been published until recently and the kits are exceptionally expensive (Elegheert et al., 2018).

Constructs of full length α_v and β_6 were first incorporated into lentivirus genomes by transient transfection of a lentivirus producer cell line (HEK293 Lenti-X, Takara/Clontech, cat. no. 632180).

To this end, a 100 μL solution of DMEM/F12 containing 5 μg each of two transfer plasmids ('pHR-[CMV or SFFV]TetO$_2$', see Elegheert et al. (2018), and Figure 2.3) encoding α_v or β_6 alongside 10 μg packaging (psPAX2; AddGene plasmid no. 12260) and 10 μg envelope (pMD2.G; Addgene plasmid no.12259) plasmids were prepared, and mixed with a 100 μL solution of PEI (1:2.5 w/w DNA:PEI ratio). This solution was then vortexed gently and centrifuged briefly before being left to incubate for 30 min. A 90 % confluent T75 flask containing HEK293T Lenti-X cells was then washed with 10 mL dPBS and the media replaced with 12 mL DMEM/F-12/2 % FBS. The DNA/PEI

mix was then added to this flask, and gently mixed by tilting. Cells were then incubated at 37 °C with 5 % CO_2.

After three days, the resulting supernatant containing the lentivirus-encoding α_v and β_6 transgenes was harvested, filtered through a 0.45 μm filter unit, diluted with SFM and mixed with 18 μL (from a 10 mg/mL 1000x stock) polybrene (polybrene is a cationic polymer that is believed to neutralise charge repulsion between virus capsids and cell surfaces) and used to transduce HEK293S cells (90 % confluent). For inducible cell lines, HEK293S GnTI$^-$ TetR cells were used. Since lentiviruses exploit host cellular machinery for the amplification of their own genetic material, constructs are expressed in infected cells. After three days, cells were split into a T175 flask for expansion.

Both HEK293 cell lines were cultured in DMEM/Nutrient Mixture F-12 (DMEM/F-12; with added L-Glutamine; Thermo Fisher) with 10 % v/v FBS (Gibco).

For screening of inducible cell lines for suitable additives, induction was performed using a 10 mg/mL stock solution of doxycycline hydrochloride (Dox; Thermo Fisher), dissolved in ethanol and filtered. Induction tests were performed using sodium butyrate (Merck) as well as valproic acid (VPA; Merck).

Enrichment of Transduced Cell Populations by Fluorescent Activated Cell Sorting (FACS)

FACS was conducted using a Sony SH800Z Fluorescence activated cell sorter.

Following two passages of stable cell lines post transduction, adherent cells were first washed, and then suspended using Trypsin-EDTA (0.05 % (wt/vol)) without phenol red (Gibco, the phenol red can interfere with fluorescence signal). Trypsin was inactivated using soybean trypsin inhibitor (Thermo Fisher, 1 mg/mL stock), which has been found to reduce clumping (Elegheert et al., 2018). Trypsinised cells were then filtered through a 40 μm cell strainer (Corning) to avoid blockage of the cell sorter nozzle.

Prior to sorting of stable cells, a compensation matrix, to take into account overlap of mRuby2 and MV emission spectra, was determined using 'positive control' cells expressing mRuby2 and MV fluorochromes only, and non-transduced cells as a negative control.

Stable cells showing both mRuby2 and MV fluorescence were then sorted, and ca. 100,000 cells collected in 2 mL DMEM with 10 % FBS (Gibco). The collected cells were then treated with

penicillin streptomycin (Thermo Fisher) and aliquoted into a 6-well plate (Corning).

Freezing Cells

For long term storage, cells in 175 mL flasks were treated with trypsin and suspended in a total of 10 mL serum free media before being spun down at 1000 g for 3 min. The supernatant was removed and the resulting pellet resuspended in 6 mL media to a density of ca 1 x 10^7 cells/mL with 10 % DMSO and frozen slowly within an insulated box placed at -80 °C.

Cells were frozen with DMSO and kept at -80 °C.

SDS Page and Western Blot

To assess sample purity, samples were initially denatured by incubation in reducing buffer for 2 min at 95 °C. In most cases, pre-cast 4-12 % SDS Page gels were run in MOPs buffer for 70 min at 165 V and stained using Instant Blue stain (Sigma), unless otherwise stated (for experiments conducted in Bilbao and Pirbright, 'home-made' 6 % or 8 % gels polyacrylamide gels were used instead with tris-glycine buffer). Typically, either unstained or pre-stained BenchMark^{TM} Protein Ladders (Thermo Fisher) were run alongside analytes to determine molecular weights.

For western blots, gels were transferred using the semi-wet transfer method onto a nitrocellulose membrane for 70 min at 10 V in NuPAGE Transfer Buffer (Thermo Fisher) with ethanol. The membrane was then blocked with 5 % milk (or 3 % BSA in the case of HRP conjugated primary antibodies, see below) for either ca. 1 hr at room temperature or overnight at 4 °C.

Blots were then incubated with a primary antibody at dilutions stated below for 2 hr at room temperatures, or overnight at 4 °C.

Primary antibodies used, and, in parenthesis, associated dilutions (v/v), were as follows:

- Penta-His mouse monoclonal antibody (Qiagen, 1 in 1000).
- Rho-1D4 purified monoclonal antibody (lot #1019, 5.8 mg/mL) for 1D4 tagged constructs (1 in 10000).
- Precision protein Strep-Tactin-HRP conjugate (Biorad) for StrepII tagged constructs (1 in 10000).

A secondary anti-mouse antibody (Thermo Fisher) was then applied for 30 min-1 hr before 3-4 rounds of 10 min washes with PBS 1 % Tween buffer.

2.1.5 Expression of Truncated $\alpha_v\beta 3$ Integrin as a FMDV receptor

Transient transfection

For large-scale transient expression of $\alpha_v\beta_3$ constructs, 12 flasks of HEK293S cells (800 mL per flask) were suspension cultured suspension with Freestyle media, supplemented with 1 % FBS, 1 % non-essential amino acids (NEAA) and 1 % L-Gln to a density of 1.9 million cells/mL at a speed of 130 rpm, 37 °C.

For each 800 mL cell suspension, the following protocol was used. 0.4 mg (0.2 mg α_v, 0.2 mg β_3) DNA was sterilised using 200 μL chloroform before being diluted with 10 mL serum free Freestyle media with supplements as above. The transfection reagent PEI was incubated in a separate 10 mL aliquot of Freestyle media, before mixing with the DNA mixture. VPA diluted in 60 mL Freestyle media was prepared separately. Each 800 mL aliquot of cells was spun down at 1000 g for 7 min with slow deceleration to improve pelleting. The supernatant was removed and resuspended in 20 mL Freestyle and PEI/DNA mix, such that the final concentrations of PEI and VPA were 1.5 μg, and VPA 0.75 mM respectively and returned to 37 °C with a higher speed of 150 rpm. This same procedure was then repeated for the 10 remaining 800 mL flasks, which were then returned to the shaking incubator. The shaking speed was reduced after 5 hr, and additional media (recycled from pelleting) was added, supplemented with 0.5 % v/v Pen/Strep. Flasks were then incubated in a shaking incubator at 130 rpm at 37 °C for three days.

Harvesting

Cells were spun at 4500 g at 4 °C, and the resulting supernatant filtered with a 0.22 μm sterifilter (Merck Millipore), before diafiltration through a micro-molecule permeable filter to remove contaminating salts and small proteins. Filtrate was then stored at 4 °C overnight.

The supernatant was then loaded onto a 5 mL pre-packed Histrap column with pre-charged Ni Sepharose (GE Healthcare) in 20 mM Tris pH 8.0, 1 mM $CaCl_2$, 1 mM $MgCl_2$ and 150 mM NaCl ('Buffer W'). After thorough washing with a gradient up to 50 mM, the protein was eluted in 20 mM Tris pH 8.0, 1 mM $CaCl_2$, 1 mM $MgCl_2$,150 mM NaCl, 250 mM imidazole.

Five 5 mL fractions corresponding to a clear peak were then pooled and concentrated with a 30 kDa ultrafiltration device (Amicon) equilibrated with Buffer W to a volume of 500 mL. Finally, samples were loaded onto a Superdex 200 Increase column (GE Healthcare) which had been previously calibrated with the same column and loops with a gel filtration standard (Biorad, 1511901).

2.2 Virus and VLP Production and Purification

2.2.1 Equine Rhinitis Virus A

HeLa Ohio or vero cells were grown on Dulbecco Eagle medium (DMEM; Gibco) supplemented with 5 % fetal bovine serum (Gibco), 2 mM L-glutamine, and penicillin-streptomycin at 37 °C in a 7 % CO_2 atmosphere. Assistance with ERAV growth and purification was kindly provided by James Kelley (Pirbright Institute). At circa 70-80 % confluency, cells were washed with 1 x PBS and media replaced with DMEM supplemented with 2 mM L-glutamine and then infected with ERAV at low MOI and maintained at 37 °C in a 7 % CO_2 atmosphere. After three days, supernatant was harvested and clarified by centrifugation at 2500 g for 10 min.

Supernatant was precipitated with PEG6000 and stored overnight at 4 °C before centrifugation at 4 °C.

Virus was then purified by sedimentation through a 15-45 % sucrose gradient (w/v PBS, pH 7.4) at 111,000 g and 4 °C for 2.5 hr. Since no, or typically at best a very faint band could be seen, gradients were fractionated and each fraction assessed by A260 and A280 absorbance readings and SDS-page gel. Peak fractions were then pooled and samples concentrated (Amicon) and desalted (Zeba).

2.2.2 EV-F3

Buffer-E refers to 50mM HEPES, 200mM NaCl, 0.5 % (v/v) NP40, pH 8.0 in the proceeding section.

BHK21 cells were grown in BHK21-GMEM (Gibco) supplemented with 10 % fetal bovine serum (Gibco) and 1 mM MEM NEAA at 37 °C in roller bottles. When cells reached 70 % confluency, they were washed and the media replaced with BHK21-GMEM supplemented with 2 % fetal horse serum and 1 mM MEM NEAA. EV-F3 from frozen stocks was then added, and left to grow for three days before harvesting.

0.5 % (v/v) NP40 was added to harvested supernatant and stored at -80 °C. Samples were then thawed and virus precipitated with 8 % (w/v) PEG 6000 and stored at 4 °C overnight. The resulting suspension was spun at 3700 g for 1 hr at 4 °C and the pellet resuspended in circa 35 mL buffer-E and spun again at 3700 g. Virus was then pelleted through a 30 % sucrose cushion at 29,000 rpm for 3 hr at 4 °C. The pellet was resuspended in 0.5 mL buffer-E and stored at 4 °C overnight. The resulting suspension was spun at at 3700 g for 1 hr at 4 °C and the resulting pellet resuspended in ca 35 mL buffer-E and spun again at at 3700 g.

Virus was then purified with a 30 % sucrose cushion at 29,000 rpm for 3 hr at 4 °C. The resulting pellet was resuspended in 0.5 mL buffer-E and stored at 4 °C overnight, before spinning at 14,000 rpm for 0.5 hr. The resulting supernatant was kept at 4 °C, and the pellet resuspended again in 1 mL buffer-E and spun as before. The resulting combined supernatants were applied to a 15-45 % sucrose gradient and spun at 29,000 rpm for 3 hr at 4 °C.

Typically, a band could be seen in the centre of the gradient. In cases where this band was less well defined, central fractions were collected, desalted, pelleted and gradient purified.

Room Temperature Crystallography Samples *

For virus samples crystallised for room temperature crystallography experiments at I24, Diamond Light Source, UK (Chapter 3, Section 3.2.2), the same protocol was used as above, except, instead of PEG 6000 precipitation was achieved with ammonium sulphate.

Production of natural empty particles was amplified by inhibition of RNA replication using GuHCl, and the resulting particles purified using the above protocol, but without detergent.

Plaque Assays

BHK21 or BSRT7 cell monolayers in 24 well plates were washed with 1x DPBS (Gibco) before incubation at 37 °C with serial dilutions of EV-F3 for 1 hr (100 μL per well).

For receptor experiments, BHK21 cells were treated with either sialidase from *Clostridium perfringens* (broad specificity, Sigma) in serum free GMEM-BHK21, or just serum free GMEM and incubated at 37 °C for 1 hr. The cells were then washed twice with DPBS, before a dilution series of EV-F3 in 2 % Fetal Horse Serum (FHS, Gibco) GMEM-BHK21 was added. Each dilution was repeated in triplicate. After 1 hr incubation with virus at 37 °C, the supernatant was removed, cells were washed, and a layer of 3 % carboxymethyl cellulose (CMC) in 10 % FBS GMEM was carefully pipetted into each well. Cells were then incubated at 37 °C for three days, before the CMC was removed, the cells washed twice, and stained with crystal violet solution.

2.2.3 Yeast PV VLP Production*

Samples were prepared and purified at the School of Molecular and Cellular Biology, Leeds by Lee Sherry. Multicistronic vectors were generated using three constructs; comprising (i) PV P1 and PV 3CD, (ii) PV P1-RhPV IRES and PV 3CD, (iii) PV P1, TaV-2A and PV 3CD, since this was found to cause lesser metabolic stress on *Pichia Pastoris* systems than co-transfections. Since, for multicistronic vectors, everything is expressed within one cassette, cells positive for an encoded detection marker (fluorescence or resistance) should be expressing all three encoded constructs. This has proved exceptionally useful for some VLP expression systems. Constructs were cloned into the pPINK vector, both low copy number and high copy number plasmid types.

Yeast was grown for 24-48 hr at 220 rpm, 27 °C then transferred into 100 mL Buffered Glycerol-complex Medium (BMGY) and incubated at 220 rpm, 27 °C until the OD600 exceeded 2. The harvested cells were then spun and resuspended in 100 mL of Buffered Methanol-complex Medium (BMMY) with steady addition of methanol until 1 % concentration was reached after 24 hr. Samples were then spun down at 1500 rpm and yeast cells were broken open using a OneShot cell disrupter (Constant Systems Limited) at 40 kpsi before centrifugateion at 10,000 g. VLPs were then precipitated with 40 % saturated ammonium sulphate overnight before pelleting in 30 % sucrose at 150,000 g for 3.5 hr at 4 °C. The resulting pellet was then resuspended and further purified on

a 15-45 % sucrose gradient and fractions were harvested and checked by WB against an α VP1 monoclonal antibody (Millipore mAb 8650).

2.2.4 Vaccinia FMDV VLP Production*

VLP samples were prepared at Pirbright Institute, UK by Alison Burman, using the protocol described in Porta et al. (2013a). Mutant VLPs were also produced by this method, notably those stabilised by a 93C mutation (Porta et al., 2013a) on the two-fold axis aimed at disulphide bonding the pentamer subunits-this will be referred to by the cc suffix.

Briefly, constructs were cloned into vaccinia virus transfer vector pBG200, in a position downstream of the T7 promoter. Recombinant viruses were prepared through transfection of pBG200-VLP into CV-1 cells that were infected with vaccinia virus (VV) strain WR. The recombinant VVs were then selected in HuTK-143 cells with 5-bromo-2-deoxyruridine.

Stable recombinant VVs were obtained via three rounds of plaque purification and screening by PCR, which were then amplified in RK13 cells, cultured in DMEM with 10 % FBS at 37 °C.

HEK293 cells were infected with both vaccinia-VLP (MOI 10) and vTF7.3 (MOI 5) and incubated for 20 hr at 37 °C. Cells were harvested by centrifugation at 3500 g for 30 min at 4 °C, the pellet resuspended in 40 mM sodium phosphate, 100 mM NaCl pH 7.6 and lysed with 1 mL 0.5 % Igepal detergent (octylphenoxypolyethoxyethanol, Sigma). Following incubation on ice for 20 min, the sample was clarified for 20 min at 10,000 g, 4 °C, the resulting pellets resuspended in small volumes of buffer and re-extracted with 1 volume of chloroform from which the aqueous phases were combined with clarified supernatant and layered onto a 30 % sucrose cushion that was centrifgued at 10,5000 g for 5 hr at 12 °C. Pellets were then resuspended in buffer to which 0.1 % Igepal and 200 mg/mL RNAse A was added for 30 min on ice. Treated sample was then pipetted onto a 15-45 % continuous sucrose gradient and spun for 22 hr at 54,000 g at 12 °C. Each gradient was then fractionated into 12 x 1 mL fractions which were analysed by SDS-PAGE. Sucrose was removed as before through desalting with a spin column (Zeba, Pierce) and concentrated by ultrafiltration (Amicon).

2.2.5 Virus Preparation for EM

For all EM sample preparations, it was necessary to remove the high concentration (typically ca. 30 %) of sucrose from sucrose-gradient purified samples. For all virus samples presented here, this was performed using a Zeba^{TM} desalting column (Thermo Fisher), followed by concentration using a 100 kDa cutoff Amicon concentration unit. This latter step was, as discussed in Chapter 6, is potentially disruptive of virus particles due to strong affinity to the concentrator membrane.

2.3 Picornavirus Concentration

Typically, A260 absorbance readings were used to determine the final concentration of purified virus preparations, but this is problematic for virus preparations, where absorbing encapsulated genome interferes with protein signal (Wilson and Walker, 2010).

While it is still necessary to provide calculated concentration, it is helpful to also explicitly report both A260 and A280 readings. For the purposes of this thesis, all complete virion protein concentrations were determined using Warburg and Christian's formulae, which takes into account contaminating nucleic acids, but A260 and A280 readings will also be stated where appropriate (see Equation 2.1) (Warburg and Christian, 1941; Layne, 1957).

$$[Protein, mg/mL)] = (1.55 \times A_{280}) - (0.76 \times A_{260}) \tag{2.1}$$

A280 and A260 readings were determined using a NanoDrop (Thermo Fisher), with a baseline reading taken using the buffer that the sample was suspended in.

2.4 Protein Crystallisation and Structure Determination

2.4.1 Crystallisation

EV-F3 and ERAV samples were concentrated (Amicon, 100 kDa cutoff) and desalted (Zeba) prior to crystallisation. FMDV A22cc VLP samples were delivered sufficiently concentrated for crystallisation.

Crystallisation experiments were set up using a 96-well vapour-diffusion sitting drop SBS plate

Table 2.1: Summary of crystallisation conditions used for experiments.

Sample	P:R Ratio	Precipitant	[Precipitant] / M	pH	A260	A280	[Virus]/ mg/mL
EV-F3	1:1	$(NH_4)_2SO_4$	1.50-1.34	8.0	9.1	5.6	1.74
A22cc	1:1	$NH_4CH_3CO_2$	3.23-3.45	7.14-8.00	4.8	5.3	5.49

format. In all cases, pre-defined conditions were screened across a narrow concentration and pH range (see Table 2.1) with 100 nL reservoir solution (R) and 100 nL protein solution (P) in CrystalQuick Greiner plates using a Cartesian robot (Genomic Solutions) (Walter et al., 2003, 2005).

For room temperature experiments, EV-F3 crystallisation was set up using sitting-drop nanolitre vapour diffusion in Greiner Crystal Quick X plates. Initial crystallisation trials used a SaltRx (Hampton Research) screen using a Cartesian robot at temperature of 20 °C. Crystals, most of bipyramidal morphology, grew in 77 out of 96 conditions. Those used for cryo-cooled experiments were grown in 1.5 M Ammonium Sulphate and 0.1 M Tris at pH 8.5 and tended to grow to a maximum size of 0.1 x 0.1 x 0.22 mm^3). Specific experimental set ups are described in Chapters 3 and 4 (Sections 3.2.2, 4.1.2, 4.1.3, 4.1.4).

Crystals of A22cc empty capsids (see Table 2.1) with maximal diameters of 50-100 μm^3 grew within 1 week at 21 °C with 3-4 M ammonium acetate at neutral pH. Experiments were performed at I23, Diamond light source. A focussed beam 70 x 70 μm^2, adjusted for crystal sizes, was used (λ = 3.17, 3.07 Å).

2.4.2 Crystal Harvesting and Data Collection

Room Temperature Experiments

Diffraction data for EV-F3 mature and empty structures were collected at room temperature (21 °C) from crystals in crystallisation plates (*in situ* data collection) at I24, Diamond Light Source (Didcot, UK). Diffraction images of 0.05 ° or 0.1° oscillation were recorded on a Pilatus6M detector (frame rate of 100 Hz) using beam sizes between 0.02 x 0.02 mm^2 and 0.05 x 0.05 mm^2 depending on the size of the crystals. The X-ray beam was homogenised with a 0.25 mm thick carbon plate and focused downstream from the crystal. To obtain maximum data from a single exposure position, radiation damage was minimised by using maximum beam flux and a short exposure time per image. With 0.1 s exposure time and 100 % beam transmission, typically 6 to 10 useful images

could be collected from a single exposure position of a crystal. For large crystals, data could be collected from up to six positions.

Cryo Experiments

Virus Crystals

The large unit cells and limited crystal contacts inherent in virus crystals renders them immensely fragile; so much so that cryo-cooling often breaks the delicate lattice. Cryo-cooling tends to increase the apparent mosaic spread of the crystal, a particular problem in the case of viruses, where diffraction lunes are closely spaced (large unit cell) (Fry et al., 1999a). Thus, it has been historically necessary to conduct experiments at room temperature. However, this limits the possible radiation dose used in data collection such that as little as 0.5 ° of data can be collected prior to radiation-induced crystal destruction. For the longer wavelengths necessary for anomalous phasing at the calcium edge, radiation damage is an order of magnitude greater and thus cryo-cooling is essential.

For GSH experiments (Chapter 4, Section 4.1.3), soaks were performed overnight with 10 mM of either GSH, CG or GSH/CG (the final concentration for this latter condition was 20 mM) prior to cryoprotection where each additive was maintained. Similarly, 6'-Sialyllactose (Sigma-Aldrich), or, for competition experiments, 3' and 6'-Sialyllactose (Sigma-Aldrich), were dissolved in the crystallisation solution to a concentration of 10 mg/mL, and ca 100 nL was added to crystallisation drops and incubated overnight before harvesting with cryoprotection.

Virus crystals are typically sensitive to cryoprotectants, and have been observed to crack when subjected to glycerol concentrations of above 15 % (v/v). Although this was found to be the case for EV-F3 crystals. However, the more gentle, so-called 'gradient-glycerol' approach to cryoprotection worked well. Here increasing concentrations of glycerol were added to the crystal drop until a concentration of approximately 20, or 25 % was achieved, as appropriate.

Crystals were mounted using loops in the range of 50-100 μm, depending on crystal size. Often loops smaller than the target crystal were used for ease of fishing. For I23 experiments, bespoke loops were used that are compatible with CombiPucks (MiTeGen), with diameters of 50, 75 or 100 μm in diameter, and kindly provided by Ramona Duman, (beamline: I23, Diamond Light Source UK).

For all EV-F3 soak experiments, data were collected at I03 (Diamond Light Source, UK) using automated X-ray centering and data collection at 100 K with an X-ray wavelength of 0.976 Å. Diffraction images were recorded with 0.03 ° or 0.05 ° oscillations with a beam size of 80(H) x 20(V) μm^2. Images were collected with 0.075 or 0.01 s exposure times with 100 % beam transmission, and, typically,180 ° of data could be collected.

2.4.3 Detectors

The majority of synchrotron-derived X-ray crystallography datasets described in this thesis were collected using Pilatus or Eiger pixel array detectors, implemented at I03. These comprise an 8 x 2 array of CMOS ASICS (complementary metal-oxide-semiconductor application-specific integrated circuit) that are arranged to form a larger detector.

Pixel array detectors hold many advantages compared to their CCD (charged coupled device) precursors, including the ability to directly detect X-rays, count single photons, and a good signal-to-noise ratio with high dynamic range, short readout time and high frame rates.

The I23 beamline has a more complex Pilatus 12M-DLS detector to enable the collection of data over an extensive diffraction angle range (2 $\theta \pm$ 100 °). Semi-cylindrical in shape, this was custom built for the beamline. Also unique to I23, both the detector and sample are kept *in vacuo* so as to eliminate X-ray air absorption.

2.4.4 Data Processing

Icosahedral structures contain 5-fold axes that cannot tesselate in three-dimensional space. For the *F*23 space group, for a particle at the crystallographic origin, this results in minimal 5-fold non-crystallographic symmetry (NCS). For a similarly positioned particle in space group *I*222, 15-fold NCS is present.

For EV-F3 room temperature experiments, images were indexed, integrated and scaled using HKL2000 (Minor and Otwinowski, 1997) and structures determined using EV-E1 (Smyth et al., 1995) (PDB code 1BEV) as the starting model by positioning and orientating it in the unit cell of EV-F3. Rigid-body refinement was then followed by cyclic positional, simulated annealing and B-factor refinement using strict NCS constraints with CNS (Brunger, 2007). NCS averaging was

performed using GAP (D Stuart, J Grimes & J Diprose, unpublished) and models were rebuilt with Coot (Emsley and Cowtan, 2004; Emsley et al., 2010).

EV-F3 sialyllactose and glutathione datasets were processed using the Xia2-3dii or Xia2-Dials pipelines available on ISPyB (Winter, 2010; Waterman et al., 2013; Delageniere et al., 2011), before scaling and merging using Aimless (Winn et al., 2011).

Note that the resolution cut off chosen for each dataset changed over the course of my PhD as my understanding of crystallography improved. Data collected at I23 were truncated based on the CC0.5 value, typically set to 20-30 %, while, for the majority of remaining datasets presented, cut offs were decided based on the appearance of the resulting Wilson plots, specifically whether they followed a smooth curved profile. The general principle is that high NCS allows useful information to be obtained from extremely weak data, since it implies oversampling of the molecular transform, however beyond a certain point the integration software fails (for instance there will be insufficient information to accurately define spot positions at higher angles), and this can be detected in the Wilson plot.

Data collected at I23 beamline (Diamond Light Source, UK) for EV-F3 and FMDV A22cc were indexed and integrated using XDS, merged and scaled with XSCALE, and then refined with either 5-fold or 15-fold NCS (for *F*23 and *I*222 crystals, respectively) (Kabsch, 2010*b*,*a*). FMDV A22cc data were refined using Phenix with strict NCS constraints, while both EV-F3 datasets were refined using CNS with cyclical positional and B-factor refinement with strict 5-fold NCS constraints (Adams et al., 2010, 2011; Brunger, 2007).

2.4.5 Anomalous X-ray Diffraction for Metal Ion determination

When an incident photon with sufficient energy hits an atom, some photons are absorbed and re-emitted (incoherently) with lower energy, and we see fluorescence, others are absorbed and then re-emitted at the same energy (elastic scattering) but with the phase retarded, while others are scattered 'normally'. For those with a retarded phase, there is an imaginary component to the scattering factor, f″. This effect as a function of X-ray energy can be seen by a dramatic increase in fluorescence, and separately, in a phase change of the elastic scattering due to f″.

The use of anomalous signal to localise and identify metal ions in three-dimensional structures has been reported previously, such as in the very recent publications by Handing et al. (2018);

Rozov et al. (2019) and is perhaps the only robust method to localise metal ions in macromolecular structures. For instance, even at atomic resolution, analysis of ion coordination and the surrounding solvent cannot provide an unambiguous answer (Auffinger et al., 2003), especially since interpretation informed by experimental preparatory steps can be muddied by the presence of multiple different ions in purification buffers. Furthermore, electron densities are still a snapshot in time and so interpretation can be further clouded if ion sites are only partially occupied. X-ray spectroscopic techniques have been applied with success to determine the coordination geometry and oxidation states of atoms to high resolution, but are unsuitable for the more complex environments of macromolecules. Alternative methods include solid state NMR and cryo scanning TEM (Wu and Zhu, 2012; Elad et al., 2017).

It is possible to exploit the different underlying electron structure of individual elements by using the resulting inherent set of absorption edges in the X-ray range. Each electron shell, termed K, L, M has a different associated binding energy. Measuring above and below a given absorption edge, we see a dramatic change in anomalous signal, such that peaks from anomalous difference Fouriers present in the high-energy site that are absent on the low-energy side can indicate the atomic position of an anomalous scatterer. Absorption edges for lighter metals reside at longer wavelengths than are conventionally used at synchrotron beamlines. For instance, the potassium edge, of interest in this thesis, has a K-edge around 3.608 keV. For data collection at such energies/wavelengths, air absorption in the beam path is high and very large diffraction angles are needed for data collection. The soft wavelength beamline I23 at Diamond Light Source, UK is designed for for such experiments, with an experimental chamber in vacuum and a detector covering more than 180 ° solid angle. I23 is currently the only synchrotron beamline capable of data collection at such energies/wavelengths.

For the purposes of our experiments, crystals were exposed to wavelengths above and below the calcium and potassium K-edges.

f′ and f″ were determined using the Cromer and Liberman approximation. This does not account for the effects of neighbouring atoms, which can be substantial near an absorption edge. ANODE may be used to provide an estimate of anomalous density using a model and merged MTZ file processed by SHELXC as input (Sheldrick, 2008; Thorn and Sheldrick, 2011). This is achieved by reversal of experimental phasing (i.e. SAD, MAD, SIR) procedures, whereby na-

tive phases, obtained from the input PDB file, are retarded so that they are appropriate for the anomalous scattering component. When combined with the measured anomalous differences and Fourier transformed, the resultant map should show the anomalous sub-structure. SHELXC, part of the SHELX pipeline for structure solution, provides the appropriate amplitude and phase terms (Sheldrick, 2008).

2.5 Transmission Electron Microscopy

Cryo-TEM has arguably become the second titan of macromolecular structure determination, with recent advances tipping some reconstructions over the atomic resolution edge. Converse to X-ray crystallography, TEM overcomes artefacts observed in harsh crystallisation conditions and allows samples to be kept in an environment closer to physiological conditions than crystallisation buffers. The upper size limit is also relatively higher for TEM experiments, while a much lower quantity of protein sample is required.

For cryo-electron microscopy, the entire system must be keep at a high vacuum to minimise sample contamination and to avoid electron scattering by gas molecules due to the strong interaction of electrons with matter.

At the top of the microscope column, an electron source generates coherent electrons used to probe the inserted sample. For the high-end cryo-TEMs, used in this work (Glacios and Titan Krios, Thermo Fisher), a Field Emission Gun (FEG) is used, where electrons are emitted by application of a strong electric field. FEGs provide a brighter and more coherent source than thermionic sources, such as the lanthanum hexaboride (LaB6) crystals in the T12 (Tecnai) used for negative stain experiments, which is nonetheless brighter and more rigid than the alternative tungsten sources.

Probe electrons are accelerated in high vacuum at voltages of, usually, 120-300 kV through an ultra-thin specimen. An electromagnetic lens system then focusses resulting scattered electrons with condenser, objective and projection lenses. Image quality is degraded by lens imperfections, which causes aberrations. Spherical aberration can be accounted for during image processing and should be microscope specific. In the case of data processed for this thesis, the spherical aberration was assumed to be 2.7 mm.

2.5.1 Negative Stain Microscopy

Negative staining involves the application of a heavy metal salt solution to a sample suspension on an EM grid, which is then blotted to a thin film and air-dried. As the name suggests, macromolecules are unstained, and thus should appear white above a grey background.

All negative stain investigations were conducted using a T12 microscope, operating at 120 kV equipped with a CCD camera (Eagle, FEI).

For negative stain experiments presented in this thesis, Formvar grids (Agar Scientific) were glow discharged for 30 s on high, before application of 4 μL of sample on the carbon surface. Formvar refers to a collection of thermoplastic resins which can stretch, whilst carbon is brittle, thus Formvar-Carbon grids are much more robust, which is essential for the manipulations involved, but at the cost of some sample drift (Watson, 1956). Sample was then incubated for 30-60 s, before blotting from the bottom edge contacting the tweezers holding the grid in place. This was then washed in MQ water and blotted, a process repeated three times to remove buffer. Where phosphate buffers were used, such as for the VLP samples, an additional step was included to minimise crystallisation of uranyl salts. Grids were washed with uranyl formate stain and incubated in circa 20 μL stain for 30 s before blotting as before.

2 % uranyl formate (UF) was typically used in preference to 2 % uranyl acetate (UA), despite the higher contrast of UA due to its larger 8-10 Å grain size, as UF stains were found to be more reliable in terms of stained grid quality. Both stains are acidic, and could cause sample disruption, but this was not a problem for the majority of preparations. Even acid-labile wild-type FMDV samples have been prepared without too much disruption using UF and UA stains (personal communication, Abhay Kotecha).

2.5.2 Single Particle Cryo-Electron Microscopy

Thin-vitrified suspensions of samples enable the visualisation of particles free from dehydration, or fixation or staining damage (Adrian et al., 1984).

Sample Preparation

Hydrated biological macromolecular samples must be immobilised on cryo-EM grids and vitrified to preserve a near-native state prior to imaging. In an ideal scenario, samples are free of contamination (such as ethane and ice crystals), sufficiently thick to not disrupt immobilised sample, but thin enough to be within the mean free path length ($<$ 350 nm for 300 keV electrons) (Grimm et al., 1996)). Sample size can also have a substantial impact on suitable thicknesses; for giant mimivirus, for instance, ice thicknesses of 700-800 nm were required, and beautiful structures were still obtained (Xiao et al., 2005b, 2009).

Grids and Grid Quality

Cryo-electron grids used in this work comprised a copper mesh covered by a carbon support. Exposure of carbon supports to an electron beam can cause deformation resulting in beam-induced specimen motion. This is largely corrected for using motion-correction algorithms, in our case MotionCor2 (see below) (Zheng et al., 2017). Sample motion can still be a problem, especially in initial movie frames, and alternative supports, for instance gold, may reduce beam-induced movements, but this was not found to be necessary here (Rhinow and Kühlbrandt, 2008; Yoshioka et al., 2010; Russo and Passmore, 2014).

Lacey carbon grids (Agar Scientific) were the grids of choice for the majority of picornavirus samples described in this thesis. They comprise an ultra-thin layer of continuous carbon; a particularly desirable feature for samples that tend to have a high affinity for carbon, but at the cost of consistent hole size and quality. This additional continuous carbon layer can markedly degrade contrast, impacting particle alignment during processing, but, for larger specimens such as the viruses and VLPs used in our studies, this was not an issue (Drulyte et al., 2018). In addition, randomly distributed 'lace' across lacey grids makes data collection set up much longer (since only crude template acquisition spots per grid can be used), and the number of useable areas highly variable. It has also been found that, in some cases, the ultra-thin carbon layer can be folded over in some sections, and thus carbon thicknesses can vary.

Some provisional analyses with Quantifoil grids (Quantifoil) with a continuous carbon layer, reported to be slightly thinner than that of lacey grids, were attempted, but we have yet to achieve good ice quality for data collection, likely due to sub-optimal glow discharge times. Graphene-

oxide coated grids have also yet to be explored.

Another alternative grid type, C-flat, was used for some PV and ERAV sample preparations (Chapters 6 and 7). The quality control of C-flat grids (Protochips) is far superior than that for Quantifoils, as evidenced by the standard deviations (S.Ds) in ice thicknesses, but their tendency to shatter, particularly during clipping, has made them less popular for screening on microscopes requiring grids to be clipped to form a 'cartridge' for loading, as for the Titan Krios and Glacios microscopes (Cho et al., 2013). Measurements of ice thicknesses for a variety of commonly used grids show that Quantifoils are best suited for larger biological specimens (> 100 nm), while lacey and C-flats are suited for smaller particles (Cho et al., 2013). Of these three grid types, EELS measurements of carbon film thickness demonstrated that Quantifoils hold the thickest carbon film layer (49.11 ± 8.50 nm; mean ± S.D.), followed by lacey carbon grids (28.36 ± 2.95 nm, 20 nm in other studies) and finally C-flats (reports of 17.3 2± 0.82 nm and 10-20 nm in other studies), with ice uniformity correlating with that of carbon thickness (Cho et al., 2013; Ermantraut et al., 1998; Quispe et al., 2007; Karlsson, 2001).

Generally, for each new batch of grids received, grids were checked for quality under a light microscope. Sometimes it was even possible to see by eye disrupted support films.

Glow Discharging

Prior to sample application, it is necessary to pre-treat carbon support films with a low-energy plasma to render the surface hydrophilic. This was achieved using air (glow-discharge) for samples in this study, but plasma cleaning, where a defined gas mixture is used, is an alternative strategy that is reportedly more reproducible than glow discharge units. In both approaches, residual gas in a chamber *in vacuo* is charged, generating radicals that react with the grid surface, thereby charging the carbon and removing surface contaminants, and the resulting hydrophilic surface should ensure a uniform sample distribution and ice thickness (Aebi and Pollard, 1987).

For Quantifoil grids, lacey grids and formvar carbon grids (used for negative stain), glow discharging was performed for 30 s on 'high'. C-flat grids were glow discharged for a shorter time, 20 s.

Blotting and Vitrification

Suitable ice thickness for cryoEM is typically achieved through blotting of excess liquid. This is poor for maintaining sample integrity and wastes much sample. Suitable ice thicknesses follow the goldilocks principle; thin ice can be detrimental to sample integrity, particularly in the case of virus specimens, but too thick ice impedes contrast. Blotting is often assumed to be non-invasive to the biological sample, but the sample imaged on a cryo-EM grid does not necessarily reflect its state in the eppendorf pre-application. Samples must endure the shear forces associated with blotting excess buffer through the grid and undergo some dehydration as water evaporates post blotting. In addition, blotting paper may be a potential toxic agent for some samples (Arnold et al., 2017) and some have used ash-free filter papers to minimise additives such as divalent metal ions to compromise specimen integrity (Parmar et al., 2018).

The majority of picornavirus samples reported in this thesis, unless stated otherwise below, were prepared using a Vitrobot, Mark IV (Thermo Fisher) with a blot force of -15, blot time of 3.5 s, 100 % reported humidity and at 4 °C with lacey ultrathin Carbon grids (Agar Scientific), blotted with Whatman filter papers.

For PV-VLP grids, 4 μL of yeast VLP PV-1 SC6b or PV-3 SC8 sample was immobilised onto glow-discharged C-flat 2/1 (Protochips), or lacey 400 mesh with an ultra-thin carbon layer (Agar scientific), respectively.

Moving away from Blotting

Despite exceptional leaps forward at the data collection and processing stages over the past decade, grid preparation remains stubbornly in the dark ages. Blotting excess liquid on EM grids is not easily reproducible, causes the loss of over 99 % of sample, is slow, and often leads to preferential particle orientation, frequently at the air-water interface which can massively disrupt delicate sample preparations (Arnold et al., 2017; Parmar et al., 2018).

Numerous blot-free methods are now emerging, with the 'spotiton', or chameleon now commercially available (https://www.ttplabtech.com/products/sample-preparation/chameleon/). This utilises 'self-blotting' grids with a piezo-electric transducer with a microcapillary (akin to an ink-jet printer), which dispenses a stream of droplets of tens of picolitres onto a grid, such that only 20-50 nL of sample is required per grid (Jain et al., 2012; Razinkov et al., 2016). Nanowires distributed

across the bespoke grids allow droplets to spread evenly to create a uniform ice (Wei et al., 2018). Recently, a spotiton was installed at eBIC, Diamond. Typically, a single line of uniform ice across the grid, 1-2 grid squares in width, can be seen on the grid, and thus no time is wasted screening grids for suitable ice thicknesses.

An interesting variant that has come to light very recently is the 'Shake-it-off' device, which uses self-wicking grids coupled with a piezoelectric sprayer and plunging solenoid, all made from readily available components, and simply controlled by a Raspberry Pi (Rubinstein et al., 2019). The 'Cryowriter' is similarly conservative, where a thin needle 'draws' a total of 2-20 nL of sample onto a suspended grid by microcapillary action (Arnold et al., 2016, 2017; Kemmerling et al., 2013), and has been successfully applied to the determination of a human 20S proteasome structure to a ca 3.5 Å resolution from only 1 μL of cultured cells (Schmidli et al., 2019).

Microfluidics or voltage-assisted spraying devices, typically used for time resolved crystallographic studies, may also be used to control ice thicknesses via the distance between the spray nozzle and grid and the pressure applied (Feng et al., 2017; Kaledhonkar et al., 2018; Kontziampasis et al., 2019).

Titan Krios and Glacios Microscopes (Thermo Fisher)

Demand for Titan Krios microscopes has risen massively over the past five years, far surpassing the those of other manufacturers, largely because no other 300 kV microscope offered the level of automation and user friendliness.

The Glacios, a more compact version of its predecessor the Talos Arctica, operates at 200 kV, but has the same autoloader arrangement and similar objective optics to the Krios. It is largely sold as a screening machine, however it is suitable for high resolution data collection although 300 kV instruments are often favoured for high-resolution data collections due to the reduction in inelastic scattering (which causes radiation damage and adds to background noise). Lower voltage instruments can provide better specimen contrast and a recent collection that I have omitted from this thesis of HRV-1B attained a resolution of 3.3 Å, despite limited usable acquisition areas, while attained resolutions as low as 2.0 Å (Danev et al., 2019) and 2.7 Å (Hamdi et al., 2019) have been published for high- and low-end 200 keV microscopes, respectively, and claims of achieving resolutions as low as 1.8 Å have been made (personal communication, Abhay Kotecha). This is

competitive with a 300 kV machine; for instance, octahedral ferritin achieved to 1.6 Å resolution on a Titan Krios (Danev et al., 2019).

Both microscopes share an almost identical method for sample insertion. Before loading, vitrified grids must be first clipped using a 'C-clip' and 'autogrid rim', forming a 'cartridge'. This serves to bolster otherwise fragile grids (note, however, that the clipping stage itself can shatter grids) and enables reliable mechanical manipulation of loaded grids by the microscope. Interestingly, it is claimed that clipping of grids prior to vitrification is possible (Peter Peters, Strubi group meeting 2018). Cartridges can then be placed into a cassette, a multi-specimen holder, which is then transported to the microscope in a small dewar, called 'nanocab'. The nanocab is simply slotted into the microscope core, and the cassette is retrieved by the autoloader. Problems can and do occur during loading, and the autoloader has been found to be the Achilles heel in both the Strubi Titan and Glacios microscopes, perhaps unsurprisingly given the number of moving parts.

Cryo-EM experiments presented in this thesis were conducted on either the Titan Krios G3 or Glacios microscopes at Strubi, with some exceptions (see Table 2.3 in Section 2.5.3 below).

2.5.3 Data Acquisition

All data collections within Strubi were set up using EPU (Thermo Fisher) using the standard pipeline, except that eucentric heights for each selected grid square for acquisition were determined using TUI (Technai microscope control User Interface) and not with the automated procedures implemented in EPU, which often failed for lacey grids. EPU2 eucentric height procedures seem to be more robust (personal communication, Bilal Qureshi), and may be adopted in future. Stigmation and coma-free alignments were performed using a beta 'AutoCTF' software that worked in conjunction with EPU. This has since been incorporated into the new EPU2 software.

Detectors

Two detectors were used for single-particle microscopy studies; FalconIII (FIII, Thermo Fisher) and the K2 (Gatan). FIII and K2 are examples of direct electron detectors (DEDs), the major technology advance responsible for the cryoEM 'resolution revolution'. Briefly, a complementary metal-oxide-semiconductor (CMOS) chip detects incident electrons and converts them to charge. DEDs have

Table 2.2: Specifications for FIII, K2 and K3 detectors, as quoted on respective supplier websites. Ny = Nyquist in this instance; EC = electron counting. Note that the K2 and K3 detectors are more radiation hard, but at the cost of DQE since the chip is thicker, and is often thought to be most appropriate for smaller particle data collection, since the energy filter improves contrast, while the DQE at lower resolutions is higher at high magnifications.

	Falcon III	K2 Summit	K3
Supplier	Thermo Fisher	Gatan	Gatan
Sensor size / pix	4096 x 4096	3838 x 3710	5760 x 4092
Pixel size / μm	14	5	5
Frame rate / frames s^{-1}	40	400	1500
DQE at 0.5 Ny EC (1 e^-/pix/s), 300 kV	0.60	0.55	0.53

a significantly lower back-scatter of electrons that tend to compromise high-resolution image information relative to their CCD forebears.

Four fundamental characteristics that differentiate these detectors are: the read out speed, the area of the detector (pixels that can be usefully distinguished from one another), detective quantum efficiency (DQE), and the modulation transfer function (MTF).

DQE reflects the probability that an incident electron will produce a corresponding signal increase on the detector, thereby reflecting the signal dampening effect of a detector (Table 2.2, Equation 2.2; SNR = signal to noise ratio; a perfect detector would give a DQE of 1), and varies with resolution (and more broadly, spatial frequency). Detector DQEs vary with exposure (i.e. $e^-/\mu m^2$), and at higher exposures, before saturation, detector noise become relatively insignificant and DQE reflects the proportion of electrons producing a detectable signal. DQE values are generally specified at half the maximum frequency, or half Nyquist (Glaeser, 2007; McMullan et al., 2016).

MTF refers to the decline in signal strength relative to that of the input as a function of the spatial frequency, and is determined by factors such as pixel size, and the spread of electrons across the detector (Glaeser, 2007). Both FIII and K2 detectors also have a high dynamic range, meaning that both low and high pixel values can be accurately recorded on the same frame. Characteristics of the two detectors are summarised in Table 2.2.

$$DQE = \frac{SNR(output)^2}{SNR(input)^2} \tag{2.2}$$

Data collection can be run using 'linear' (integrating), or 'counting' modes. This simply means

Table 2.3: Summary of data processing parameters used for single-particle cryo-EM data collections. All collections were conducted in Strubi using a 100 μm objective aperture. The "Strucs" column refers to the number of refined reconstructions from a given collection presented in the given chapter. For the two respective modes, L = linear/integrated and C = counting mode. Mag = magnification. The defocus step used was 0.3 μm for all listed collections.

Dataset	Chapter	Microscope	Nominal Mag / kX	Pixel size Å/pix	Camera	Mode	Defocus range / μm	Total Dose $e^-/Å^2$	Strucs
ERAV	7	Krios	59	1.36	FIII	L	0.9-3.0	50.8	3
EV-F3 (heat)	3	Glacios	92	1.54	FIII	L	1.0-2.8	42.1	1
A22cc	5	Krios	130	1.05	K2	C	0.9-3.0	41.0	1
A22cc, $\alpha_v\beta_3$	7	Glacios	120	1.22	FIII	L	0.9-3.0	38.5	2(3)
A22cc, DTT	5	Krios	75	1.06	FIII	C	0.9-3.0	30.0	1
AsiahShcc	5	Krios	75	1.06	FIII	C	0.9-3.0	30.0	4
AsiaSh, DTT	5	Krios	75	1.06	FIII	C	0.9-3.0	30.0	3
PV-1	6	Glacios	120	1.22	FIII	L	0.9-3.0	51.7	2
PV-3	6	Glacios	92	1.54	FIII	L	0.8-2.9	40.9	2

that the signal of electrons reaching the detector is integrated, or individual electron events are counted, such that resulting images either show integrated signals or a sum of electron count. The former approach offers a more rapid readout, but at the cost of attainable resolution. Counting mode collections are much slower, especially for FIII cameras, since per frame exposures must be limited such that individual electron events may be detected and thus longer exposure times are needed to achieve desired total doses. Table 2.3 lists the modes used for cryoEM data presented in this thesis. A really interesting comparison between results from counting and linear mode collections on the FIII are presented in Song et al. (2019), where it was shown that, for a 4.8 MDa sample (cf picornaviruses 8-9 MDa), no significant differences in the reported resolutions from counting and linear mode collections were observed, whereas, for a much smaller complex (465 kDa), a substantial difference was found.

The FIII was used for most structures in this thesis owing to its availability on both Glacios and Krios microscopes in Strubi, and my lack of training on the K2.

2.5.4 Data Processing

The resolution revolution sparked a keen interest in cryo-EM data collection, and has spurred a revolution in user-friendly cryo-EM software. Of the multitude of software available, Relion and cisTEM have been the most widely used in the last few years (Scheres, 2012; Zivanov et al., 2018; Grant et al., 2018). Cryosparc, designed for GPU-based processing, has also gained much popularity owing to its rapid processing speed and straightforward user interface, and hands-off, on-the-fly processing pipeline (Punjani et al., 2017). All programs follow roughly the same pipelines, guiding users from movie alignment and motion-correction through to particle polishing.

In this work, the majority of data processing was performed using the latest version of Relion (3.0), with particle picking mostly performed using Warp (see below) (Scheres, 2012; Zivanov et al., 2018; Tegunov and Cramer, 2018). Cryosparc was often used in parallel to get a quick initial reference for reconstructions, and to judge the success of data collections (Punjani et al., 2017). Motion correction was performed with MotionCor2 and contrast transfer function (CTF) parameters were determined by Ctffind4.1.5 (Zheng et al., 2017; Rohou and Grigorieff, 2015). 2D classification of extracted particles was performed with Relion (Zivanov et al., 2018). Initial references for reconstructions were derived either using Relion, or Cryosparc (Scheres, 2012; Punjani et al., 2017; Zivanov et al., 2018).

Particle Picking

Warp is a neuronal network-based particle picking program (Tegunov and Cramer, 2018). Briefly, movies are loaded, aligned and CTF-estimations performed on the fly before AI-driven particle picking. For FMDV VLP datasets, fifteen movies were initially loaded into the Warp GUI, and particles were manually picked, while a mask was drawn on areas with thick lace. This was then used to run a round of training and the resulting algorithm assessed on a subset of typically 20 images. It was checked that the majority of particles were picked, excepting overlapping particles. Sometimes it was necessary to run another round of training with manually selected particles. This learning was then applied to the entire dataset. For the latter VLP datasets, one further round of training on a subset of 10 particles was required. For most datasets, particle picking was checked in random selections of micrographs, and, occasionally, erroneous particles were removed, or potential particles added. Plots of particle numbers were inspected and where an unusually high

or low number of particles were selected, were investigated.

Heated EV-F3, FMDV-integrin VLP and PV-3 VLP datasets all used this initial VLP training as a basis, which was re-trained on a subset of 10 images. For heated EV-F3, it was necessary to repeat this training twice more, on different subsets of images, so as to ensure that all particles, whether filled or empty were picked. In this latter case, all micrographs were quickly inspected following the Warp run, to ensure that no particles were missed.

PV-1 VLP micrographs required additional training of Warp before sensible particle selections were made, likely due to the fact that C-flat grids were used in this case. As with EV-F3 heated samples, every micrograph was quickly inspected after the Warp run and any missed particles were selected manually.

ERAV particles were picked manually using the Relion 'Manual Picking' GUI (Zivanov et al., 2018).

The Relion Pipeline

Coordinates from Warp were then copied over to local Relion processing directories, but not motion and CTF-corrected images. This was because motion and CTF correction performed in Relion usually resulted in a better resolution (and looking) map, and Relion3 particle Polishing required motion correction to be performed in Relion (Zheng et al., 2017; Zivanov et al., 2018). Simple scripts were made to convert coordinates to a Relion-friendly format. Specifically, removing additional columns and headers not recognised or needed by Relion (Scheres, 2012; Zivanov et al., 2018).

Exposure of samples to incident electrons triggers a number of processes that collectively cause subtle rotational and translational movements of particles embedded in the vitreous ice layer, resulting in a blurring effect. It is possible to correct for this by tracking particle movement over the course of a series of movie frames, and doing so greatly improves image quality (Zheng et al., 2017). This may be achieved using MotionCor2.

For CTF correction, CtfFind4 was used with exhaustive searches on aligned movies without dose weighting (this has been claimed to provide better CTF estimates). Following CTF correction, fits were inspected within the Relion GUI, and any that were appreciably poor fits were excluded (Scheres, 2012; Zivanov et al., 2018).

2D classification

Relion utilises maximum likelihood (ML) approaches for classification. For each particle, the probability of it being in every possible orientation is calculated iteratively and the particles are x- and y-shifted as well as in-plane rotated. 2D classification was typically used to simply filter out broken, blurry or non-existent particles, and the number of classes set as 50-100, depending on the number of micrographs used, and the anticipated particle heterogeneity. Particles were deemed acceptable by visual inspection and based on rotational and translational accuracy scores (ideally below approximately 2). Where accuracies seemed poor, as was the case for the ERAV dataset, an additional 2D classification was performed, and particles were selected based on visual inspection alone, before further classification. For PV-1 and PV-3, the rotational accuracies of these classes were poor, and, particles selected at the 2D classification stage were instead aligned, with minimal or no local searches before classification. This procedure worked very well, and enabled the distinction between expanded and non-expanded particles for the PV-1 dataset.

Particles for PV-1 and EV-F3 heated were then auto-refined with an initial reference generated *ab intio* in Warp before low pass-filtering in Relion, and, because imported references were not on an absolute greyscale, the Relion flag to correct for this was selected in each instance (Scheres, 2012; Zivanov et al., 2018; Tegunov and Cramer, 2018).

The reference and masks used for refinement and post-processing were iteratively improved by fitting coordinates in Chimera, and using the 'colorzone' model to extract the map fitting the coordinates (Pettersen et al., 2004). Reference masks for post-processing were generated using the Relion MaskCreate module, with an extension of 5-8 pixels, and a raised-cosine soft edge of 8 pixels. The initial binarisation threshold was determined based on appropriate contour levels in Chimera (Pettersen et al., 2004). Masks were then checked in bshow prior to application (Heymann and Belnap, 2007; Heymann, 2001).

PostProcessing

Auto-refinement provides a resolution estimate based on a 'gold standard' Fourier Shell Correlation (FSC) between 3D volumes, but using a circular mask with a user-input diameter. This is generated by refining two models, one for each half of the data, independently (by contrast to standard FSC procedures where just one model is used for particle orientations for the two halves of the data (Saxton and Baumeister, 1982)), so as to reduce inflation of resolution estimates by

noise (Henderson et al., 2012; Scheres and Chen, 2012). These resolution estimates tend to be extremely pessimistic since they are based on a volume much greater than the protein density, and this disordered density will dilute resolution estimates. To provide a more 'realistic' resolution estimate, post-processing is performed, where half datasets are combined into a single map and a tight mask is applied to the structure. The data are also B-factor weighted and detector MTFs used for B-factor estimation. Selection of a reasonable mask is therefore essential for the success of this procedure. Particles were picked using Warp (Tegunov and Cramer, 2018). Movies were initially motion and ctf-corrected using Relion3.0's wrappers for MotionCor2 (Zheng et al., 2017) and Ctffind4.1.5.(Rohou and Grigorieff, 2015).

CtfRefinement and Bayesian Polishing

Relion3.0 has also implemented new CtfRefine (per-particle CTF correction) and Bayesian Polishing modules (Zivanov et al., 2018). For our datasets, following an initial round of refinement, particles were subjected to CtfRefine with, in the case of VLP datasets, where particle numbers were much higher, per micrograph astigmatism correction. CtfRefined particles were then re-refined, and this procedure (refinement->post-process->ctf-refinement/polish) was run iteratively until no map/resolution/stated B-factor improvements could be seen. Initially both Polishing and CtfRefine procedures were run in parallel, but it was found that appreciable resolution and map improvements were generally attained when an initial CTF refinement was performed, and, only later, polishing.

For some of the smaller datasets, no improvement in map quality, or quoted resolution was observed with either procedure. Numbers of CTF refinement and polishing cycles are summarised in Table 2.4.

2.5.5 Model Fitting

Models of a single protomer were initially fitted using Chimera, and then through rigid body fitting of an entire protomer in Coot (Emsley and Cowtan, 2004; Emsley et al., 2010). The model was then extended to an 8mer, such that the central protomer was surrounded by at least one other protomer.

Phenix real-space refinement was used with local searches with iterative model building. In cases where resolution was poor and/or initial clash scores were high, hydrogen atoms were added

Table 2.4: Summary of the number of per particle CTF and Bayesian Polishing cycles performed for each dataset presented in this thesis. Cycles of first CtfRefine, refinement and then Polish and CtfRefine in parallel before further refinement were performed, and, at each stage, assessed for improvements in map quality (visual inspection in Chimera), reported resolution and B-factor. Where statistics and map quality no longer improved, cyclic refinement was stopped. For datasets were over 10,000 particles were used, polishing was performed using a trained parameter set, and, an additional astigmatism correction was included in the first round of CtfRefine.

Sample	# Particles	CtfRefine	Polish
FMDV A22cc	19961	2	1
FMDV A22cc DTT	34743	1	0
FMDVAsiacc (U-i)	36288	2	2
FMDVAsiacc (U-ii)	11856	2	1
FMDVAsiacc (U-iii)	11396	2	1
FMDVAsiacc (U-iv)	7701	1	0
FMDVAsiacc DTT (D-i)	9278	2	0
FMDVAsiacc DTT (D-ii)	2211	2	0
FMDVAsiacc DTT (D-iii)	4047	2	0
FMDV AIRN	6896	2	0
FMDV Otur	4913	2	0
EV-F3 B	8104	2	1
EV-F3 I	783	1	0
EV-F3 A	1620	2	0
PV-1 C	619	0	1
PV-1 D	242	0	0
PV-3 I	6749	0	1
PV-3 II	3539	0	0

to the model before each cycle, and were removed prior to determination of the final statistics.

Following each round of refinement, the most central protomer was inspected for major clashes, and overall density fit in Coot, and refined as required. This was then regenerated into an 8mer using phenix.apply_ncs, with NCS matrices generated from the previous 8mer with phenix.simple_ncs.

Goodness of fit of map to model was assessed by visual inspect in Coot, and CC for side chains, main chains on a per chain basis (as determined by the Phenix validation module). Stereochemistry was assessed using MolProbity scores.

Namdinator

For the medium resolution FMDV A22cc 10 mM DTT treated pentamer structure presented in Chapter 5, substantial structural rearrangements were observed in the VP3 in particular, which were fitted into the map using a Molecular Dynamics Flexible Fitting (MDFF) simulation followed by Phenix real space refinement using the Namdinator pipeline (Kidmose et al., 2019) online (website: https://namdinator.au.dk).

Prior to uploading to Namdinator, a protomer from the crystallographic structure of FMDV A22cc (PDB: 4IV3, Porta et al. (2013a)) was rigid body fitted into the pentamer map using Chimera, and then checked again in Coot. The PDB was uploaded to Namdinator along with the associated map.

As recommended in the Namdinator documentation (https://namdinator.au.dk/manual/), a starting and final temperature of 298 K were used, with a 0.3 G-force scaling factor, 2000 minimisation steps, 20,000 simulation steps and 5 Phenix real-space refinement cycles. Implicit solvent was excluded. Details and statistics are presented in Chapter 5, Section 5.2.2.

The resulting fit was then inspected in Coot, and extended to a pentamer, before two more iterations of real space refinement in Phenix (Adams et al., 2010, 2011).

2.6 Time-Resolved Cryo-EM

'Caging', in the context of photoactivatable protecting groups for biological molecules, was coined by J.F. Hoffman in 1978 in reference to 'caged ATPs'. Caged compounds comprise an often photolabile 'protecting group' appended to a compound of interest (Kaplan et al., 1978), such that irradiation at a certain wavelength promotes irreversible separation of this photolabile group and

target molecule. Although at the time photolabile protecting groups was old-technology in the context of synthesis, Hoffman et al. pioneered the notion of their employment in spatiotemporal control in biochemical experiments, which has since been applied by other groups (Adams and Tsien, 1993; Kaplan et al., 1978; Kawahama, 2007; Kaplan, 1990; Kaplan and Somlyo, 1989). Since then, a variety of biologically relevant groups, such as cAMP, have been synthesised with activatable caging groups (Engels and Schlaeger, 1977; Walker et al., 1988).

Controlled proton release is of interest for our studies, with a view to capturing FMDV capsids in transient uncoating states (see Chapter 7).

2.6.1 Caged compound Control Experiments

To establish the suitability of NPE-caging for proton release, a sample holder for test solutions was designed using AutodiskFusion360 software, and 3D-printed (https://www.autodesk.co.uk/products/fusion-360/).

For standard curves, stock solutions containing Succinic Acid, Sodium Dihydrogen Phosphate and Glycine (SPG, Molecular Dimensions) buffer at different pHs, and, unless otherwise stated, 25 % v/v of 0.04 % w/v bromothymol blue (BTB, Sigma Aldrich) were prepared, and mixed thoroughly. A water background was taken for each set of readings, with 30 accumulations of 0.01 s to minimise errors in background estimation. For each reading, 4 μL aliquots of SPG/BTB stock solution were placed onto a custom-designed 12-well plate, and 30 accumulations of 0.01 s over a wavelength range of 400-800 nm were taken. For cage-compound solutions, solutions containing HEPES pH 7.5 buffer, caged compound, and BTB were made directly before measurement. An aliquot of 4 μL solution was also pipetted on the custom made well-plate, and, at the fifth cycle, solutions were exposed to a 280 nm UV light, for a given time. The LED exposure was controlled by a fast shutter, kindly lent by the I24 beamline team at Diamond. Spectroscopic readout was taken using a 'kinetic' measurement of 50 cycles of 0.545 s (for 30 accumulations of 0.01 s for each), and over a wavelength range of 400-800 nm was taken.

Measurements were recorded on an Andor detector for both control, and caged experiments, in 'Full Vertical Binning'. Raw data were exported as 'asc' files using Andor Solis software (https://andor.oxinst.com), and parsed using a combination of R and Python2.7 scripts. Since the background varies on a per reading basis for micro-spectrometers, it was necessary to correct the background by subtraction

of the average absorbance readout for the last 100 wavelengths for each set of measurements.

For pH estimation of caged compound solutions, absorbance readings across three repeats were background corrected, and averaged. The $\log_{10}$ ratio between the two peaks at 450 nm and 620 nm was then determined. The initial pH was assumed to be that of the HEPES buffer, pH 7.5. The absorbance for the first cycle was adjusted to the absorbance for this pH, and the same correction was applied to the remaining readings (see Chapter 7).

2.7 Thermofluor Assays

During a thermofluor assay (aka thermal shift assay; differential scanning fluorimetry, DSF) (Pantoliano et al., 2000; Ericsson et al., 2006; Geerlof et al., 2006), protein sample is slowly heated and its state monitored using a reporter fluorescent dye, typically a repurposed gel stain such as SYPRO orange (Invitrogen), whose fluorescence quantum yields increase upon a particular environmental change, such as exposure of hydrophobic residue patches as proteins unfold. Thermofluor experiments are conducted using a quantitive PCR (qPCR) machine and for the purposes of this thesis, a MX3005p RT-PCR instrument (Agilent) was used.

This method has recently been adapted to measure the thermal stability of virus capsids (Walter et al., 2012; Kotecha et al., 2016) . Termed 'PaSTRy assay', virus capsids are heated with nucleic-acid sensitive dyes to detect the release of viral genome in addition to the conventional hydrophobicity-sensitive dyes.

The particular nucleic acid dyes selected for PaSTRy were done so on the basis of their being sufficiently large to prevent viral capsid penetration, sufficiently sensitive to detect low concentrations of nucleic acid release and robust against changes in temperature and pH and also to be seperable from the dyes used to detect the exposure of protein hydrophobic surfaces. SYTO9 (Invitrogen) has been observed to be quenched more readily than SYBR green II (Invitrogen), purportedly as a consequence of SYBER greenII being a stronger RNA binder than SYTO9. However, SYTO9 is often favoured for PaSTRy assays owing to its further removed emission spectrum relative to SYPRO orange (Walter et al., 2012).

Conventionally used dyes sensitive to hydrophobic environments include SYPRO orange (optimally excited by the 540-700 nm) and SYPRO red. Nucleic acid sensitive dyes include SYTO9 and

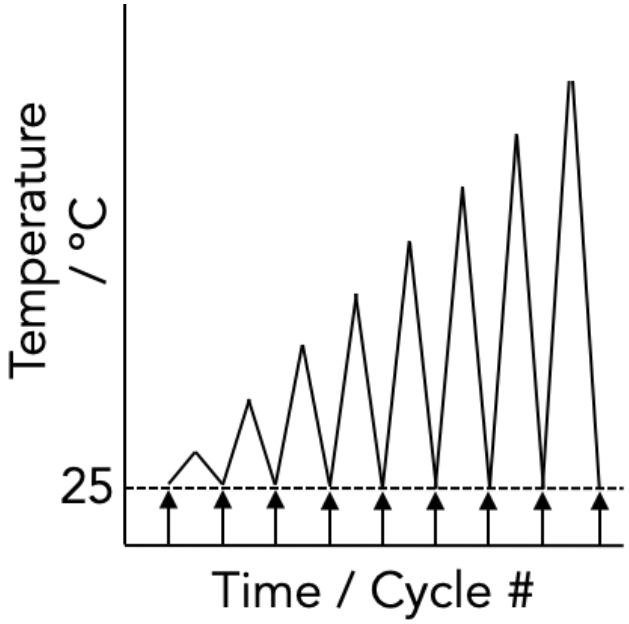


Figure 2.6: Thermal profile for thermofluor cycles used in this thesis. Arrow correspond to when a reading is taken. The saw-tooth shows the temperature ramp up to a given temperature, followed by cooling to 25 °C upon cooling. This is not to scale.

SYBR green (I or II depending on whether DNA, or RNA is to be detected, respectively), and all, except for SYBR green I, were used in this work.

For the majority of cases, 5000 x stocks of SYPRO orange and SYPRO red, and 5 mM stocks of SYTO9 and 10,000 x of SYBER green II in DMSO were used.

Assays were set up in a thin-walled PCR plate (Agilent) and, unless otherwise stated, 1 μg of virus or VLP was added to wells containing the desired concentration of dye (5 mM SYTO9 and 3 x SYPRO Red or 3 x SYPRO Orange, unless otherwise stated in text) with 50 mM HEPES, pH 8 200 mM NaCl buffer to a final volume of 50 μL. GloMelt^{TM} (Biotium), used in Chapter 5, was also provided in DMSO and diluted to a 10 x working stock, and diluted as described in Chapter 5, Section 5.2.4.

A saw-tooth thermal profile was used to minimise the collisional quenching effects that are seen at high dye concentrations (Figure 2.6). This means that, following each temperature increase, the sample was cooled back to 25 °C before a recording was taken.

Prepared wells were then mixed by pipetting and covered with optically clear foil (BioRad Micro-seal B, MSB-1001) before centrifugation at 4 °C for 1-2 min at 2000 g to remove any bubbles. Every condition was performed in triplicate.

Temperatures were ramped from 25 to 99 °C and fluorescence measurements recorded at 1 °C intervals.

For analyses, the inflection point, corresponding to the melting temperature, was determined using a five parameter fit of the average of the three first derivative denaturation curves using Prism (GraphPad Software version 7 for Mac, La Jolla California USA, www.graphpad.com), with a precision of typically < 0.5 °C. This was chosen in preference to a four parameter fit, unless otherwise stated, since this takes into account the asymmetry seen in melting curves and thus should provide a more accurate estimate of the inflection point.

2.8 Illustrations

Sequence alignments were performed using Clustal Omega (Sievers et al., 2011) and either edited and displayed with nedit (SourceForge), MAFFT (a multiple sequence alignment program) or with the program ESPript (Katoh et al., 2017; Gouet et al., 1999).

Twodimensional interface diagrams were prepared using LigPlot+ (Wallace et al., 1995; Laskowski and Swindells, 2011).

Structures were presented using coot (Emsley and Cowtan, 2004; Emsley et al., 2010), PyMOL (DeLano et al., 2002), Chimera (Pettersen et al., 2004) or ChimeraX (Goddard et al., 2018).

Chapter 3

EV-F3: from Entry to Uncoating

3.1 Introduction

Bovine enteroviruses (BEV), spanning species Enterovirus E and Enterovirus F, are a group of largely non-pathogenic members of the genus Enterovirus (see Chapter 1, Section 1.1.6), and, as such, serve as functional surrogates for the study of more pathogenic enteroviruses (Jiménez-Clavero et al., 2005; Smyth et al., 1995).

As described in Chapter 1 Section 1.1.6, BEV species have recently been reclassified into two clusters, each containing two and three geno/serotypes, respectively (Zell et al., 2017). Enterovirus species names were then changed to remove reference to the host species name, and the bovine enterovirus clusters became species Enterovirus E and Enterovirus F (Zell et al., 2017). Structures are available in the PDB for both species; EV-F3 (PDB: 5OSN, determined using XFEL data) and EV-E1 (PDB: 1BEV) (Smyth et al., 1995; Roedig et al., 2017) and may represent different species, rather than just separate serotypes as suggested previously (see Chapter 1) (Dunne et al., 1974; Knowles and Barnett, 1985). Indeed, previously identified neutralising antigenic sites for EV-E1 occur at (I) VP1 GH-loop: D1198-Y1203, (II) VP2 EF-3.10 helix: Y2150- G2159; (III) VP3 BC-loop: E3054-G3063 (Smyth et al., 1990, 1995; Smyth and Martin, 2001), and, when compared with EV-F3, no cross-neutralisation was observed for site (I), nor site (III) (Smyth et al., 1992). However, some evidence of cross reactivity was observed for site (II) (note that site (II) is similar to that of both PV-1/3 site 2b and HRV14 neutralising site NIm-IIA and thus is likely to be less specific) (Minor et al., 1986; Sherry et al., 1986). EV-E1 and EV-F3 share 64 % overall sequence identity between

structural proteins VP1-4 (Uniprot ID: P12915, Q65480) (Madden et al., 1996).

To use EV-F3 as a model system for other enteroviruses, it is necessary to establish similarities and differences to existing enterovirus structures, both for native and uncoating intermediates; differences that occur at these intermediate stages may inform the diversity within the enterovirus group.

In the following chapter, structural characterisation of EV-F3 at different stages during infection will be described, specifically that of the mature virion (for which no published structural description exists), and 'natural' empty particles that are devoid of genome and have yet to undergo maturation cleavage of VP0 into VP2 and VP4. This is in addition to heat-induced intermediates, which may represent A-, A/B and B-like particles (Chapter 1, Section 1.1.6).

3.2 Results

3.2.1 EV-F3 production and purification

EV-F3 (PS-87 strain) was grown using established protocols in either BSRT7 or BHK21 cells maintained at 37 °C (Walter et al., 2012; Roedig et al., 2017), harvested three days post infection and clarified by centrifugation. Typical final titres of over 10^9 pfu/mL (for instance, Figure 3.1(a)) were determined by plaque assay for both BSRT7 and BHK21 cells. The supernatant was then concentrated by pelleting the particles through a sucrose cushion, and subsequently purified on a 15-45 % sucrose gradient (Chapter 2, Section 2.2.2), before fractionation and characterisation on a reducing SDS-page gel (see Figure 3.1(b)). Typically, 260/280 ratios of around 1.70 were recorded, confirming the presence of genome (pure RNA should have a 260/280 of 2.00, but absorbance by the viral proteins, particularly phenol rings from Tyr residues which have a peak absorbance at 270 nm, diminish the ratio). Promising fractions were then pooled, desalted and concentrated to over 5 mg/mL for analysis (Chapter 2, Section 2.2.2).

The natural empty particles were prepared by collaborators at the Asbury Centre for Structural Molecular Biology, University of Leeds (Chapter 2, Section 2.2.2).

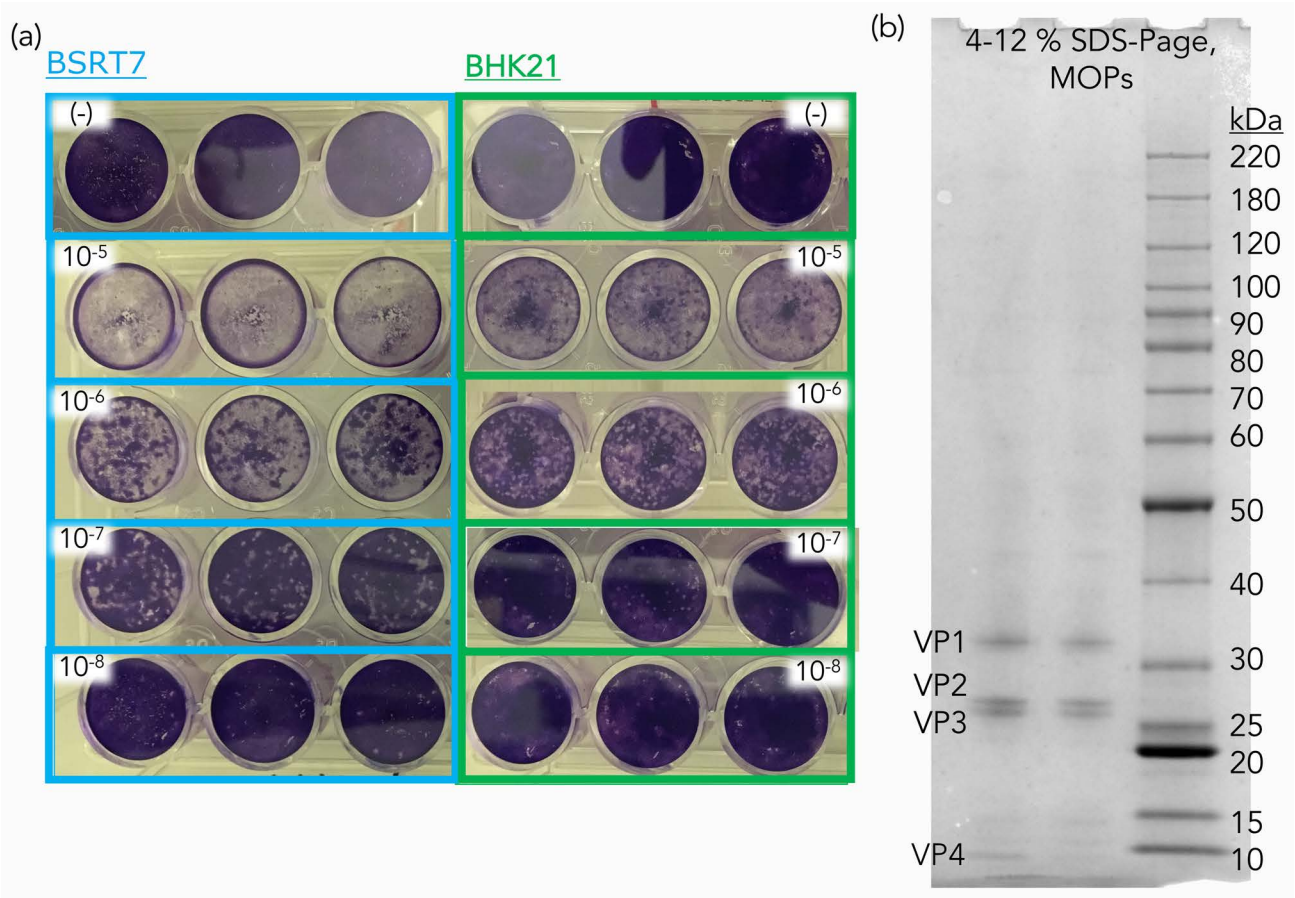


Figure 3.1: (a) Example plaque assays (Chapter 2) for the same EV-F3 sample, performed in parallel with BSRT7 and BHK21 cells, with respective virus dilutions displayed in white boxes. (b) Typical reducing SDS page gel from two 0.5 mL fractions from a EV-F3 purification. Structural proteins are marked on the left hand side of the gel image, MW markers are on the right

3.2.2 Crystallographic structures of Mature and Immature EV-F3

EV-F3 routinely crystallises in space group *I*222 or *F*23, with the particle centre at the crystallographic origin, and 15-fold or 5-fold NCS, respectively. Crystallographic data for mature and 'natural' empty EV-F3 were collected from multiple, isomorphous, crystals at room temperature, in-situ, at a microfocus beamline (I24, Diamond Light Source), and merged (see Table 3.1). The data collection and processing strategy is outlined in Chapter 2, Section 2.4.4.

The mature and natural empty EV-F3 (the latter in parenthesis in what follows) data sets were collected from 254 (54) positions from 88 (48) crystals to 2.10 (2.65) Å resolution. The mature EV-F3 structure was solved by molecular replacement using EV-E1 (PDB: 1BEV) with 4 virus particles per unit cell and one icosahedral pentamer in the crystallographic asymmetric unit (5-fold NCS) (Smyth et al., 1995). EV-F3 mature was determined by positioning the EV-E1 capsid at the origin and aligning icosahedral 2-fold axes along each of the unit cell edges, this left two possible orientations to be tested. For data processing statistics, see Table 3.1.

Overall Structure

EV-F3 adopts the canonical picornavirus organisation, Figure 3.2, as described in Chapter 1, Section 1.1.2. Overall, EV-F3 and EV-E1 share a very similar surface topology, with a slight pitting at the 2-fold axes where the two VP2 protomers meet (Figure 3.2). Respective sequence identities for VP1-4 for EV-E1 and EV-F3 are as follows: 54.3, 75.2, 63.0, 75.0 %. Strands and loops for VP1-3 are labelled in Figure 3.3.

VP1

The most striking differences between EV-E1 and EV-F3 structures occurs at the VP1 N-terminus inside the capsid, and at the surface loops that reside close to the five-fold axes; DE, HI and BC, Figure 3.4.

For all enteroviruses, the VP1 N-terminus, located in the capsid interior, plays an essential role in stabilising intra-protomer and inter/intra-pentamer interactions. EV-F3's modelled VP1 N-terminus (from the first visible residue, 1004, to residue 1013, Table 3.2) traverses the interior of the capsid in the opposite direction to EV-E1, which instead occupies a region modelled as VP4 in EV-F3. A potassium ion site, corroborated by anomalous experiments (see Chapter 4, Section

Table 3.1: Data processing statistics for EV-F3 mature and natural empty room temperature data collection. Values in parentheses are for the highest-resolution shell for data collection.

Data set	Native	Empty
Data Collection		
No. of crystals	88	48
Space group	*F*23	*F*23
Unit cell a=b=c= Å	437.0	436.1
Resolution range / Å	50.0-2.10 (2.18-2.10)	50.0-2.65 (2.74-2.65)
Unique reflections	396379 (18815)	325071(13851)
Completeness / %	91.5(87.3)	82.0(35.1)
Redundancy	3.1(2.7)	1.9(1.2)
Average I/σI	3.8(1.0)	2.4(0.7)
Rmerge	0.194(0.694)	0.476(0.946)
Refinement		
Resolution range / Å	50.0-2.10	50.0-2.65
No. of reflections(working/test)	308141/16466	169354/8998
R-factor (Rwork/Rfree)	0.186/0.197	0.208/0.223
Ramachandran plot / %		
Preferred regions	95.9	95.6
Allowed regions	3.2	3.5
Disallowed regions	0.9	0.9
No. of atoms (protein/other atoms)	6230/433	6223/344
R.m.s. deviations		
Bond lengths / Å	0.01	0.01
Bond angles / °	1.7	1.6

Table 3.2: Residue number from which the N-terminus for each structural protein was modelled for our room temperature (RTP) EV-F3 structure and that of our cryo-cooled crystal derived structures.

VP	N-Terminal Start	
	RTP	Cryo
1	5	4
2	5	1
3	5	1
4	27	22

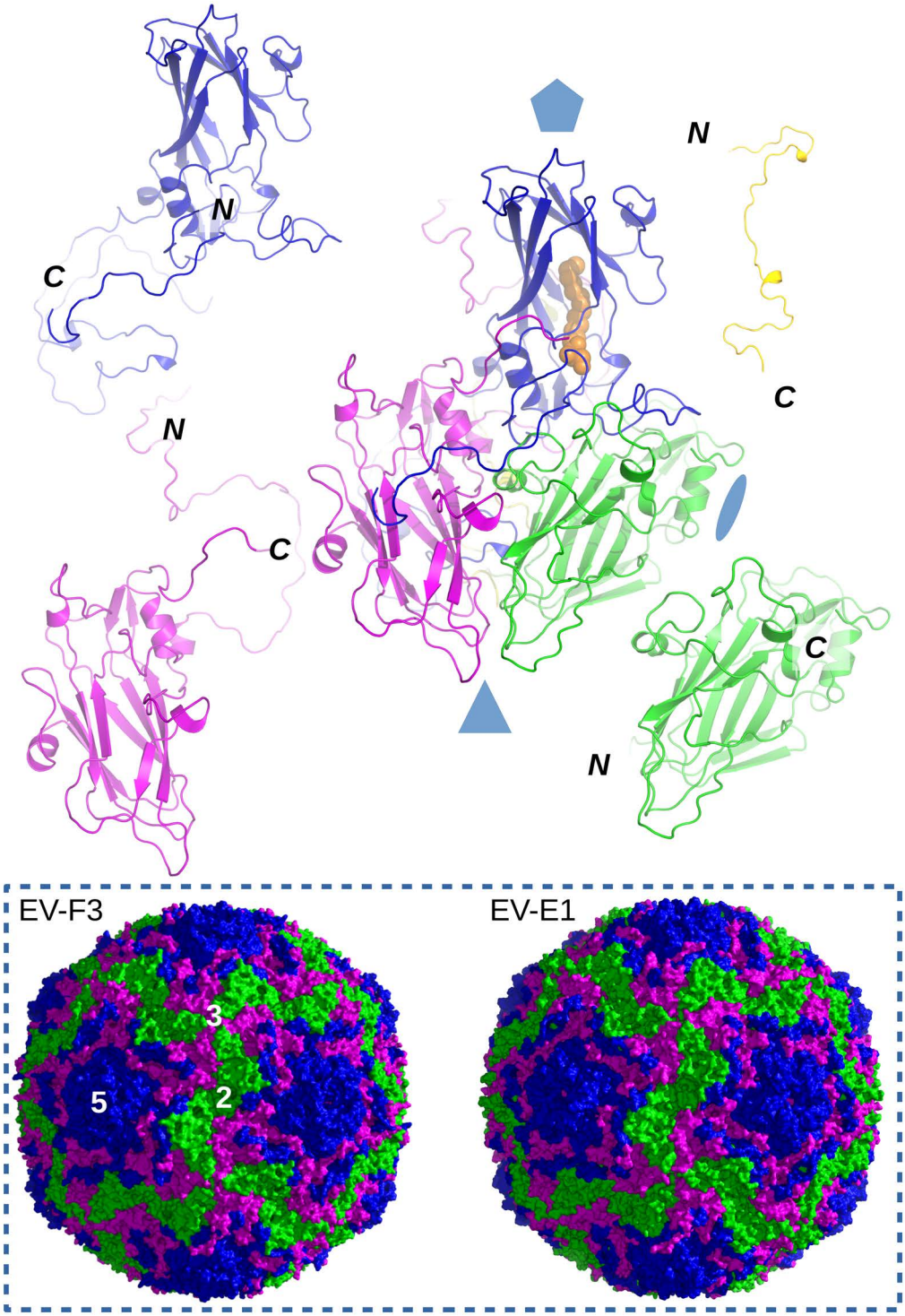


Figure 3.2: Centre top: biological protomer of EV-F3 coloured by VP; VP1 blue, VP2 green, VP3 magenta and VP4 yellow. Within the P-site, a modelled sphingosine pocket factor is shown by orange spheres. This protomer is surrounded by individual images of EV-F3's four structural proteins, with N and C termini labelled in bold itallic text. Each follow the canonical fold seen for all picornaviruses. Blue dashed box: Surface representation of EV-F3 with symmetry axes labelled in white text. The surface for an EV-E1 model (PDB: 1BEV) is shown on the right (Smyth et al., 1995).

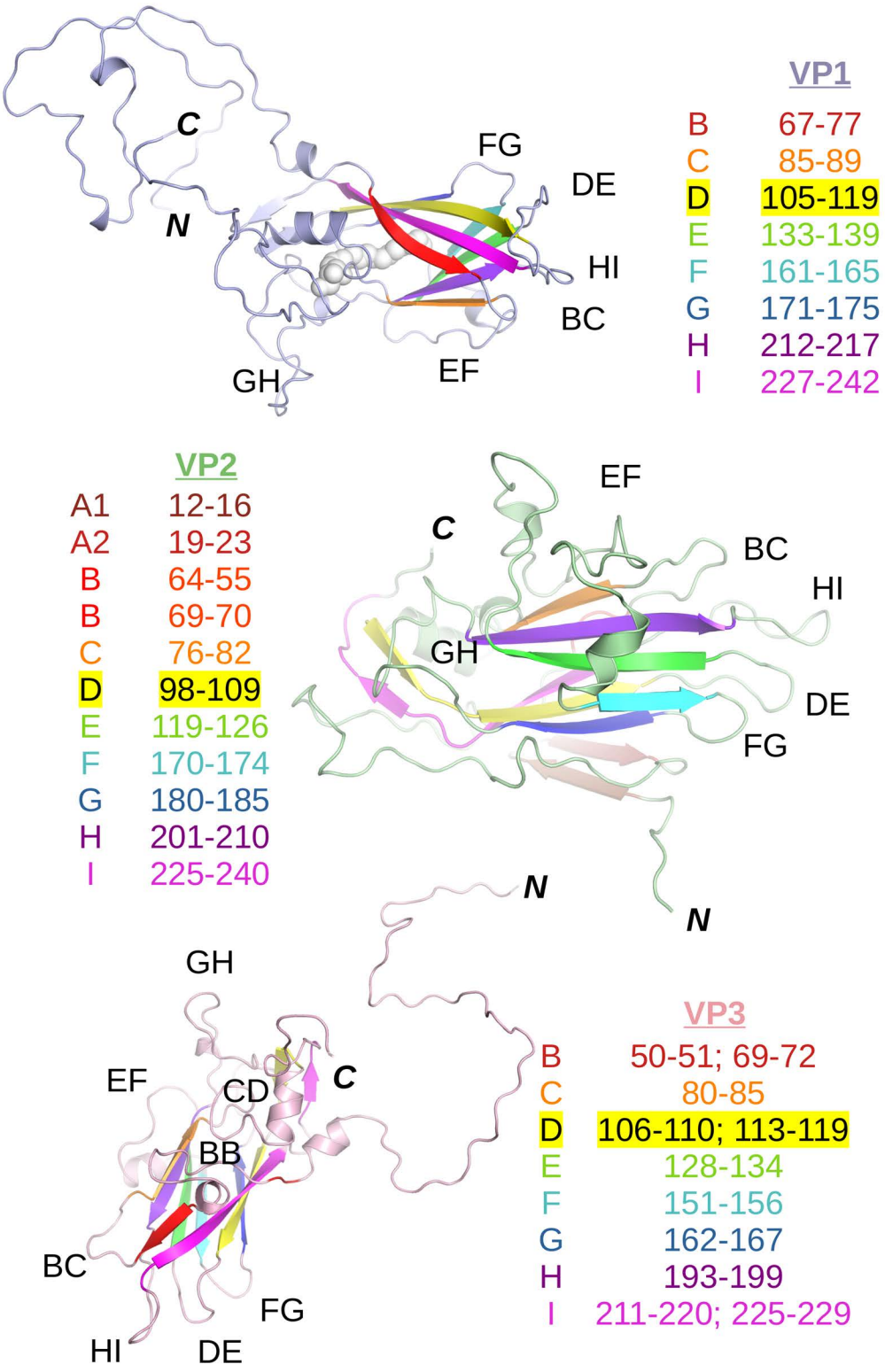


Figure 3.3: Top to bottom: VP1 (blue), VP2 (pale green) and VP3 (pale pink). Strands are coloured according to the key adjacent to each VP, and associated sequence numbers (note that the 1000-3000 notation is omitted here for clarity) are listed. Loops are labelled in black type, while N and C termini are labelled in black, bold italic type. For VP1, the P-factor is shown as pale-grey spheres for reference.

4.1.2), locks the side chains of residues T1014, N1057 and N1017 in place. This cannot occur in EV-E1, where T1014 is replaced by a hydrophobic leucine residue.

At the capsid surface, the positioning of VP1's BC-loop is similar to that in EV-E1, except with a wider spacing between EV-F3's B and C strands, allowing optimal stacking interactions between B-strand F1079 (in place of P1084 in EV-E1) and Y1086 (a Thr in EV-E1) on the neighbouring C-strand, Figure 3.4. EV-E1's BC-loop is also more extended than that of EV-F3 and reaches towards the adjacent VP1 HI loop. EV-F3's DE loop, located on the uppermost lip of the 5-fold axis, sits further away from its adjacent protomer than EV-E1, presumably to maximise contact with the more extended VP1 HI loop, Figure 3.4.

A highly-conserved N-terminal nine residue motif (typically PxLxAxExG, where 'x' denotes any amino acid, also referred to as a 'hook' in literature) (Hovi and Roivainen, 1993; Airaksinen et al., 1999) that can be externalised in enteroviruses during uncoating and has been shown to be important for CV-A9 infectivity (Roivainen et al., 1993; Airaksinen, 2000; Airaksinen et al., 2001; Ren et al., 2013) (1031-PALQAAETG-1040 in EV-F3, here referred to as the 'PAL-motif') ,and rests below VP4 (Figure 3.5(b)), is almost identical between EV-F3 and EV-E1, Figure 3.5. However, there are additional intra-chain H-bonds in EV-F3 between Q1039 (Q1034) and N1043 of VP1, which may serve to stabilise the loop conformation, see Figure 3.5(a,c). The motif conformation differs somewhat for PV-1 (Figure 1.5(c)), but is very similar in EV-A71 and somewhat similar in EV-D68, Figure 3.4(d,e).

P-site

The P-site (see Chapter 1, Section 1.1.2), nestled within VP1, contains density associated with a longer pocket factor (modelled as stearic acid C18:0) than in EV-E1 and other enteroviruses (Figure 3.6). Longer fatty acids can be accommodated owing to the presence of residues V1173 and L1213 in EV-F3 which replace bulky residues F1180 and M1216 in EV-E1 (Figure 3.7(a) and Figure 3.7(b)).

Residues R1096, F1216 and Y1260 of EV-E1 were previously identified as instrumental in constricting the P-site entrance for EV-E1, compared to less bulky residues in PV-1 (Figure 3.7(a)), potentially conferring a thermostabilising effect by making P-factor release less favourable (Smyth et al., 1995). For EV-F3, we see Y1089, Y1209 and Y1254 occupying these positions. Y1089's phenol side chain is modelled angled away from the capsid surface relative to R1096, opening the pocket entrance (Figure 3.7(a), Figure 3.7(c)).

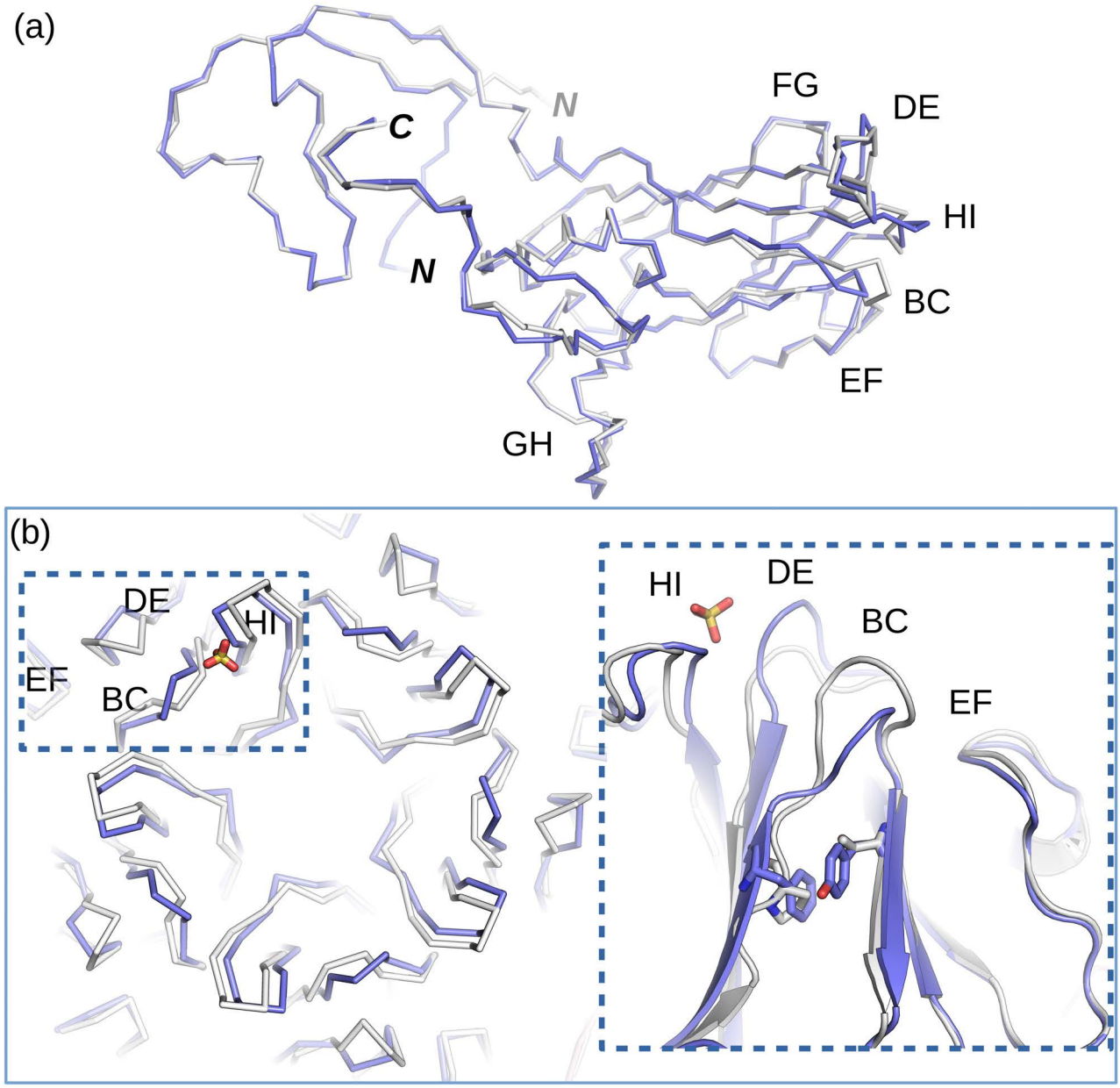


Figure 3.4: (a) Comparison of EV-F3 (blue) and EV-E1 (grey) VP1 aligned in ribbon representation The N and C termini are labelled in bold italics. The N-terminus, which differs for EV-E1, is shown in light grey italics. Surface loops are labelled by convention in black typeface. (b) Comparison at the five-fold icosahedral axis for EV-F3 and EV-E1 from the top (left) and from the side (right), with colouring as before. The sulphate ion modelled for EV-E1 (sticks, coloured by atom) and VP1 loop labels are only shown for one of the protomers for clarity. For the capsid slice, the pertinent $\pi - \pi$ stacking interactions between F1079 and Y1086 is shown in stick representation.

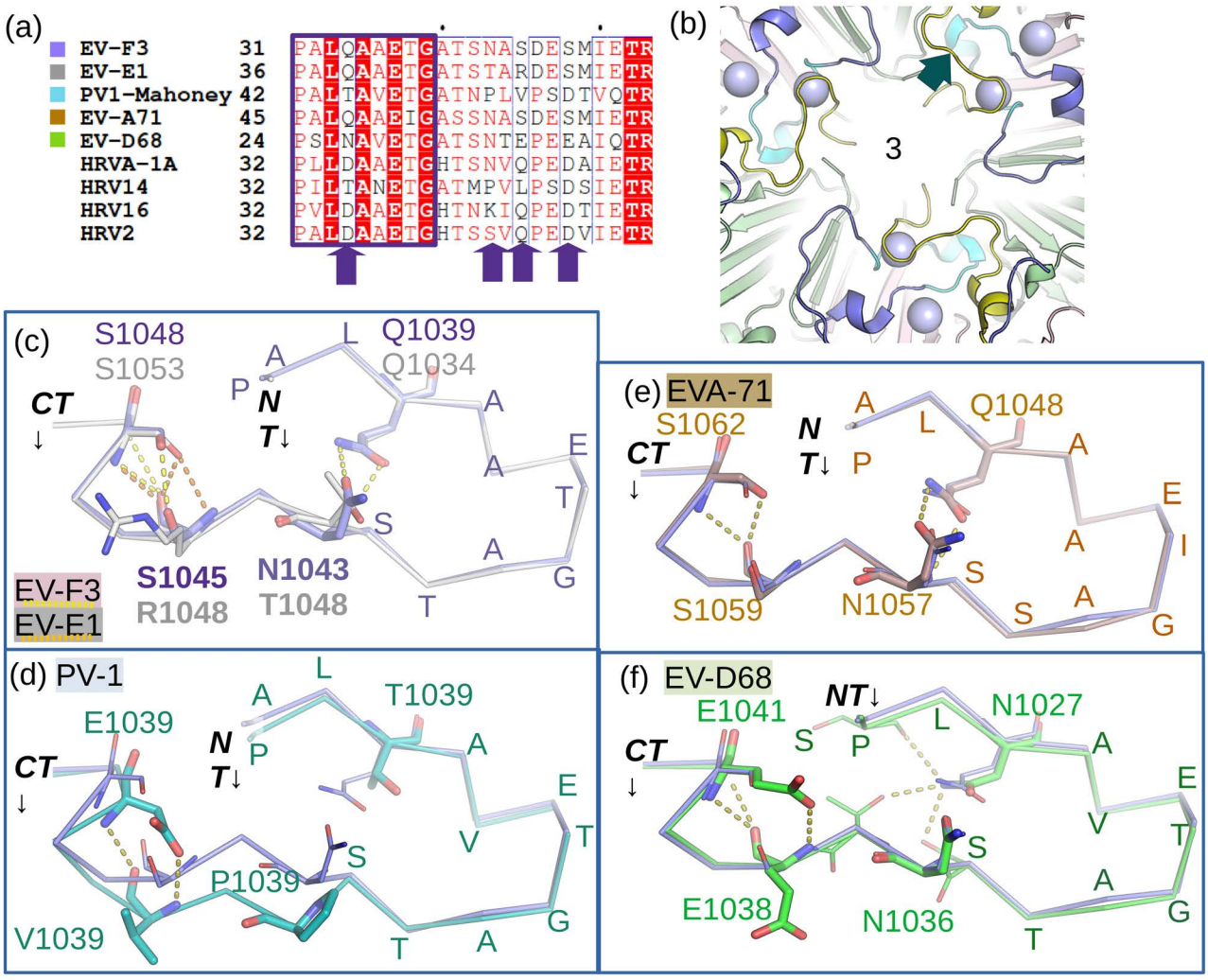


Figure 3.5: (a) Sequence alignment of enteroviruses, as listed, around the highly conserved nine-residue motif towards the VP1 N-terminus.The coloured boxes for the first five viruses correspond to the colours used for structural comparison below. The purple arrows highlight four interesting residues that differ between EV-F3 and the panel of enteroviruses listed. (b) PAL site in the context of the capsid inner surface in cartoon representation. K^+ ions are shown as lilac spheres. The nine-residue PAL site is coloured cyan, and the PAL site for one protomer is indicated with a deep teal arrow. (c) Comparison of EV-F3 (blue) and EV-E1 (grey), with residue labels for the highly conserved nine-residue motif for EV-E1 shown as single letters in blue. Pertinent residues are shown in stick representation, with thinner sticks for EV-E1 and their corresponding number is shown. Notable H-bonding interactions for EV-F3 are shown in yellow, and for EV-E1 in orange. (d) Comparison of the same region as before with PV-1 Mahoney (PDB: 1HXS) (Miller et al., 2001) with EV-F3 in blue and PV-1 in turquoise. (e) As before, except with EV-A71 shown in light brown. (f) As before, but with EV-D68 in pale green.

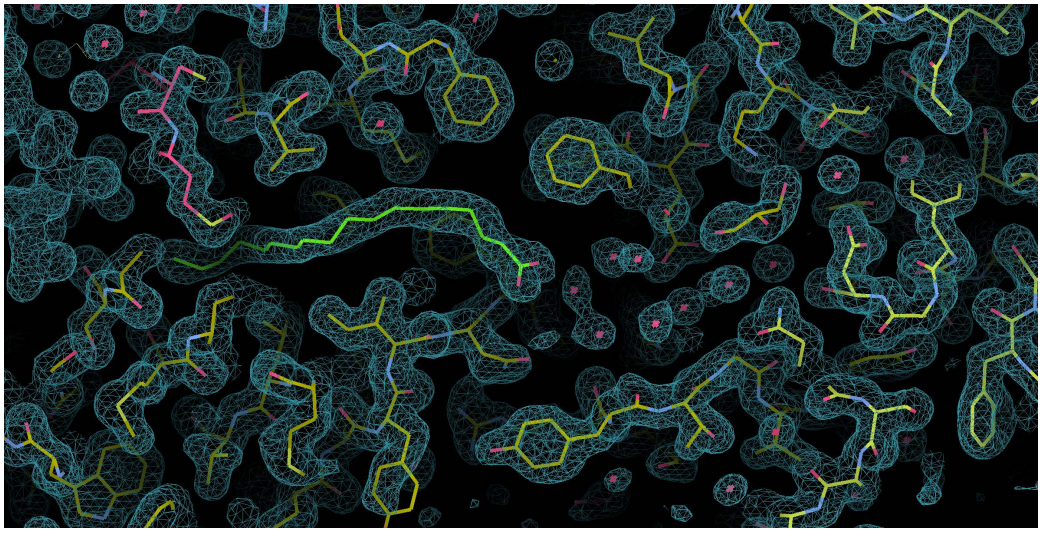


Figure 3.6: 5-fold NCS averaged density with a protomer of our EV-F3 model showing stearic acid (green sticks). VP1 is shown in yellow sticks and VP3 in magenta.

VP2

Substantial differences between EV-F3 and EV-E1 occur at the N-terminal portion of VP2 (Figure 3.8). EV-E1's VP2 N-terminus seems to have been mis-assigned, since it clashes with EV-F3 VP4 residues 4064-4069, which agree with other enterovirus VP4 structures, such as PV-1 (PDB: 1HXS, Miller et al. (2001)), Figure 3.8(b). The position of the first 10 residues of PV-1 Mahoney VP2 differ from EV-F3 (Figure 3.8(b)), reflecting the structural variability on the inside surface of picornavirus capsids.

The majority of surface loops appear similar for EV-F3 and EV-E1 (Figure 3.8(a,c,d)), except the antigenic EF loop, and the HI loop near the three-fold axis. The EV-F3 HI loop rests closer to the three-fold axis channel than in EV-E1 (2217-2224 EV-E1, 2214-2222 EV-F3), perhaps due to relaxation following a change from a rigid Pro to Thr at position 2219. At the tip of the EF loop, a subtle change in conformation makes EV-F3 flatter in shape (Figure 3.8(d)). Residue differences within the modelled EF loop, Figure 3.9, where typically conserved hydrophobic residues on the C-strand (Trp) and EF loop region (Phe) for EV-E1, PV1, EV-A71 and EV-D68 are replaced by H-bond capable residues His and Tyr, respectively, for EV-F3, moving the EF loop towards the C-strand in VP2.

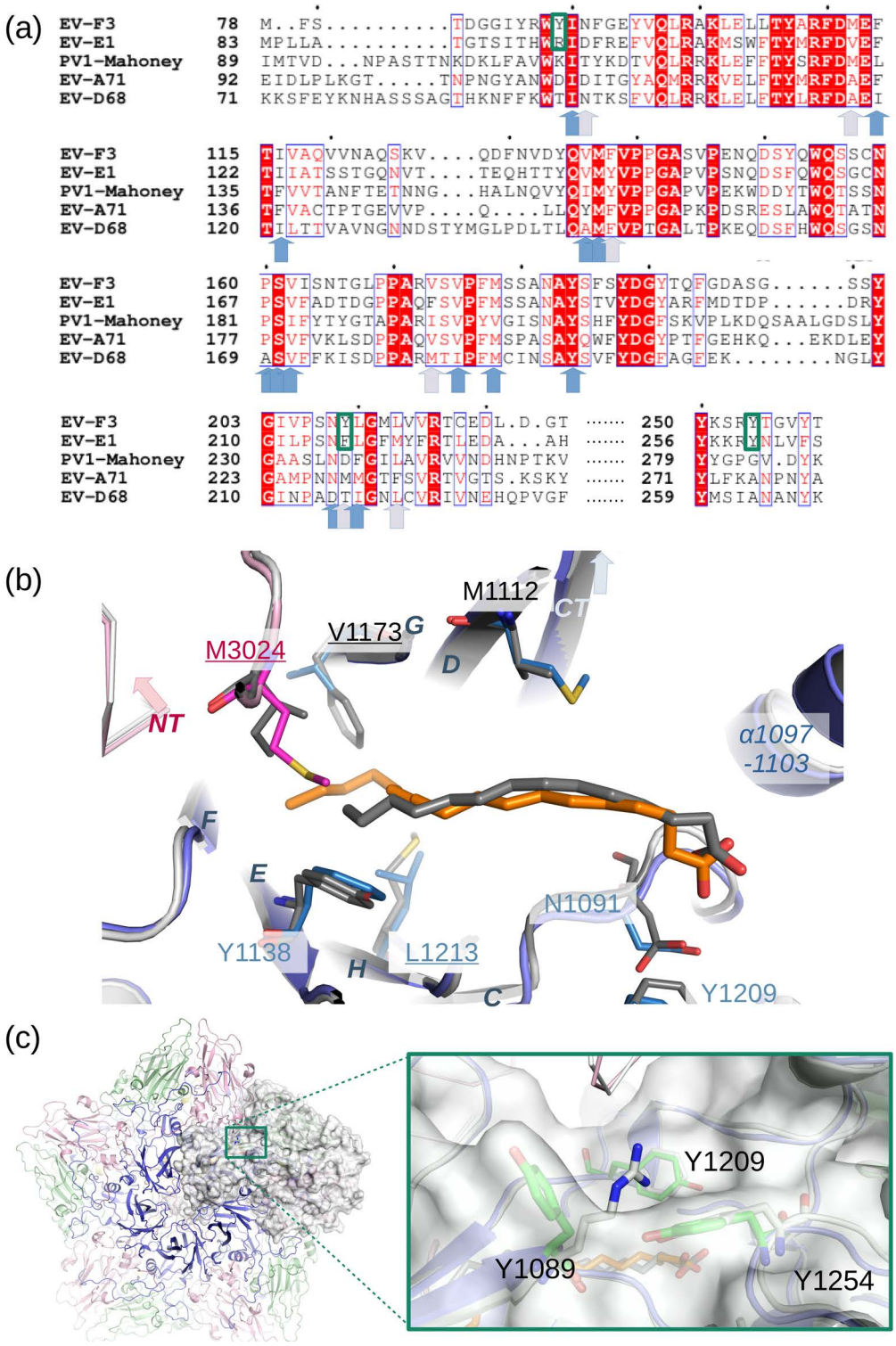


Figure 3.7: (a) Sequence alignment for a small selection of enteroviruses. Residues interacting with the P-site are indicated with blue (conserved between EV-E1 and EV-F3) or pale blue arrows (not-conserved). (b) P-site for EV-F3 (coloured blue for VP1, magenta VP3, and orange for the stearic acid P-factor). Strands are labelled with single letters, and C and N termini as 'CT' and 'NT', respectively. Key residues are shown in stick representation, with pertinent mutations underlined. (c) Pentamer with surface for one protomer showing the location of the P-site entrance. This is shown enlarged in the rightmost panel, with pertinent EV-F3 residues labelled and shown in green, EV-E1 in grey.

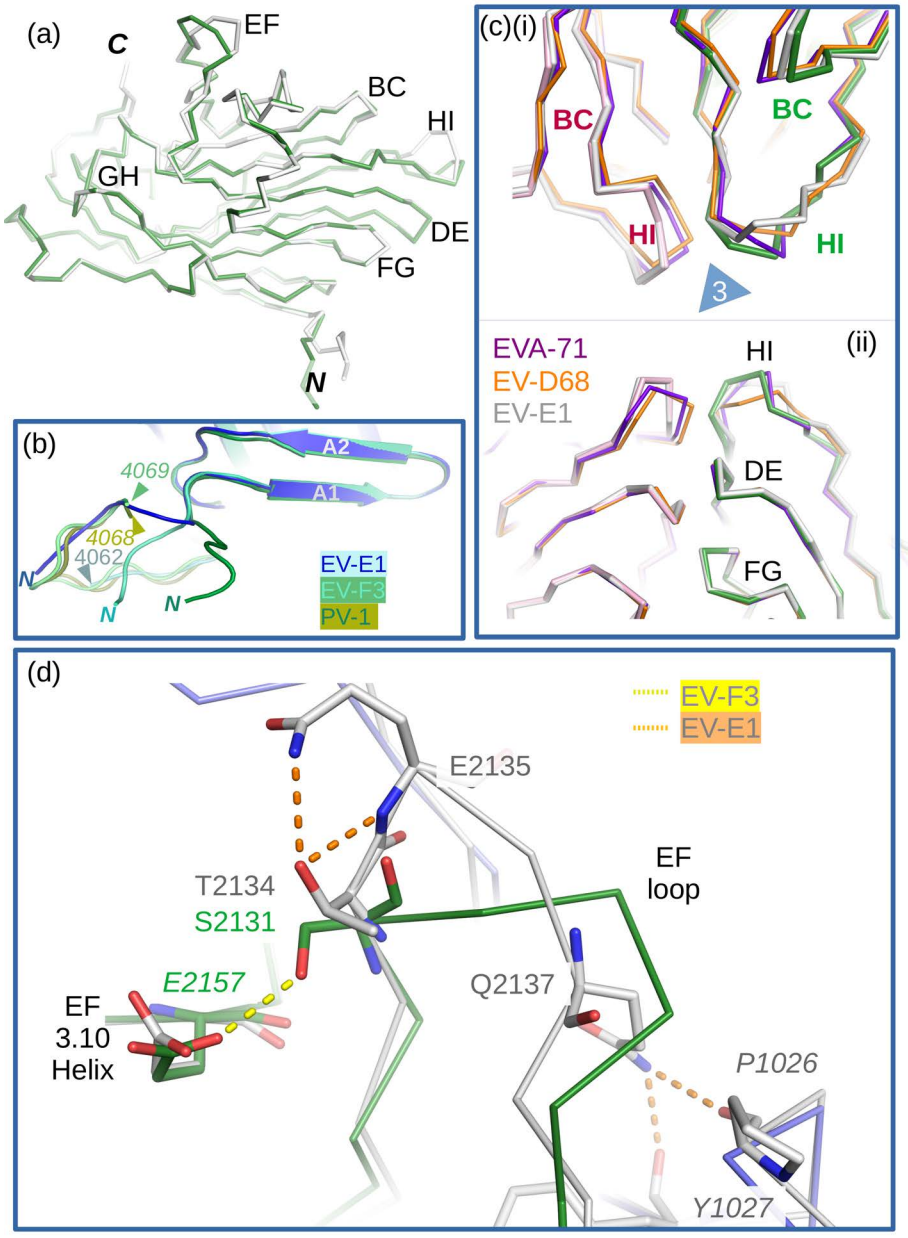


Figure 3.8: (a) Ribbon overview of VP2 for EV-F3 (green) overlaid with the model for EV-E1 (grey). Loops are labelled in black text, C and N termini in bold italic typeface. (b) Cartoon representation of the N terminal portion of VP2, with EV-E1 in blue, EV-F3 in pale blue and PV-1 in dark green. The VP4 is shown in sky blue for EV-E1, aquamarine for EV-F3 and dark gold for PV-1, as indicated by the key (text colour pertains to strand colouring and highlight colour corresponds to VP4 colouring). VP2 A1 and A2 strands are marked. (c)(i,ii) View of a protomer at the three-fold axis viewed from (i) the top of the capsid surface and (ii) through the three-fold channel, with loops labelled. EV-A71 is in purple, EV-D68 in orange and EV-E1 in grey. EV-F3 VP1 is in pale pink and EV-F3 VP2 is in dark green. (d) View of the VP2 EF loop for EV-F3 (green) and EV-E1 (grey). Pertinent H-bonding interactions are shown in orange for EV-E1, and in yellow for EV-F3. EV-E1's EF loop conformation is partly influenced by an intra-loop contact between T2134 and E2135. For EV-F3, S2131 replaces Thr, instead contacting residue E2157 of the adjacent EF 3.10 helix. In addition, EV-E1's loop may be narrower to fulfil a stabilising contact between Q2137 of the EF loop and the main chain of VP1 N-terminal residues P1026 and Y1027. No such contacts were observed for EV-F3.

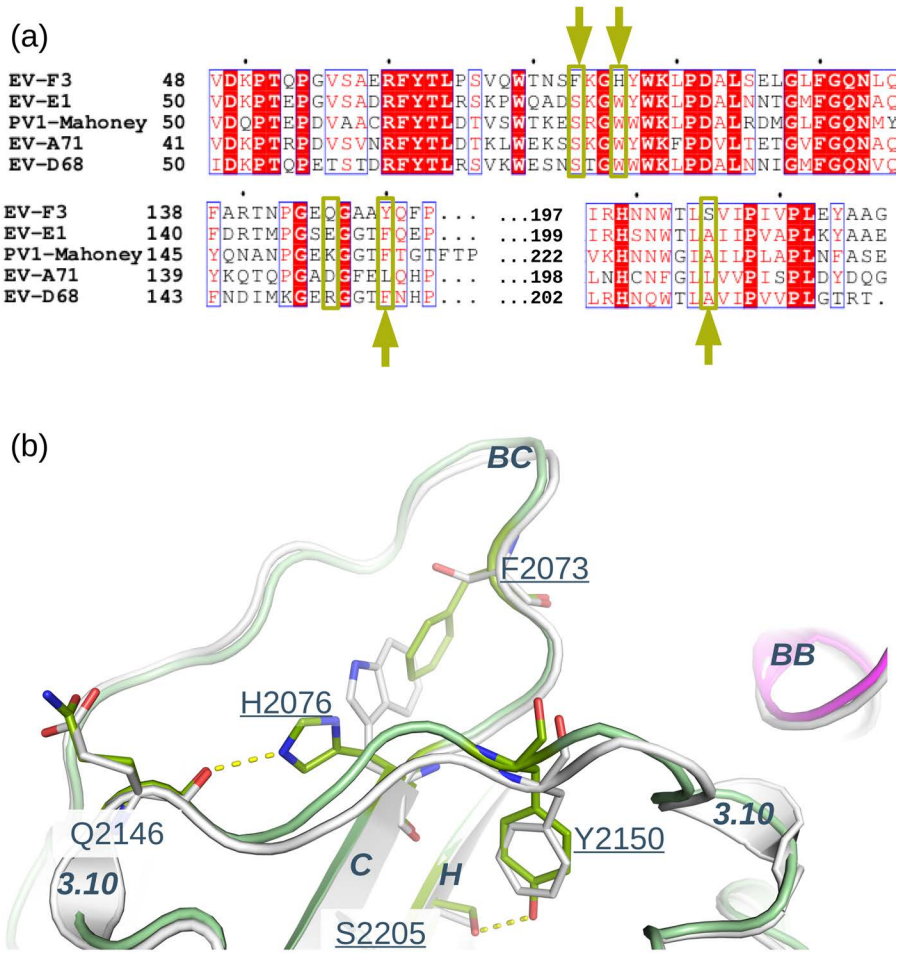


Figure 3.9: (a) Sequence alignment for a small selection of enteroviruses for the VP2 regions displayed in (b). Notable residue types that differ for EV-F3 are indicated with green arrows, and correspond to underlined residue numbers shown in (b). (b) Differences in EF loop region between EV-E1 (grey) and EV-F3 (green, VP3 BB knob in magenta). Loop, strand and helix labels are in bold face, italic type. Two interactions, with Y2150 of the EF loop region and H2076 of the neighbouring C strand, draw the C-terminal end of the EF loop towards the adjacent C-strand. C-strand H2076 is unique to EV-F3; for EV-E1, PV-1, EV-A71 and EV-D68, we instead see a Trp residue (W2078 for EV-E1), and its side chain sits rotated in the opposite direction to H2076's indole ring, instead occupying the space filled by residue F2073 in EV-F3. H2076 contacts the main chain carboxyl group of Q2146 of the neighbouring EF loop, securing a secondary interaction to shift the EF loop slightly closer to the C and H strands. Y2150 draws the loop towards Ser of the nearby H strand to satisfy a H-bonding contact with Ser's hydroxyl group, an interaction that cannot occur for EV-E1 owing to F2152 and A2207 substitutions in place of Tyr and Ser.

VP3

EV-F3's VP3 N-terminal portion and the majority of surface loops agree well with those of EV-E1 (Figure 3.10(a)), except the GH-loop and D-strand.

In both EV-E1 and EV-F3, the GH-loop is extended compared to PV-1 and EV-A71, filling the canyon base, Figure 3.11(a,b)). In EV-E1, the loop is more twisted and slightly shorter (EV-F3 has an insertion at residue 3181, Figure 3.10(b)), Figure 3.11(c,d)(i). Both EV-E1 and EV-F3 GH loops include an extensive intra-loop hydrogen bonding network, largely mediated by residue 3181 (E and D for EV-F3 and EV-E1, respectively, Figure 3.11(c,d)(ii-iii)). The C-terminal portion of the GH loop of EV-F3, but not EV-E1, also makes multiple contacts with strand D (Figure 3.12(a,b)).

VP4

VP4 snakes along the capsid interior. The modelled VP4 chain is incomplete and the density is generally not as strong as other chains, with relatively high average B-factors (31 $Å^2$, cf 18, 15, 15 $Å^2$ for VP1-3). There are marked differences compared to EV-E1 (Figure 3.13(a)). Residues 4032-4059 are similar to EV-E1, but EV-E1's VP4 C-terminus is closer to the bulk capsid, possibly due to mis-assignment of residues 2005-2011, as described in the 'VP2' section above.

No ordered RNA can be seen in the capsid interior except for the cryo-cooled crystallographic structures of EV-F3-GSH and EV-F3-sialyllactose presented in Chapter 4 (Sections 4.0.3, 4.0.5), where density consistent with a guanosine ring contacts the N-terminal extension of VP2 at residue W2036 via a π-stacking interaction with W2036's indole ring (Figure 3.13(b)), although there is no density for the associated phosphate. Interestingly, this is consistent with the guanosine modelled previously near this site for CV-A21 (PDB: 1Z7S) (Xiao et al., 2005a) and observed in some HRVs, PV-3, SVDV and CV-B3 (Arnold and Rossmann, 1990; Filman et al., 1989; Fry et al., 2003; Muckelbauer et al., 1995; Hadfield et al., 1997).

Antigenic Loops

We see a number of differences at the antigenic loops, previously identified as neutralising EV-E1 epitopes (Smyth et al., 1990, 1995; Smyth and Martin, 2001), labelled I-III for the purposes of this discussion; (I) VP1 GH-loop: Y1191-D1196 (D1198-Y1203), (II) VP2 EF-loop: Y2150- G2159 (F2152-G2161); (III) VP3 BC-loop: E3054-G3063 (E3054-G3063), see Table 3.3, Figure 3.14.

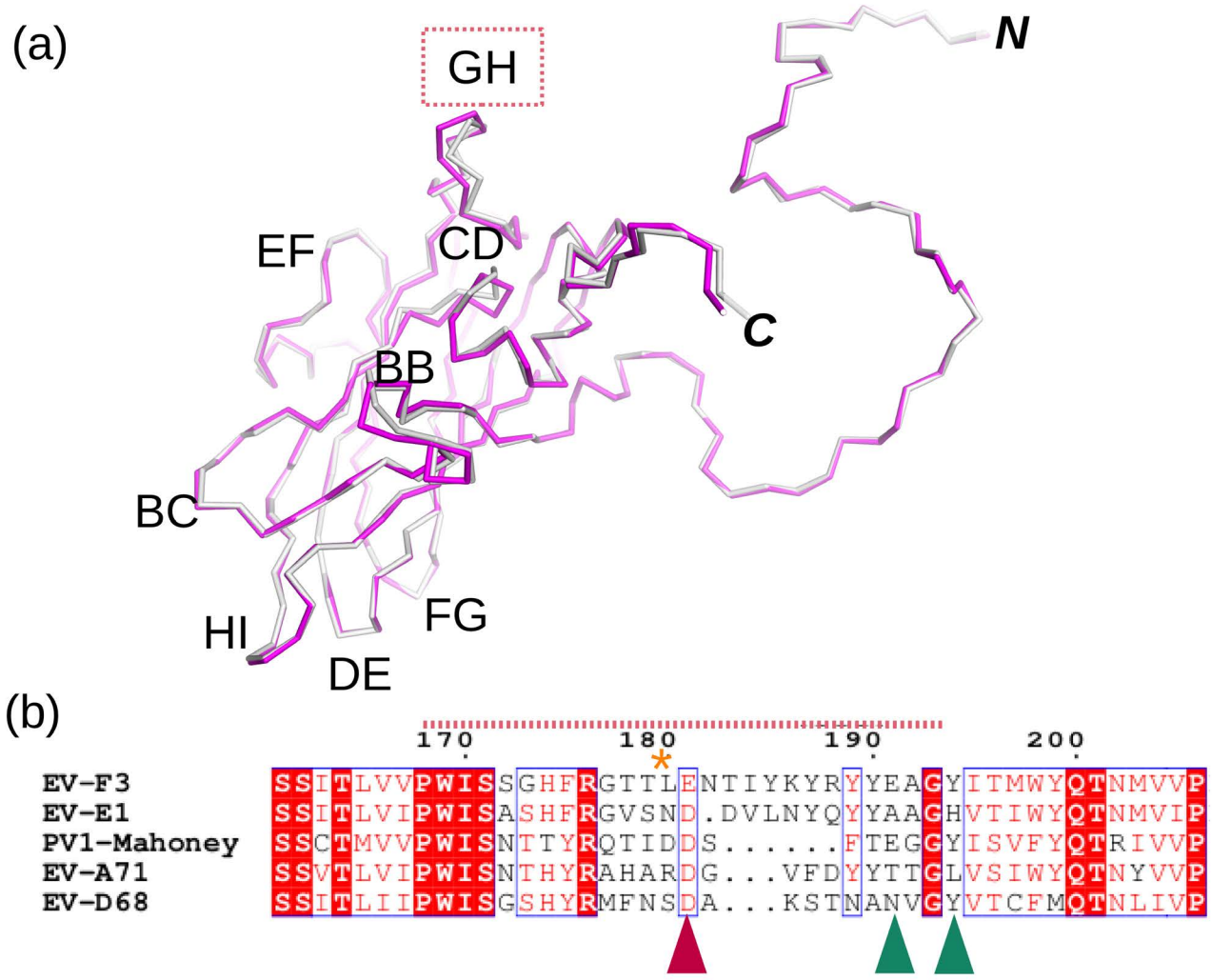


Figure 3.10: (a) Overlay of VP3 models for EV-F3 (light pink) and EV-E1 (grey). (b) Sequence alignment for a small selection of enteroviruses, with residues line the P-site indicated with arrows. The GH-loop is shown with a magenta dashed line above the sequence and the orange star indicates the position of an insertion for EV-F3 relative to EV-E1. The keystone residue that makes many of the intra-GH-loop contacts is indicated by the magenta arrow and green arrows show major contact residues.

Table 3.3: Neutralising antigenic sites for EV-E1, and corresponding sequences for EV-F3 from (Smyth et al., 1990, 1995; Smyth and Martin, 2001)

Antigenic site	Capsid protein	Residue numbers	Strain	Sequence
1	VP1	1198-1203	EV-E1	Y A R F M D
		1191-1196	EV-F3	Y T Q F G D
2	VP2	2152-2161	EV-E1	F Q E P F W L E D G
		2150- 2159	EV-F3	Y Q F P F T F E D G
3	VP3	3054-3063	EV-E1	E A N N R E G V E G
		3054-3063	EV-F3	E I N N V E S A E G

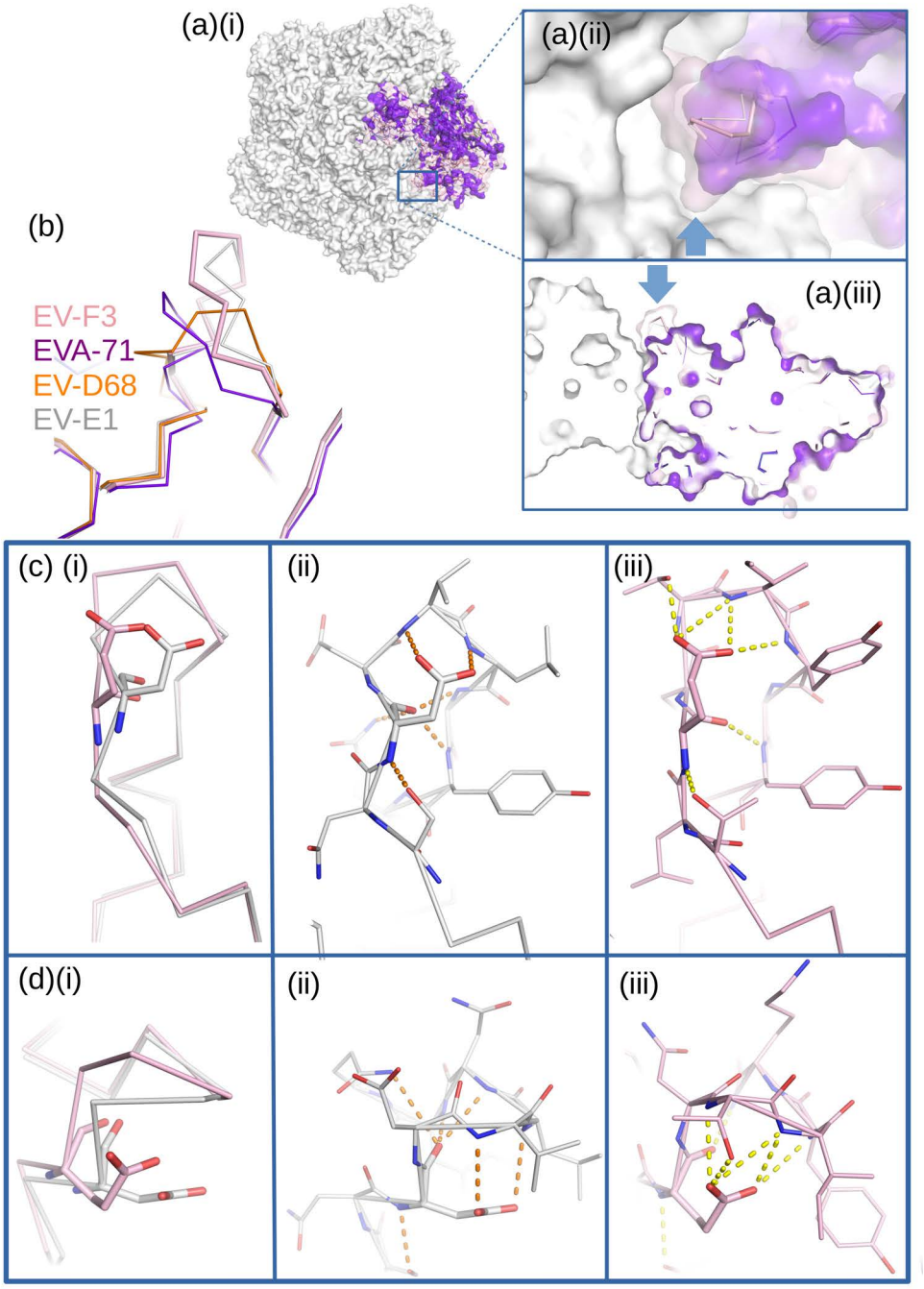


Figure 3.11: Overview of EV-F3's VP3 GH loop. (a)(i) Pentamer with grey surface for EV-F3, and a protomer highlighted in pink. EV-A71 (PDB: 3VBS, Wang et al. (2012)) is overlaid with this EV-F3 protomer and shown as a purple surface . (a)(ii) Zoomed in portion of (a)(i), around the VP3 GH loop. The extended EV-F3 surface, relative to EV-A71, is indicated with a blue arrow. (a)(iii) Section through the pentamer shown in (a)(i). The position of the VP3 GH loop is indicated with a blue arrow. (b) Side view of VP3 GH loop with EV-D68 (orange), EV-E1 (grey), EV-A71 (purple) and EV-F3 (pale pink and broader ribbon) overlaid. (c)(i-iii) side view of VP3 GH loop with EV-F3 (pink) and EV-E1 (grey). (i) EV-F3 and EV-E1 overlaid, with key E3181 (or D3181 for EV-E1) shown in stick representation. (ii) and (iii) show key intra-loop contacts between residue 3181 and other loop residues for EV-E1 and EV-F3, respectively. (d)(i-iii) As for (c), except looking down onto the outer capsid surface of VP3 GH loop with EV-F3 (pink) and EV-E1 (grey).

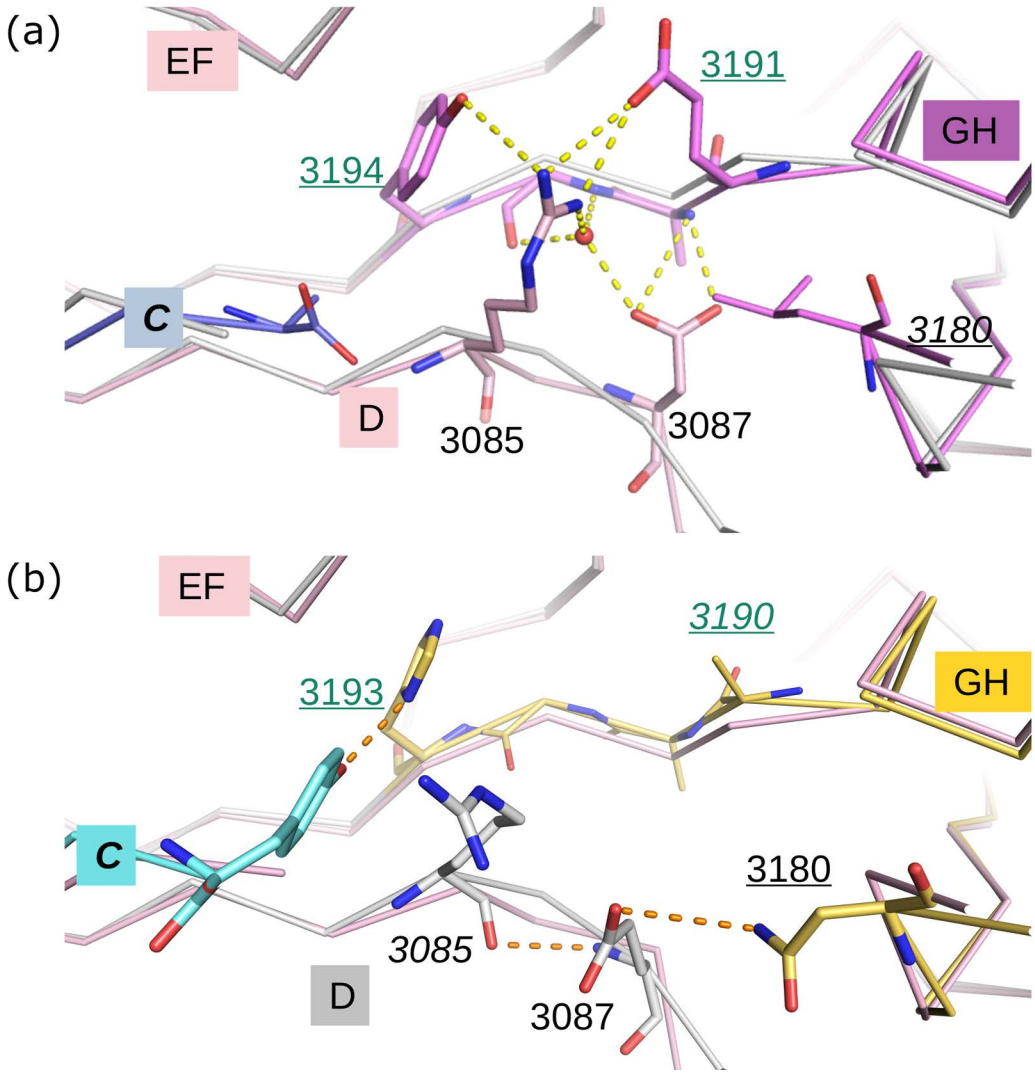


Figure 3.12: Contrasting interactions between EV-E1 and EV-F3's D strands and VP3 GH loops. (a) EV-F3 in ribbon representation, coloured by structural protein (VP1 blue, VP3 magenta), with key residues shown as sticks and H-bonds as yellow dashes. EV-E1 is shown in white ribbon representation for contrast. (b) As for (a), except EV-E1 is shown in grey, yellow and cyan ribbons and pertinent residues in stick representation. The VP1 N terminus is shown in cyan, with the N terminal residue, Y1281 shown as sticks. H-bonding interactions are displayed as orange lines. Residue positions deemed pertinent for EV-F3 interactions in this region are shown as thin sticks. For both (a, b), the respect VP1 C termini are marked in bold italic type, with the C-terminal residue A1274 shown as sticks. For EV-F3, we see residues 3085 and 3087 of the D strand form extended H-bonding interactions with residues Y3194, S3193 and E3191 that form the C-terminal portion of the the adjacent VP3 GH loop. These interactions are partly mediated by a bridging water molecule (red sphere). Residue L3180 is shown as thinner sticks for reference.

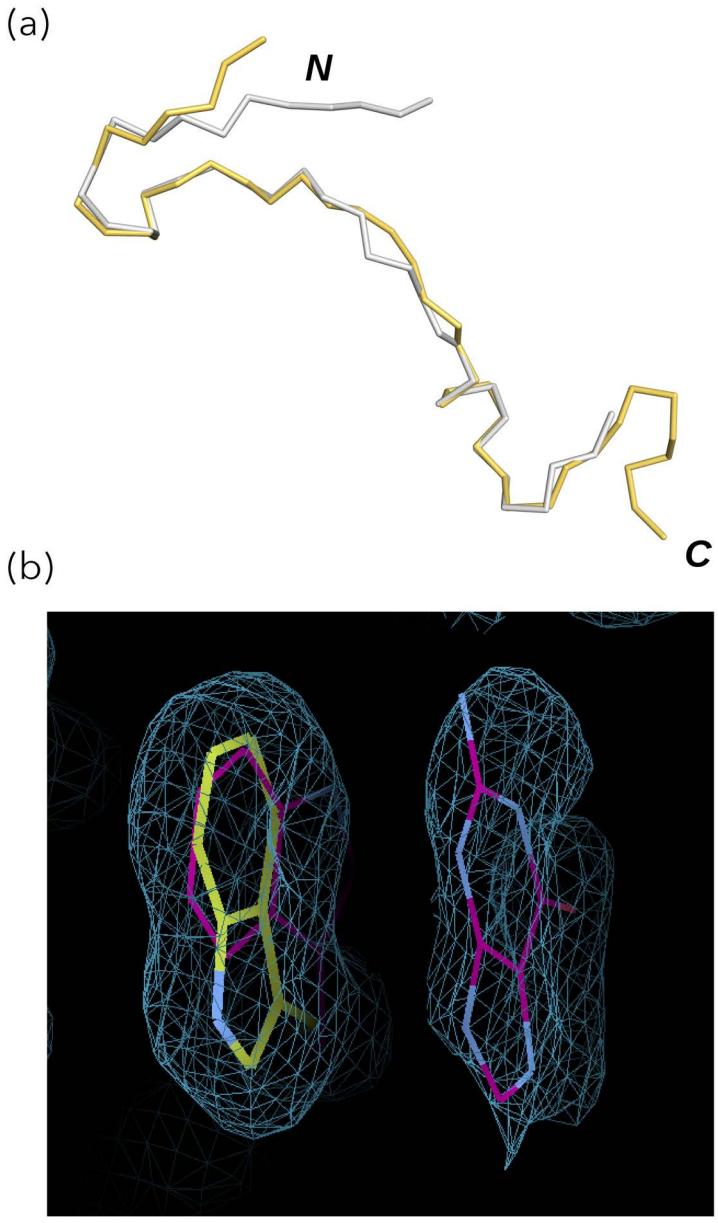


Figure 3.13: (a) Ribbon representation of EV-F3 VP4 (yellow) and EV-E1 (grey), with N and C termini labelled in bold italics. (b) NCS averaged map showing density for W2036 (yellow sticks) and additional density that sits in the same position as a previously modelled guanosine for CV-A21 (PDB: 1Z7S) (Xiao et al., 2005a) is shown in purple, thin sticks.

(I) The VP1-GH loop conformation is similar in both serotypes, with some deviation in residues following those defined as antigenic (1197- 1200). There is considerable charge variation, and the reduced side chain bulk due to the R1193Q and M1195G substitutions results in a flatter loop for EV-F3.

(II) The antigenic residues in the EF loop of VP2 form the least accessible end of this extended surface loop which is, for the most part, structurally conserved, except between residues 2132 and 2136, where the EV-F3 loop packs closer to residue 1199 in the VP1 GH loop (see above).

(III) The antigenic residues reported for the VP3 BC loop are accessible and there is a small structural deviation at residue 3060. In this region, the VP1 N-terminal loop rises 0.5 Å higher from the capsid surface for EV-E1, to make optimal contact with the adjacent 3.10 helical residues 3061-3067 of VP3, and to avoid steric clashes with VP2 residue W2157, Figure 3.15.

These differences collectively account for the lack of cross reactivity of antisera raised against EV-E1 with EV-F3 (Smyth et al., 1992).

The Icosahedral 5-Fold Axis Channel

Many picornaviruses contain a largely anionic channel at the 5-fold axis, comprising mostly VP1 strands from adjacent protomers, with VP3 N-terminal portions sitting at its base. Usually the VP3 portions from five surrounding protomers form a β-tube, part of a so-called 'plug' that rests atop the five surrounding myristate linked (VP4myr) VP4 amino termini and VP1.

For EV-E1, Y3011 residues from encircling protomers are oriented such that the phenol hydroxyl groups form a hydrogen-bonding network across the 5-fold axis channel, effectively forming a plug (shown as white sticks in Figure 3.16(a,b)). Contrastingly, in EV-F3, Y3006 side chains 'rearrange', relative to EV-E1, widening the channel to 8 Å diameter (cf 5 Å for EV-E1), Figure 3.16.

In the XFEL EV-F3 structure, electron density extending through this Y3006 ring was modelled as an octyl-sulfate (OSF) molecule, with OSF's sulphoxide contacting the T3004 hydroxyl groups at the channel base (Figure 3.16). Given the proximity of this density to the N-terminus of VP4, it is conceivable that the density corresponds to a myristoyl moiety from a VP4 molecule, caught amongst the faces of hydrophobic VP3 Y3006 residues at the 5-fold axis. Although our plug of density, shown in Figure 3.16)(c) fails to support a chain as long as a myristoyl group (C14:0), this may be due to disordering of the remaining carbon chain. Typically, this VP4myr moiety has been

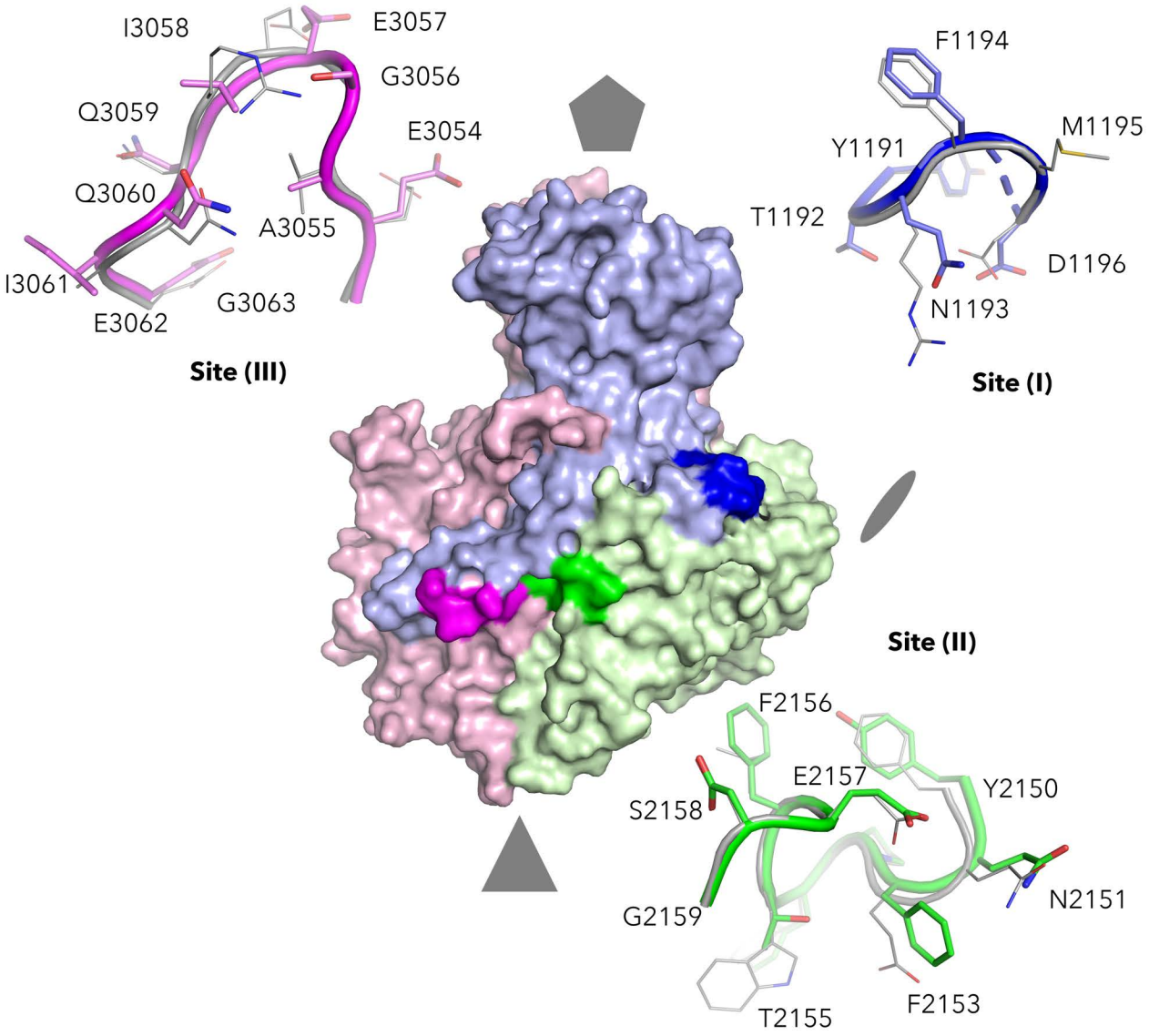


Figure 3.14: Three antigenic sites (I-III, coloured blue, green and magenta, respectively) previously identified for EV-E1 mapped onto a protomer (centre) with surface colouring by structural protein; pale blue VP1, pale green VP2, pale pink VP3. Zoomed in views of each site surround the protomer, and show the composite amino acids for EV-F3 in colour. EV-E1 is displayed in grey, thin sticks for comparison.

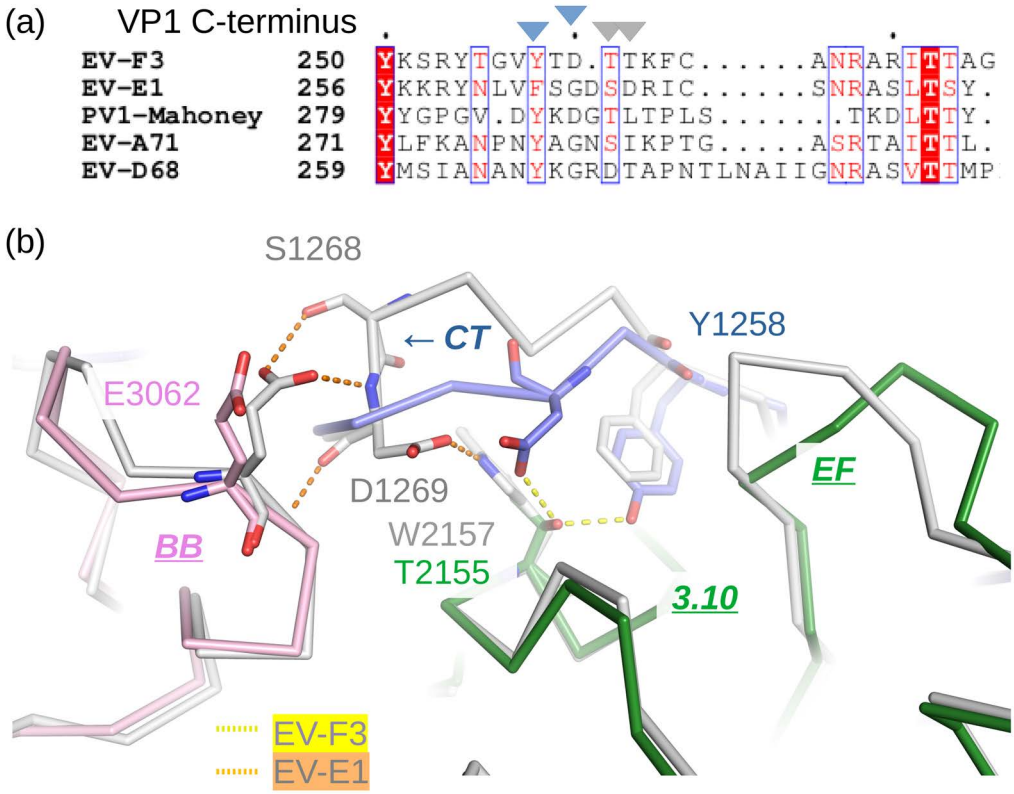


Figure 3.15: Interactions between the VP1 C-terminus and neighbouring loops. (a) Sequence alignment for the VP1 C-terminus. Residues that may form significant contacts leading to the differing C-terminal conformation between EV-E1 and EV-F3 are indicated with grey and blue arrows, respectively. (b) Overview of contacts between the VP1 C-terminus of EV-F3 (slate) and EV-E1 (grey). EV-F3's VP3 is shown in light pink and VP2 in green. The C-terminal portion of EV-F3 sits closer to the EF loop, in part due to an insertion in EV-E1 at residue 1267, satisfying a H-bond contact with the EF loop region and 3.10 helix residue T2155, shown as yellow dashes. This contact is not present in EV-E1, where a Phe and Trp replace the Tyr and Thr residues of EV-F3. EV-E1 instead contacts the VP3 BB loop E residue (orange dashes) via its S1268 and D1269 residues.

modelled, for instance, for HRV16, PDB: 1AYM (Oliveira et al., 1993) within the capsid interior, in the vicinity of the 5-fold pore, but blocked from entry into the 5-fold axis channel by the 5-fold plug (Figure 3.16).

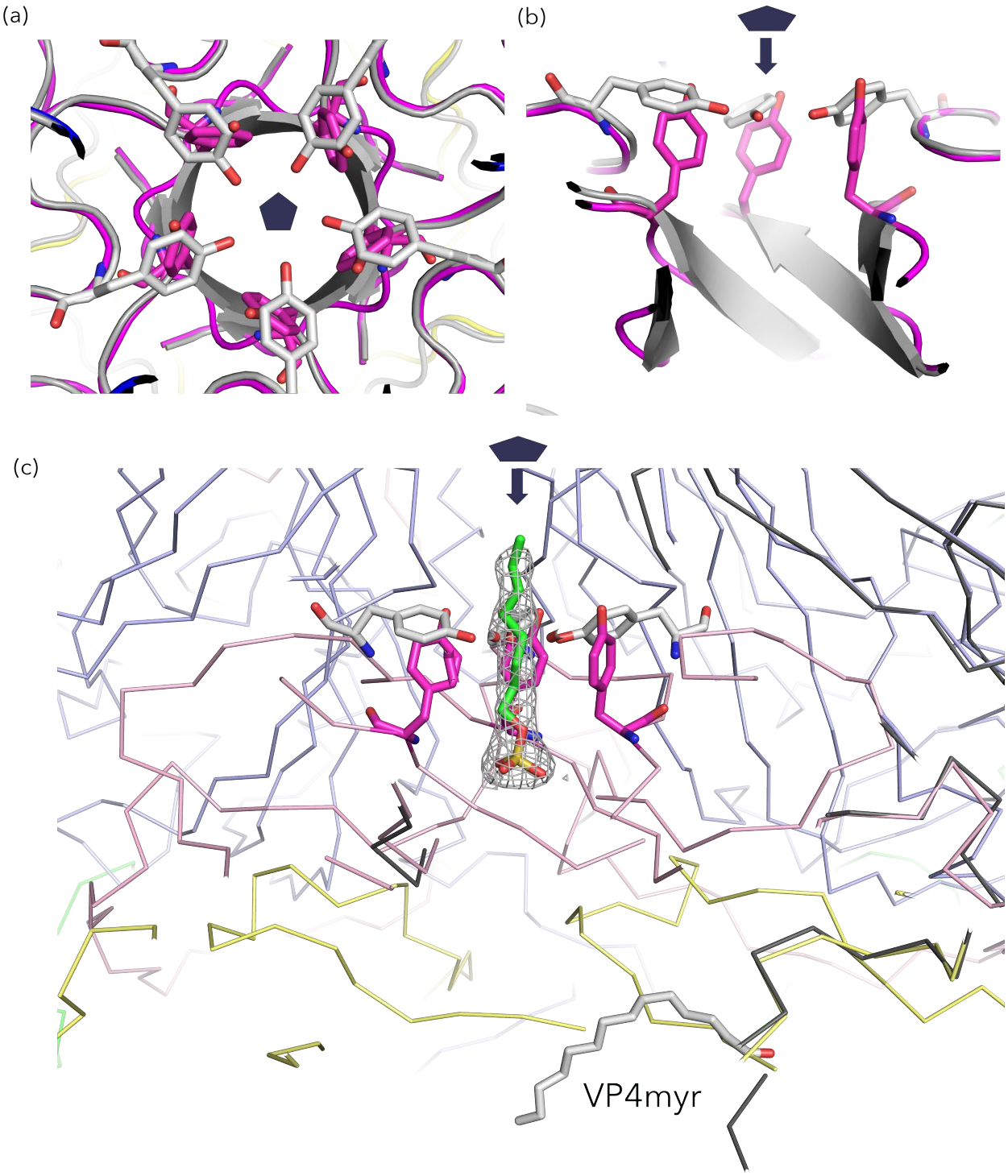


Figure 3.16: (a) and (b) Comparison of 5-fold Tyr residues for EV-F3 in cartoon representation, with VPs coloured by convention (Y3006, magenta sticks), and EV-E1 in grey cartoon (Y3011 as light grey sticks) from a central and a side cross-section of five-fold axis. (c) Cross section (side) of the 5-fold symmetry axis base for EV-F3 (ribbon representation, coloured by VP), with Y3006 of EV-F3 and just Y3011 of EV-E1 shown as magenta and white sticks, respectively. Our modelled OSF molecule is shown as green sticks, with associated 2Fo-Fc density shown in dark grey mesh. The VP4 and myristic acid modelled for HRV-16 (labelled VP4myr), PDB 1AYM, is shown in grey ribbon, and VP4myr as grey sticks (Oliveira et al., 1993)[115]

Table 3.4: Residue number from which the N-terminus for each structural protein was modelled for our room temperature (RTP) EV-F3 structure and that of the RTP empty EV-F3 structure. Note that values marked '*' were taken as effective N termini, compared to the mature virion (for empty capsids, VP2 and VP4 are combined to give VP0). A comparison with the cryo-cooled structures is shown in Table 3.2 above.

VP	N-Terminal Start	
	Mature	Empty
1	5	10
2	5	9*
3	5	1
4	27	30*

Empty EV-F3

Empty EV-F3 sample was provided by David Rowlands (Asbury Centre of Structural Molecular Biology, University of Leeds), who confirmed the presence of uncleaved VP0, consistent with particles being 'natural' empties. For the purposes of this discussion, VP0 residue numbers will be preceded by a '4' or '2', so as to follow the typical capsid protein numbering system and for ease of comparison with the mature virus.

Our empty and mature EV-F3 structures are very similar (Figure 3.17). The empty particles are not expanded, and are essentially identical in morphology to their mature counterpart (RMSD 0.18 Å for 743 aligned residues). VP1 is truncated by five residues (Table 3.4), relative to mature EV-F3, but located in an identical position. This contrasts with PV-1 natural empties, where the N-terminal loop was shifted relative to the mature capsid, to avoid steric clashes during rearrangement of the VP4 loop (Basavappa et al., 1994).

In addition, the surface topology of mature and natural empty EV-F3 models appears indistinguishable (Figure 3.18(a,b)). Within the capsid, the P-site is present, and there is a bound pocket factor, Figure 3.18((c).

There are slight differences in the capsid interior, however; gaps may be seen in the empty particles at the VP1 N-terminus and around the 5-fold axis, where the N-terminal portion of VP4 is missing (Figure 3.18(b,c)). The modelled residues closest to the VP0 scissile bond are shown in Figure 3.18(a,b).

Notably, the cross-Picornaviridae conserved motif PHQ, believed to be involved in autocatalytic cleavage of the VP4-VP2 junction (Figure 3.19(a)), located at residues 2169-2171 of EV-F3's

VP0, is modelled far from the scissile bond at the VP4-VP2 junction of EV-F3's VP0 (20 Å from K4069 to H2170 C_α-C_α, Figure 3.19(b)). This contrasts the PV-1 natural empty capsid model (PDB: 1POV 3.19), where both residues either side of the scissile bond (Ser and Asn) have been modelled, and situated where the loop is relatively close to the putative catalytic PHQ motif, 21 Å away from EV-F3 empty's modelled K4069.

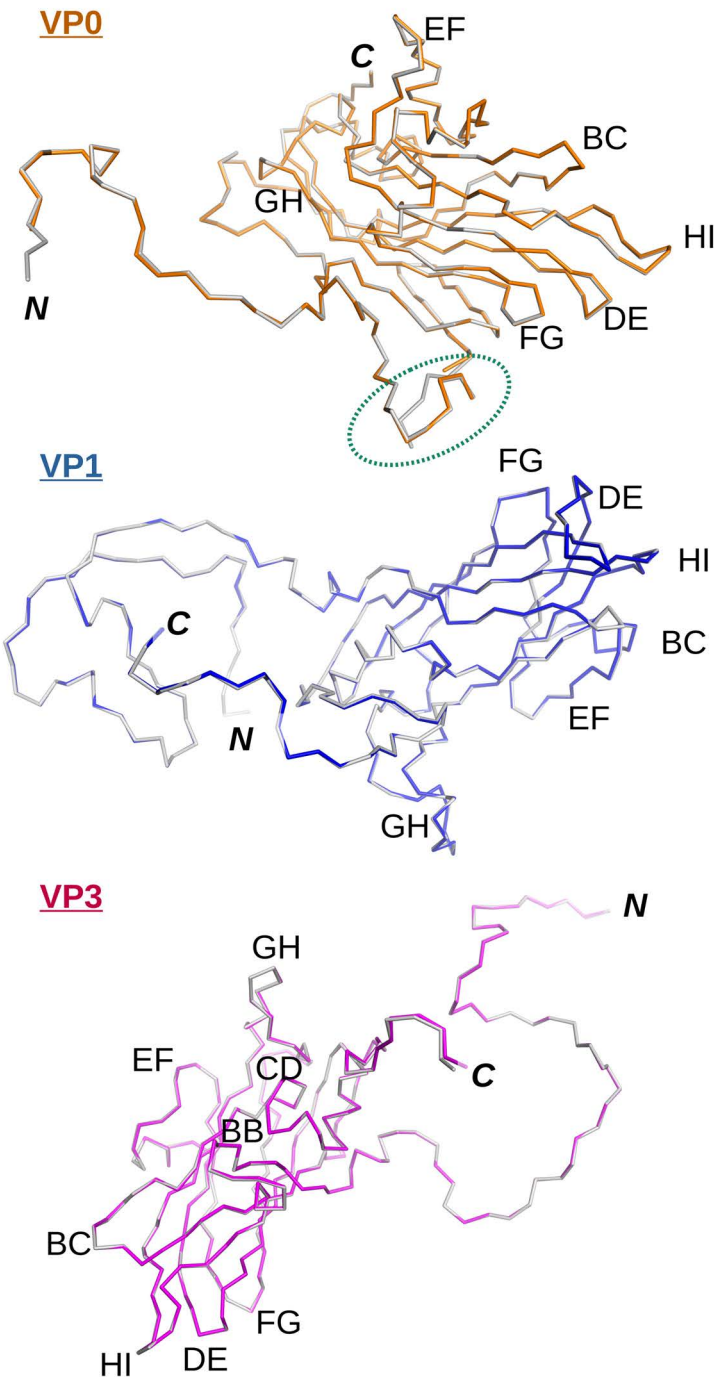


Figure 3.17: EV-F3 empty VPs (coloured orange for VP0, blue for VP1 and magenta for VP3) overlaid with EV-F3 mature (grey), in ribbon representation. Loops are marked in black text, N and C termini in bold italic typeface. EV-F3 mature's VP2 and VP4 are shown with EV-F3 empty's VP0, and the break point between these two portions is circled with green dashes.

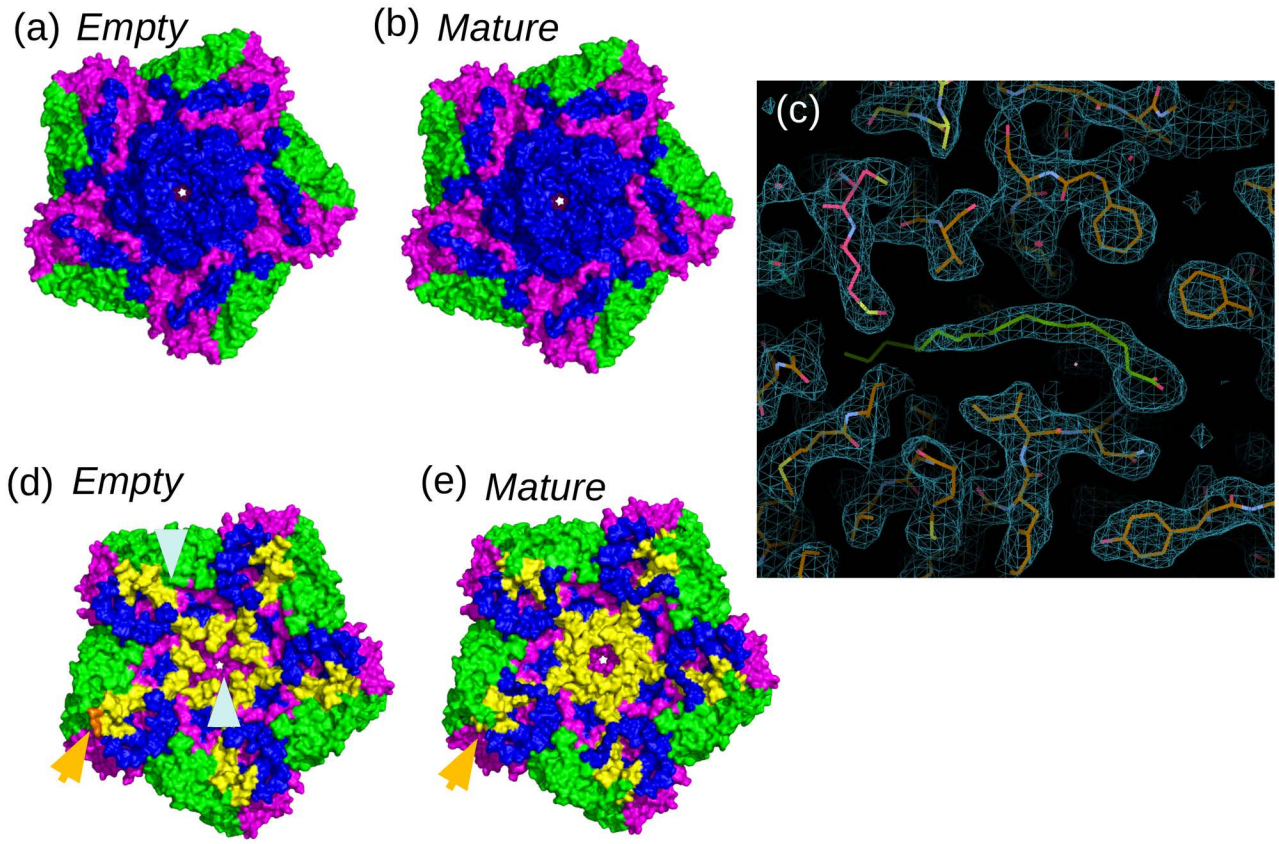


Figure 3.18: (a, b) Pentamer surfaces for empty and mature particles. VP1 is coloured blue, VP2 in green and VP3 in magenta. (c) 5-fold NCS averaged density showing the P-site for EV-F3. (d,e) Inner surfaces of empty and mature EV-F3, coloured as in (a,b), with VP4 in yellow. The VP0 for empty EV-F3 has been coloured according to the VP4/2 divide. Pale blue arrows in (d) point to missing regions relative to the mature particle, and the orange arrows highlight the position of the scissile bond.

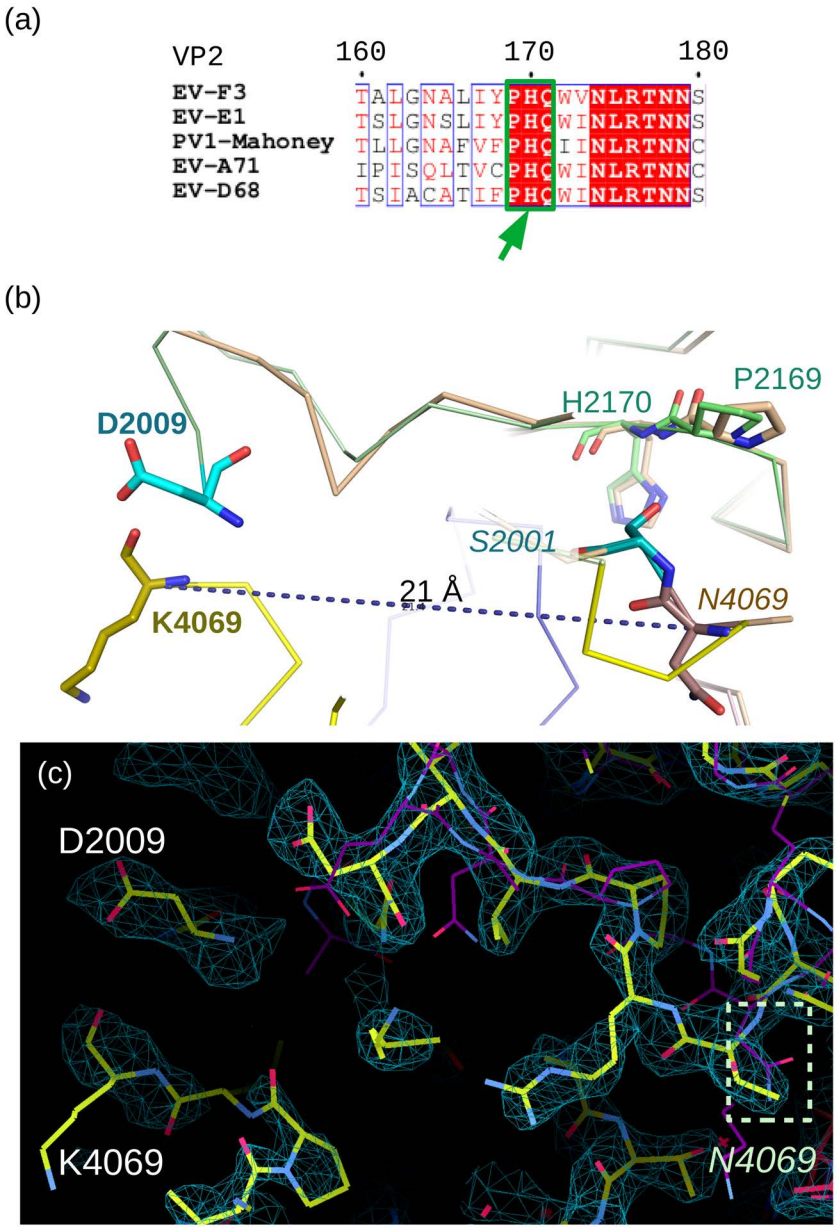


Figure 3.19: (a) Sequence comparison between selected enteroviruses, with the highly conserved PHQ motif, involved in VP0 cleavage, boxed in green. Residue numbering is based on EV-F3. (b) Comparison of EV-F3 empty's modelled VP0 and that of PV-1 (PDB: 1POV) (Basavappa et al., 1994). Residue numbers in italics correspond to PV-1, which is shown in brown ribbon representation except for the two residues either side of the scissile bond and the His and Pro catalytic residues (all shown as sticks). EV-F3 chains are coloured by convention (yellow for VP4, green for VP2 and blue for VP1). The two modelled EV-F3 residues closest to the scissile bond (not modelled) are labelled in bold text and are shown in stick representation, while the catalytic residue H2170 and P2169 are shown as green sticks. The blue dashed line shows the distance between the EV-F3 and PV-1 residues that precede the scissile bond. The inner capsid surface is at the far left of the image. (c) 2Fo-Fc map of EV-F3 empty, contoured to 1 sigma around the modelled VP0 for EV-F3 (thick sticks). The model for PV-1 empty (as above) is shown as thin pink sticks and associated pertinent residues are written in italics.

3.2.3 A continuum of Uncoating Intermediates

Genome containing A-particle and empty B-particle states for other enteroviruses, such as EV-A71 and PV, have been induced via heating (Shingler et al., 2013; Levy et al., 2010). To see if it was possible to isolate EV-F3 A and B-particles, a similar procedure was followed.

A suitable heating temperature was chosen based on thermofluor measurement at the point of RNA release, where A-particles transition into empty B-particles. Previously, hydrophobic dye signal transitions were observed at 46 °C and 58 °C before final denaturation at 89 °C (Walter et al., 2012). The first melting point could represent a transition to an expanded A-particle, consistent with the observation that EV-F3 infectivity diminishes when subjected to temperatures of 40 °C and above for over 15 mins (Walter et al., 2012). This was corroborated using RNA-release thermofluor assays (see Chapter 4, Sections 4.1.1 and 4.1.3), showing genome release to occur at 55 °C (Walter et al., 2012). The temperature chosen to investigate genome release was taken as the midpoint of the RNA release curve, where half of the particles have released their genomic material and the remaining half are likely A-particles. Thus, for cryo-EM investigation, purified virus was heated to 55 °C for 8 mins using a PCR machine, and then returned to 4 °C before plunge freezing as described in Chapter 2, Section 2.5.2.

Data Collection

A total of 2173 micrographs were collected using a Glacios microscope (Thermo Fisher) and FIII detector operating in linear mode (Table 3.5) at 92 kX nominal magnification, see Chapter 2, Section 2.5.3.

Electron micrographs showed a mixture of 'filled' and empty capsids, with patches of darker background outside of the particles consistent with released genome, and perhaps representative of particles captured at differing points of genomic egress, Figure 3.20. Notably, no pentamers were observed in collected micrographs. As may be seen in Figure 3.20(a), partly or totally 'unfilled' particles appeared to have a rougher outer surface (for instance, those marked by a black '>'), while some particles, such as that marked with a magenta '>', the genomic interior seemed to be more featured, as if it were slightly ordered.

Given the diversity and the relative sparsity of particles, the 'quantity not quality' approach was taken to data collection; linear mode at relatively low magnification, since particle heterogeneity

Table 3.5: Data processing statistics, as determined by Relion and Phenix for EV-F3 B, I and A particles. 'Particles used' refers to particles used in the final reconstruction.

**FSC=0.143 criterion, as calculated in Relion3.0 **FSC=0.143 criterion, from Phenix.*

	B	I	A
Particles used	8104	783	1620
Resolution*, FSC=0.143 / Å	3.71	4.08	4.32
Resolution**, masked / Å	3.59	3.99	4.18
Resolution**, unmasked / Å	3.71	4.09	4.31
RMSD angles / °	0.71	0.79	1.2
RMSD lengths / Å	0.006	0.004	0.004
Ramachandran / %			
Most favoured	93.0	93.6	93.3
Allowed	6.53	5.97	6.22
Outliers	0.46	0.46	0.46
CCmask / Å	82	76	72
CCmain chain / Å	81	76	70
CCsidechain / Å	80	74	70
Rotamer outliers / %	0.53	0.71	0.53
Clash score	5.81	4.79	5.75
Number of Atoms (monomer)	677	677	675

and thus numbers were anticipated to be the greatest limiting factor for resolution.

Processing

A total of 10,957 particles were selected using Warp, as before (see Chapter 2, Section 2.5.4), without discrimination between filled and empty particles.

Three distinct particle forms were observed at the classification stages; particles devoid of genomic density (8104 particles), and two different sets of particles with clear genetic material; one (783 particles) with a more expanded internal density cloud than the other (1620 particles), Figure 3.21. Interestingly, an appreciable number of 2D classes showed projections of particles with 'chunks' of missing capsid, suggesting that heating triggered irregular capsid degrada tion, Figure 3.21(a). This is in accord with the symmetry-relaxed CV-B3 A-particle structure reported recently (Lee et al., 2016).

These three forms were refined separately with icosahedral averaging, yielding resolutions (for FSC 0.143) of: 4.08 Å, 4.32 Å, and 3.71 Å for filled ('A'), part-filled ('I') and empty ('B') particles respectively (Figure 3.20(b,c)).

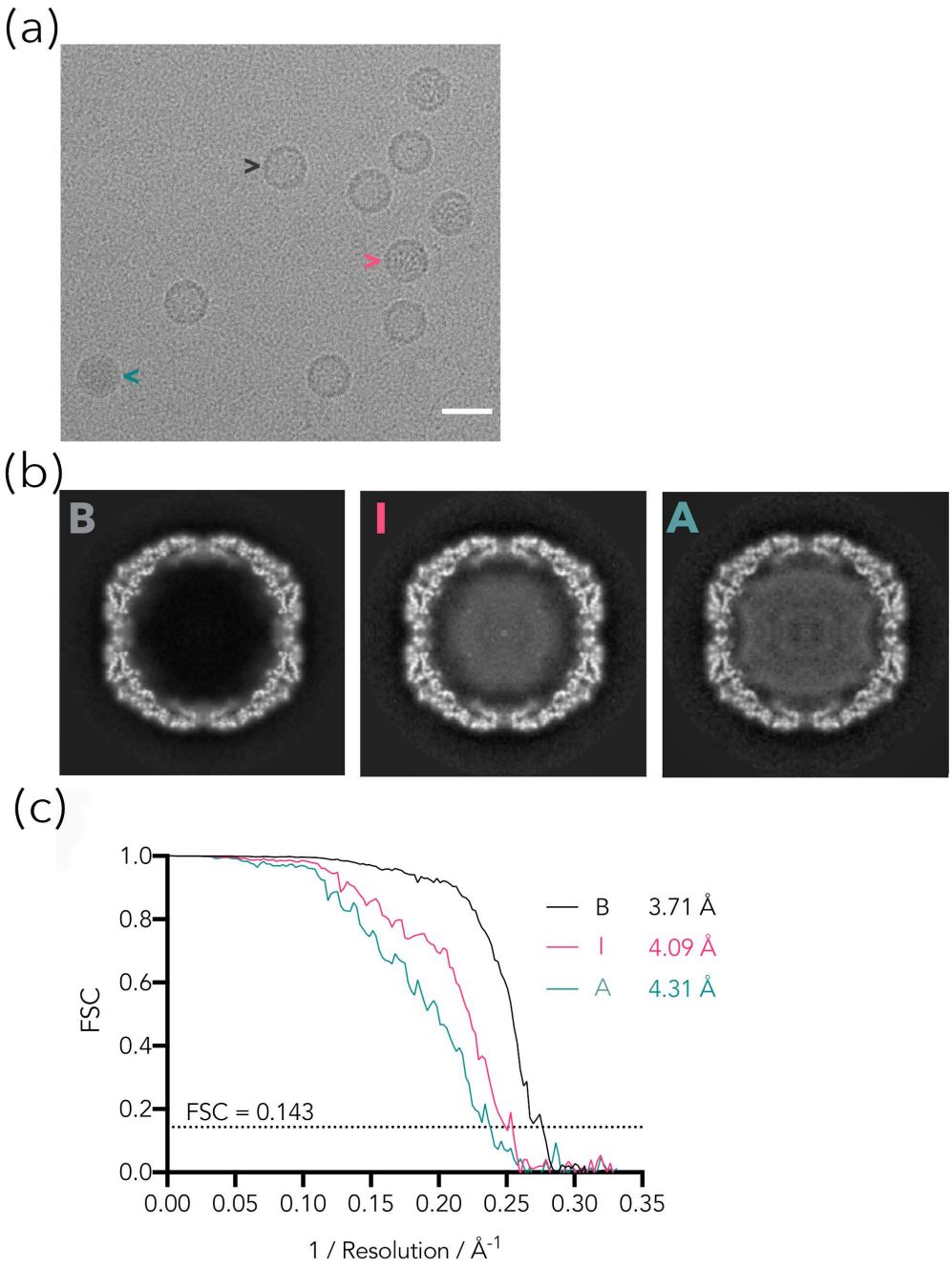


Figure 3.20: (a) Example micrograph with each particle type marked with an arrow head, coloured according to particle type. Scale bar (white bar) = 50 nm. (b) Central slices from density maps from the final refinement of particles B, I and A. (c) FSC plot for A, B and I, with FSC = 0.143 marked as a dashed line. Resolutions determined by the gold standard are marked for each particle.

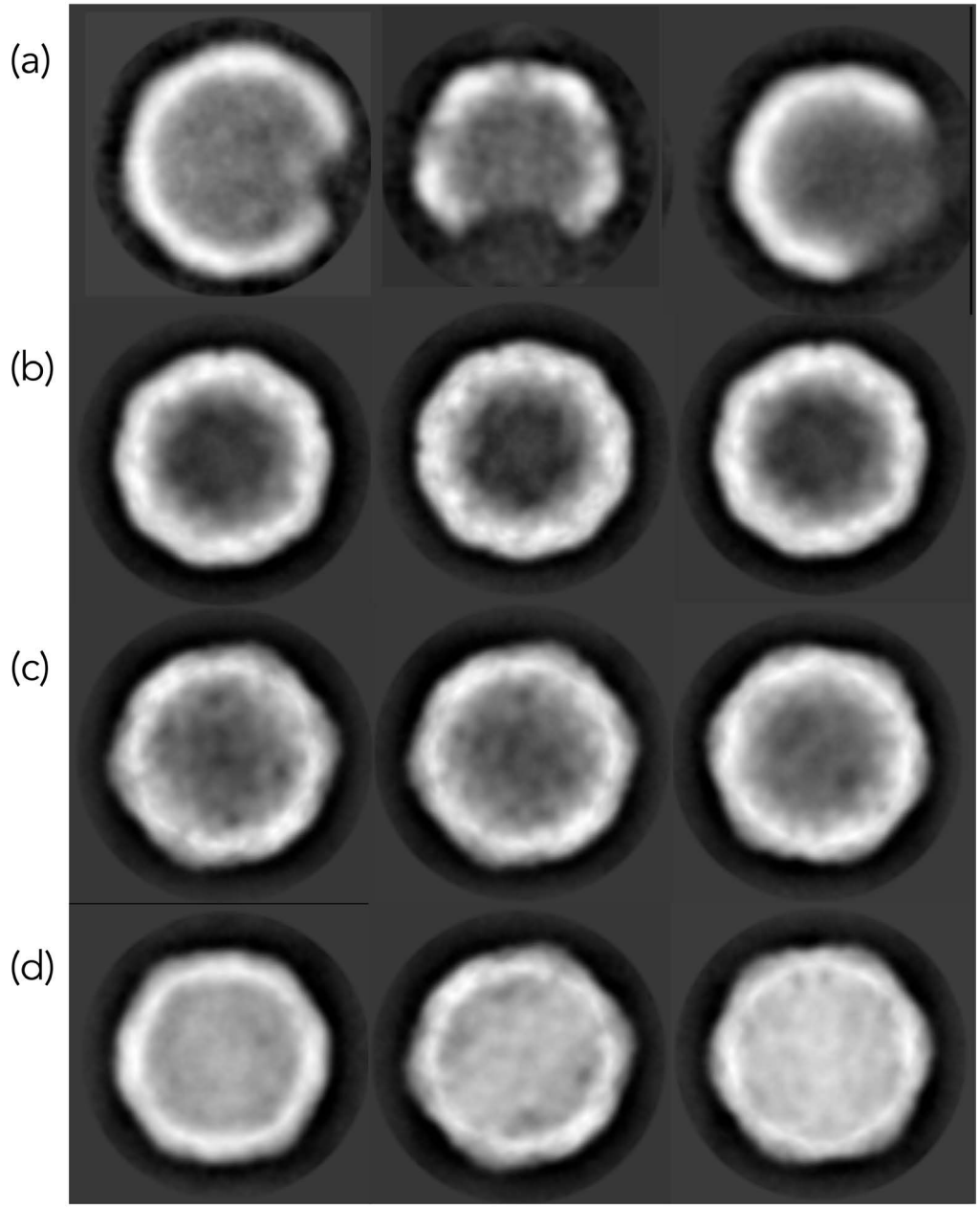


Figure 3.21: Example 2D classes from EV-F3 processing. (a) Broken particles; (b) Empty particles, B. (c) Partly-filled particles, I. (d) Filled particles, A.

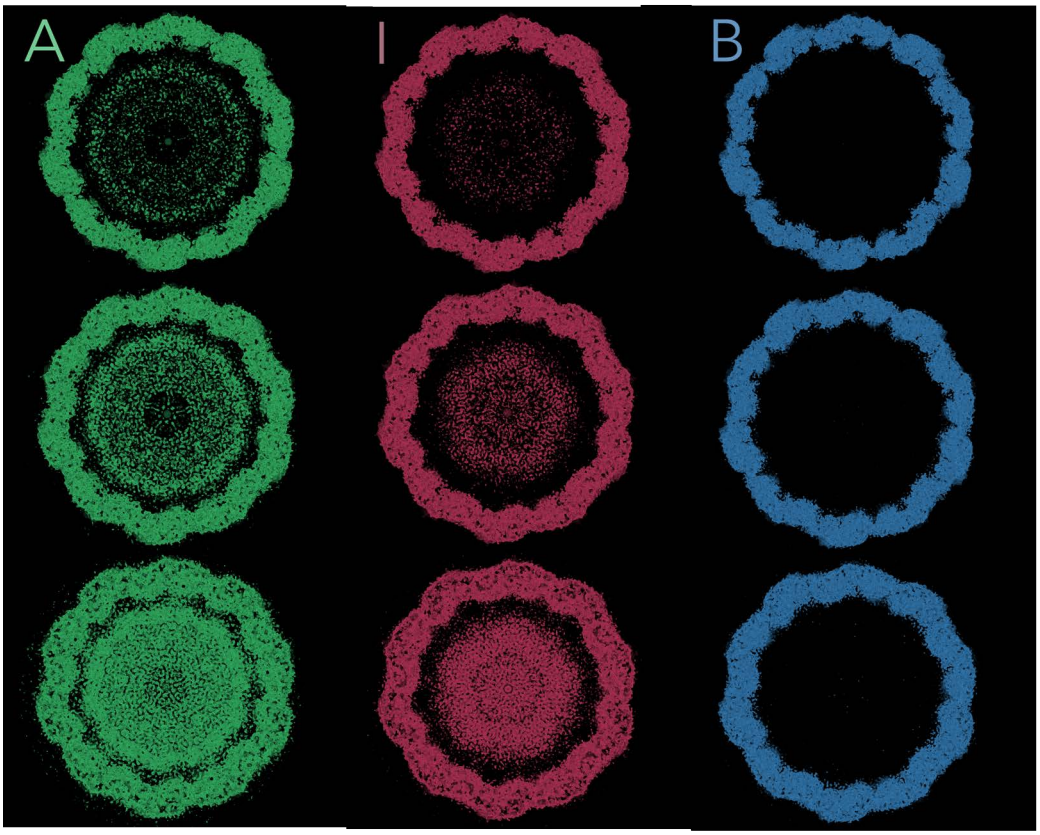


Figure 3.22: Central slice of maps for A, I and B particles (as labelled), coloured green, pink and blue, respectively, and shown at three contour levels; from high (top) to low (bottom).

Genomic Content

The majority of particles were completely devoid of internal genomic density i.e. B-particles, consistent with observations with PV-1, where heating to 56 °C resulted in a predominant empty B-particle population (Levy et al., 2010), but unlike EV-A71, where A particles were isolated following heating for 12 min at 56 °C (Shingler et al., 2013).

Central sections of the density maps show strong internal density, potentially corresponding to retained genomic material for EV-F3 A, and EV-F3 I, Figures 3.20(b), 3.22. Intriguingly, the genomic interior for EV-F3 A appears layered at medium contour levels (top left and middle left, Figure 3.22), consistent with ordering of internal RNA as observed for CV-A16 and CV-B3 A-particles (Ren et al., 2013; Lee et al., 2016).

Table 3.6: Calculated radii as determined by Pymol using distances between C_αs and the capsid origin (see main text) are in the first table segment (first three rows); and radii determined through the ViperDB (Carrillo-Tripp et al., 2008) in the second (middle set of rows) segment. Note that 'average' is shortened to 'av'. PDB codes, where appropriate, are in parentheses. Associated publications from left to right are: Roedig et al. (2017); Wang et al. (2012); Shingler et al. (2013); Verdaguer et al. (2000); Garriga et al. (2012). Percentage differences i.e. (expanded - native) / native * 100 are stated in the final table segment. 'Reported' percentage expansion values are those quoted in associated publications. The inner radius is not particularly reliable as a measure, since this is highly dependent on whether VP4, for instance, has been modelled, which could account for the highly variable values for the percentage differences.

	EV-F3m (5OSN)	EV-F3-B	EV-A71 (3VBS)	EV-A71-A (3J22)	HRV2 (1FPN)	HRV2-B (3TN9)
Av / Å	132.2	137.0	135.8	139.0	127.5	142.3
Max / Å	157.0	161.9	154.3	165.9	159.5	165.8
Min / Å	108.0	113.4	115.7	116.3	107.2	119.9
ViperDB Inner / Å	107.0	-	105.0	115.0	104.0	115.0
ViperDB Av / Å	153.0	-	155.0	161.0	157.0	163.0
ViperDB Outer / Å	158.0	-	159.0	167.0	160.0	166.0
Diff Av / %		3.7		2.4		11.6
Diff Min / %		4.9		0.5		11.9
Diff Max / %		3.1		7.5		3.9
Diff Av VDB / %		-		3.9		3.8
Reported / %		-		4.0		3.8

Modelling

To check the magnification calibration, a protomer of the XFEL EV-F3 structure (PDB: 5OSN) was fitted into map B (the highest resolution of the three particle classes) in Chimera scaled to pixel sizes ranging between 1.50-1.55 Å/pix (Roedig et al., 2017; Pettersen et al., 2004). The correlation of fit, as reported by Chimera, was then recorded in each instance. This was repeated for an expanded enterovirus particle PDB: 3J22, and results corroborated with the map for particle A (Shingler et al., 2013). The best fit occurred for 1.54 Å/pix in both cases and hence was assumed to be correct. This method was simple, but probably reduced the error in magnification to no more than about 0.3 %.

A protomer of the XFEL EV-F3 structure (Roedig et al., 2017), excluding VP4, was fitted into the B particle map using Chimera (Pettersen et al., 2004), before being extended into an 8-mer, such that, for the protomer used for model building, all contacts with surrounding protomers were represented. Each protomer was initially treated as a separate rigid body in Coot (Emsley

and Cowtan, 2004; Emsley et al., 2010), and the resulting model inspected to establish the quality of the fit based on the reasonableness of inter-protomer contacts, and agreement with the map. This per-protomer rigid body fit was essential for an initial model, since the separation between adjacent protomers was clearly substantially greater than for the native virus. Associated portions of EM maps were then extracted with the resulting 8-mer using phenix.map_box before real-space refinement in Phenix (Adams et al., 2010, 2011). Subsequent modelling was performed in Coot on the central protomer (Emsley and Cowtan, 2004; Emsley et al., 2010), from which a new 8-mer was generated for each refinement cycle. To avoid clashes, refinement was performed with hydrogens added, but, final validation statistics were checked with the hydrogen atoms removed.

Overall Structure

A, I and B maps are consistent with very similar underlying structures, although the higher resolution of EV-F3 B's map means that β-strands are better separated, and bulky side chains mostly resolved. For EV-F3 B's map, there was generally good model/density agreement for VP1-3. Quality was, as expected (based on resolution estimates, Figure 3.20(c)), poorer for EV-F3 A, with the majority of side chains not resolved (Figure 3.24). Consistent with VLP structures described in Chapters 5 and 6, there was no density for the majority of the N-terminal portion of VP1, contrasting with the CV-A16 and PV A-particle models, where the VP1 N-terminus could be seen exiting via the off-axis channel (Ren et al., 2013).

Radially coloured surfaces shown in Figure 3.23 demonstrate a clear expansion relative to the native virion for maps A, B and I (note that, for the purposes of this section, EV-F3 mature will be referred to as EV-F3 native). The average radius was determined as ca. 137.0 Å (Carrillo-Tripp et al., 2008; Roedig et al., 2017) by finding the average of distances between the capsid centre and every C_α in the capsid using Pymol (DeLano et al., 2002) (Table 3.6), and corresponds to an expansion of under 4 %, relative to the native EV-F3 model. This is consistent with previous reports (Shingler et al., 2013; Garriga et al., 2012).

A number of notable features were observed around each of the symmetry axes for A, I and B, and will be outlined in the sections below.

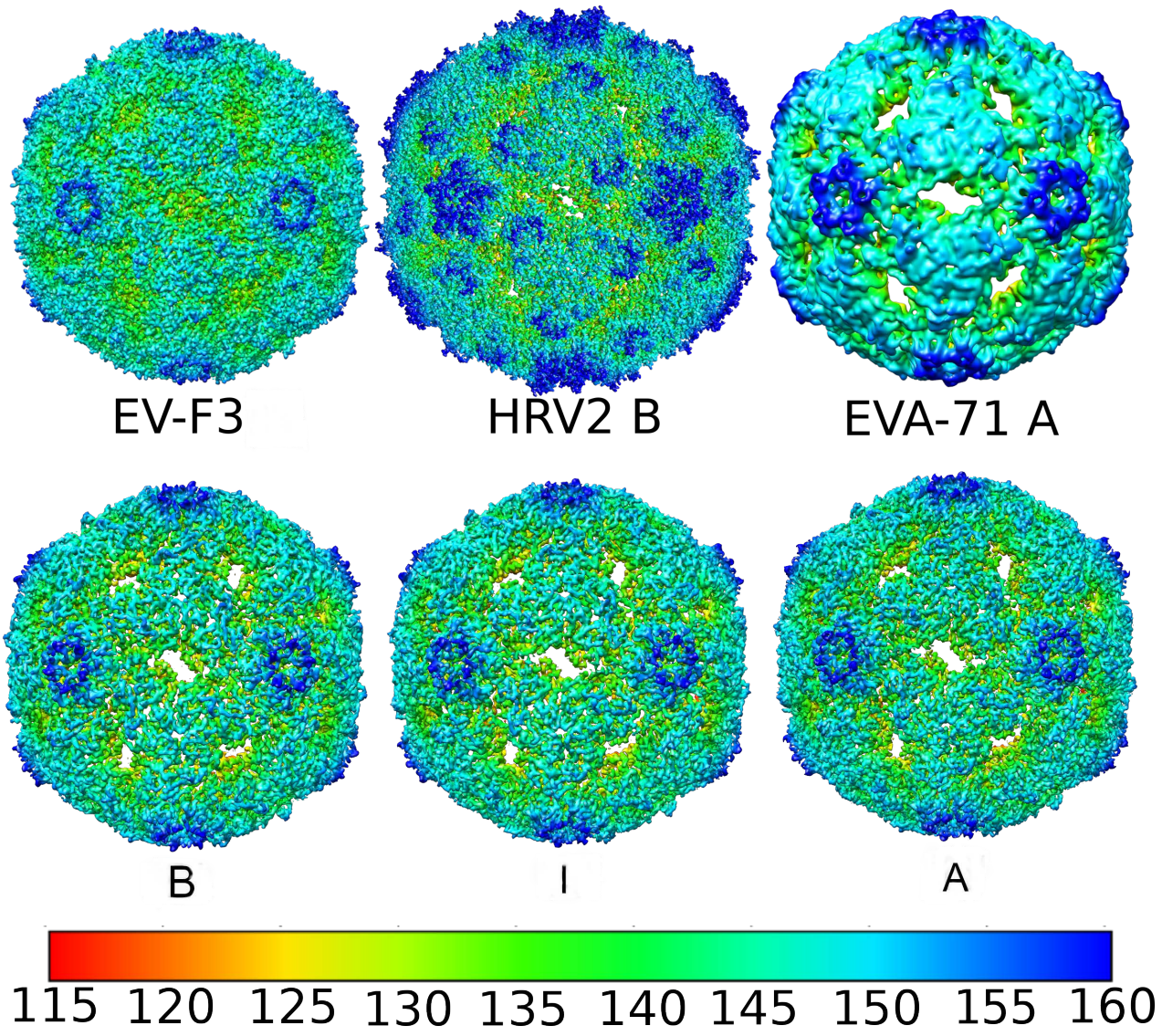


Figure 3.23: Radially coloured surfaces generated by Chimera showing the extent of EV-F3 capsid expansion. The multi-coloured bar represents radius in Å. Bottom: our heated reconstructions, B, I, A; top left native EV-F3, and middle and left taken from the PDB and EMDB, respectively, PDB codes: 3TN9 and EMD-5364 corresponding to 3J22 (Garriga et al., 2012; Shingler et al., 2013). For the HRV2 structure, molmap in Chimera was used to generate a surface (Pettersen et al., 2004).

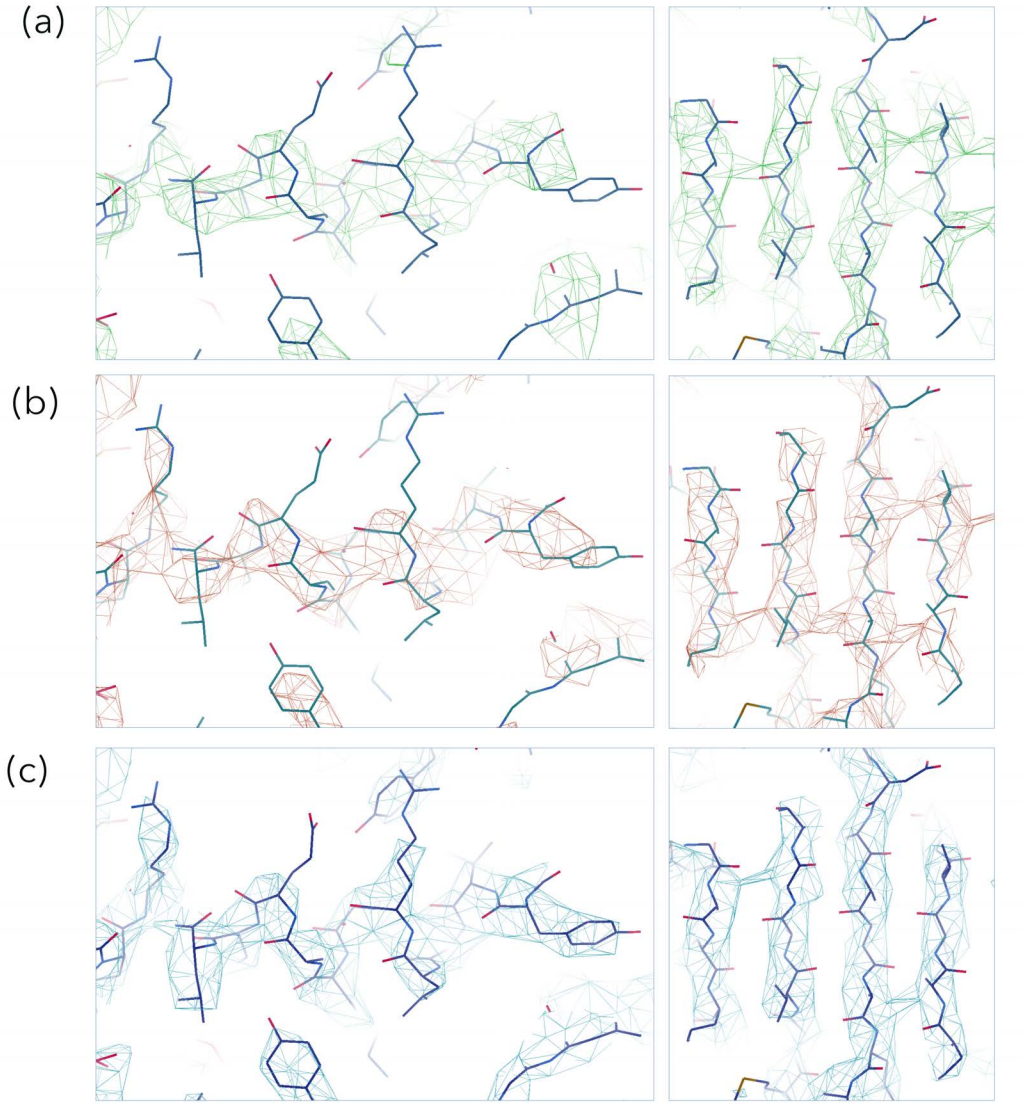


Figure 3.24: (a-c) Examples of density quality for EV-F3 for A (green), I (red) and B (blue) particles, respectively. Images on the left-hand side shown an example helix; on the right an example set of sheets.

Pseudo-3-fold Axis

The overall topology of B, A and I is very similar, and differs substantially from that of the native EV-F3 model (Figure 3.25(a-d)). Akin to observations with the PV-1 C-antigenicity particle described in Chapter 6, and the previously characterised 135S particles, such as in Butan et al. (2014), we see an opening at the pseudo three-fold axis (Figure 3.25). This seems to be largely due to remodelling of the VP3 GH-loop. Density for particles A, I and B support a continuum of VP3 GH-loop conformations, with significantly stronger and more extended density for A (Figure 3.25(a,e)) than I and B (Figure 3.25(c-g)).

As may be seen in stereo-representations of this loop (Figure 3.26), density for A actually partly follows the loop conformation modelled for mature EV-F3. Density for the tip of this loop diminishes for particles I and B (green and blue density, Figure 3.26), and becomes clearer around the leg towards the C-terminal portion of the loop at residues E3191-A3192 (Figure 3.26(a) indicated by blue and purple arrows, respectively), consistent with the loop transitioning into a hairpin-like structure. This hairpin-like arrangement is similar to that observed for CV-A16 (previously a helix for the native virion), which has been seen to vary in extent between 80S (B) and 135S (A)-like particles (Ren et al., 2013). Using CV-A16 as a guide, this loop was modelled into the density, but had to be truncated due to map quality (Figure 3.26).

2-fold Axis

Another striking feature shown in Figure 3.25 is at the 2-fold axis, located close to the proposed site of genomic egress, formed from a separation of helices (residues 2087-2096) from opposing VP2 protomers by (based on Q2092-Q2092 C_αs in opposing protomers) 13 Å for EV-F3 B (EV-F3 A and I models are almost identical). The C-terminal portion of VP2, which may facilitate 2-fold axis expansion (Ren et al., 2013), is totally disordered from residue 2237 onwards in all three structures, and was removed from the models. Particles A and B differ slightly with respect to density for the VP2 hairpin (residues 2041-2057) that sits below the VP2 2-fold helices (Figure 3.27). For particle A, some density may be seen at low contour levels; only a little for I and none for B, even at very low contour levels.

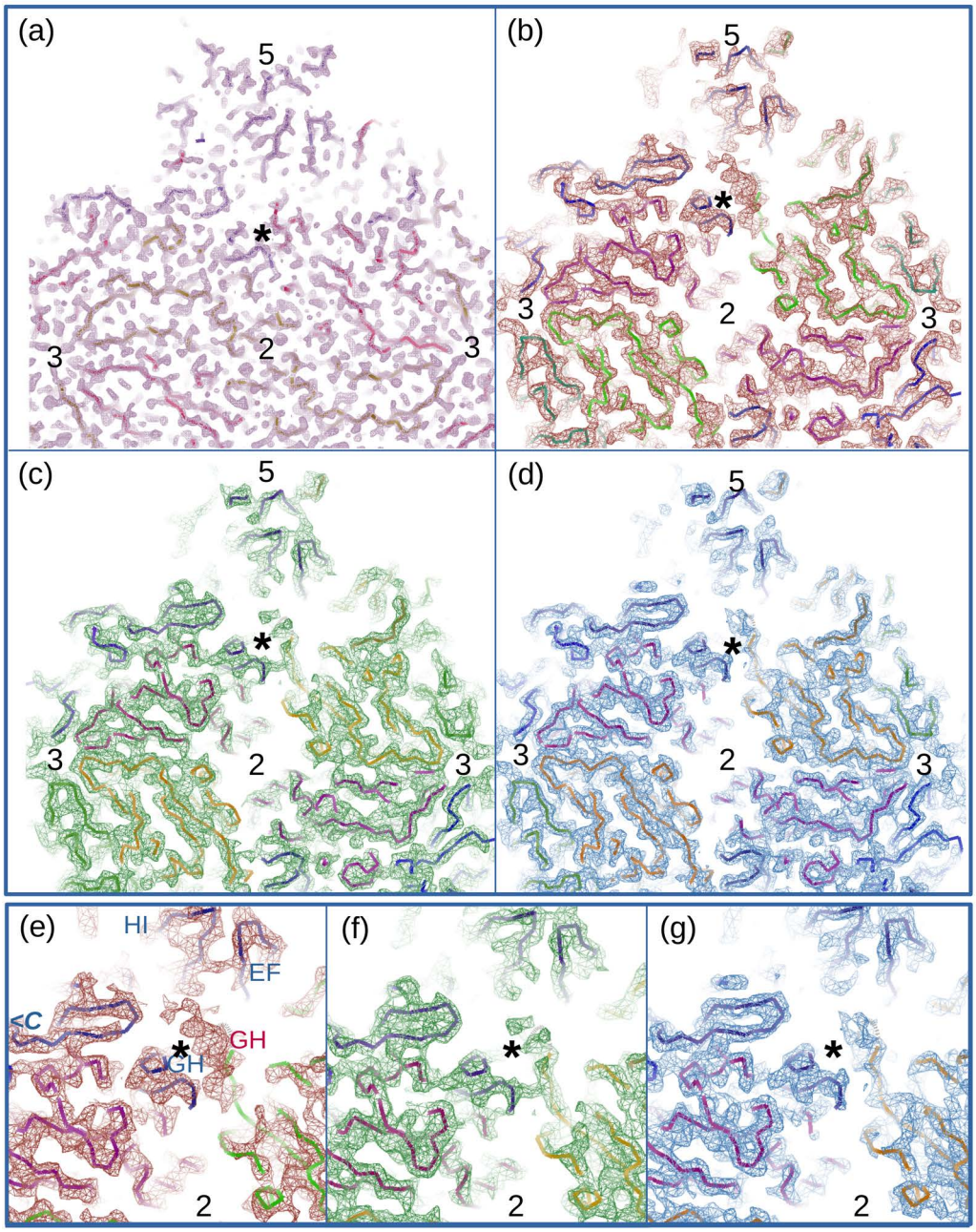


Figure 3.25: Capsid surface topologies, with the pseudo-threefold axis marked with an asterisk and numbers corresponding to the symmetry axes, for (a) Native EV-F3. (b) EV-F3 A; (c) EV-F3 I; (d) EV-F3 B. A zoomed in view of (b-d) can be seen in (e-f), with pertinent loops labelled in (e) in magenta text for the VP3 GH loop and blue text for the VP1 loops. The VP1 C terminal region is marked with a blue *C*.

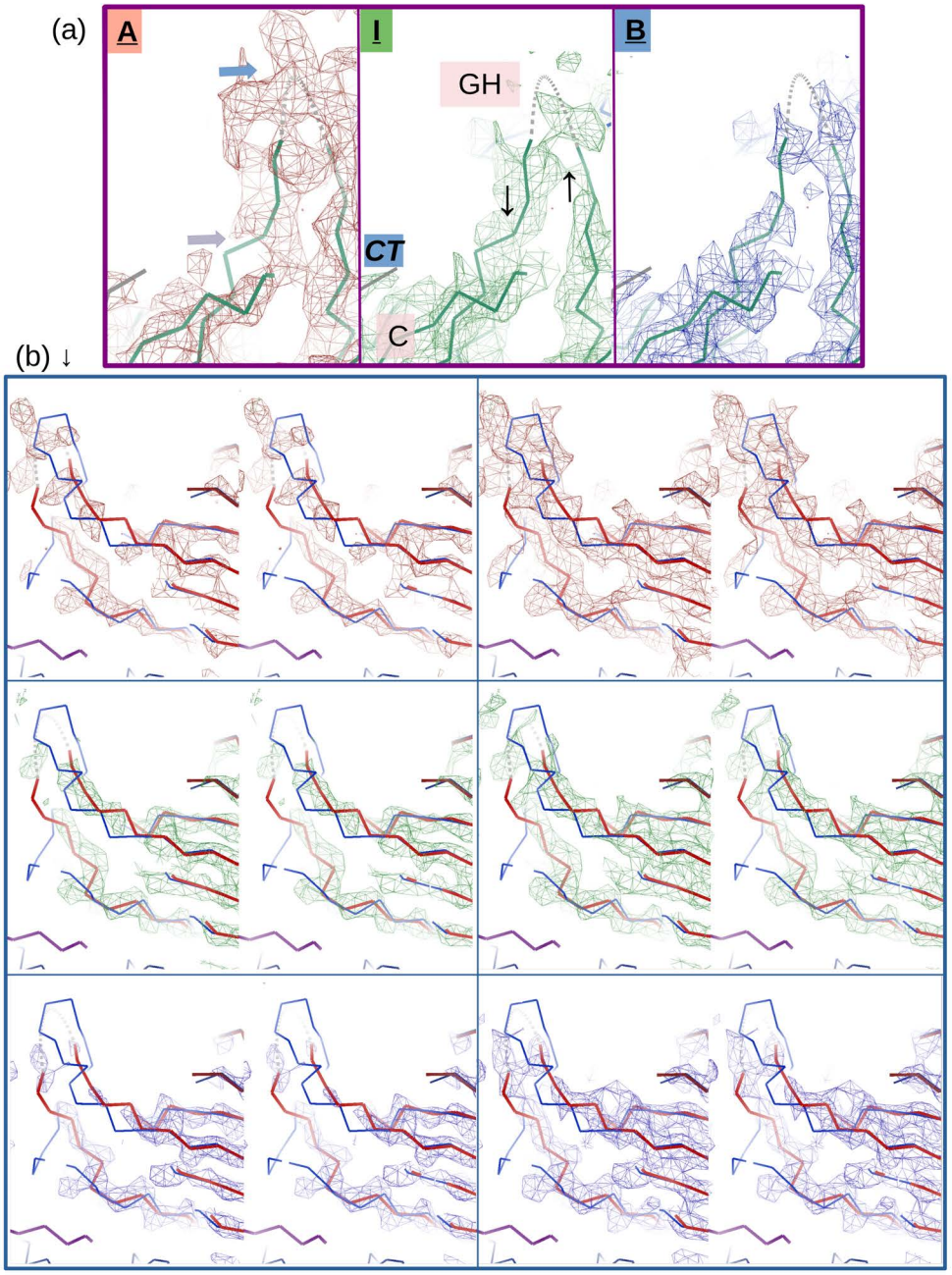


Figure 3.26: Changes at the VP3 GH loop. (a) Purple box: Front view of VP3 GH loop with EV-E1's model shown in green. The directionality of the residue numbers (from N to C termini) is shown as black arrows for I, and the C strand and VP1 C-terminus that can be seen towards the bottom left of the image are labelled on I. Two key changes in the density between particles A-B are indicated with purple and blue arrows in A. (b) Blue box: Wall-eyed stereo representation of the VP3 GH loop for particles A, I and B (red, green and blue maps, respectively) at low (right) and high (left) contour levels. The native EV-F3 model is shown as blue thin ribbon and the loop modelled for particle B in thick red ribbon. The dashed line for this latter model indicates the portion of residues omitted due to the map quality being too poor to confidently model a different conformation. The bottom right box is not in stereo representation, and indicates the two crucial points of change in density with blue and purple arrows. The C-terminal direction is marked with a small arrow above the blue boxto show the direction of the loop.

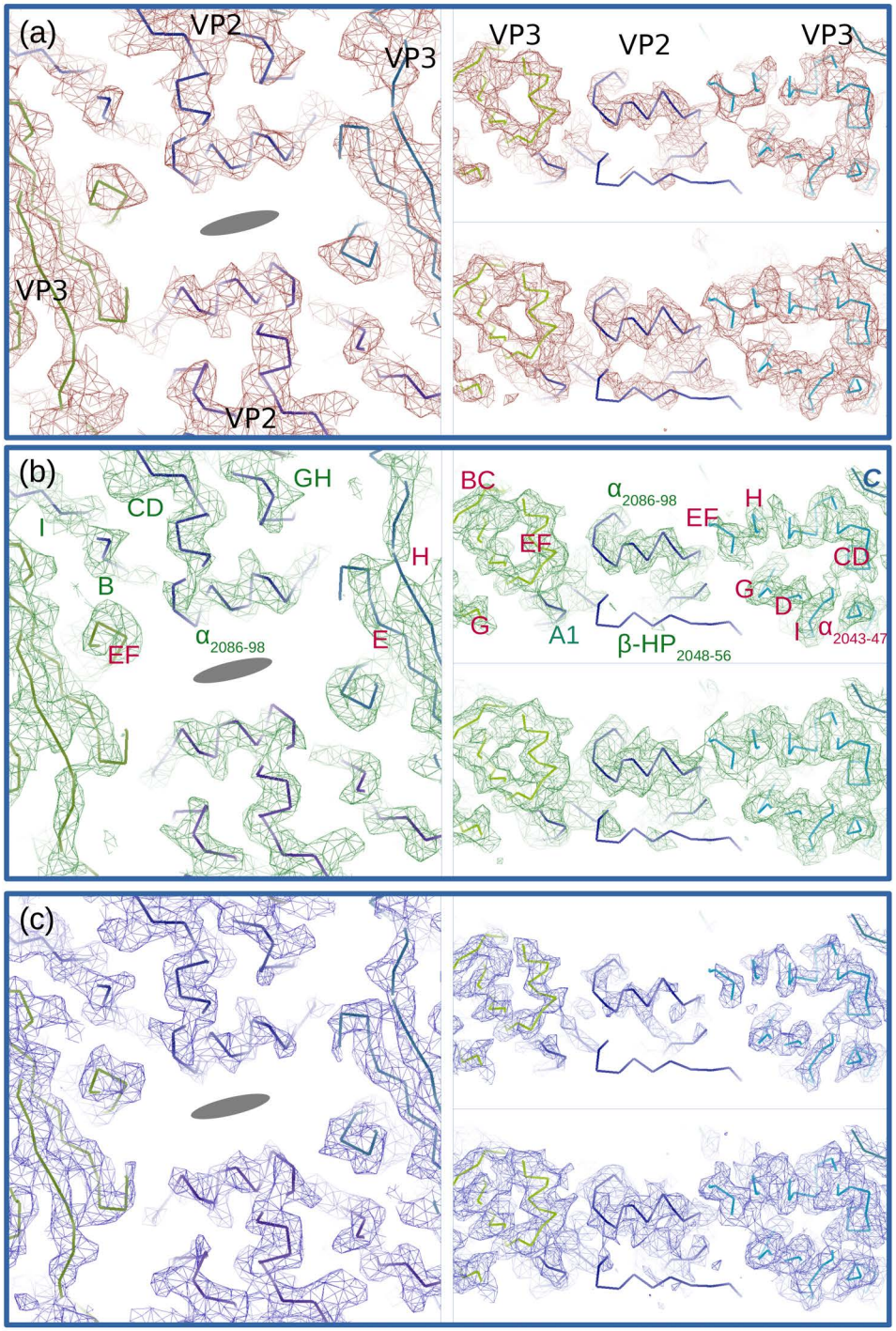


Figure 3.27: Comparison of (a) Particle A, red density, (b) Particle I, green density and (c) Particle B, blue density at the 2-fold axis. Images in the left-most column look down from the capsid surface onto the two-fold axis channel, whilst the two images on the right-hand side for each particle show a side view of one 2-fold VP2 helix (dark blue) and the β hairpin model that sits below (dark blue) at high (top of box) and low (bottom of box) contour levels. VPs are labelled in (a) and apply to (b-c), while loops and strands are labelled in (b) in green type for VP2 and magenta type for VP3. The model for B is shown in ribbon representation for all structures to aid comparison.

5-fold Axis

At the base of the 5-fold axis, we see a curious plug of density, towards the base of the channel. For I, two additional 'blobs' of density can be seen further up the axis channel (Figure 3.28), but the significance of this is unclear.

P-site

VP1 surface loops DE, HI, EF and BC that surround the 5-fold axis channel appear slightly shifted relative to native EV-F3 (Figure 3.29(a)), such that the VP1 β-barrel (P-site) is slightly collapsed. Within the P-site, strands C, D, G (residues 1172-1177) are pushed closer into the pocket by approximately 1 Å relative to native EV-F3, thereby constricting the pocket, Figure 3.29.

3-fold axis

VP2 and VP3 HI-loops (2216-2217, 3206-3208) that encircle the 3-fold lip are clearly flattened relative to native EV-F3 (Figure 3.30), and appear to sit closer towards the 3-fold axis, with a diameter (defined by distance between the C_αs of opposing VP1 and VP3 loops) measuring ca. 14 Å (measured as 15 Å using the same method for native EV-F3).

Within the capsid interior, there seems to be a shift in the VP2 N-terminal residues 2011-26 that form the A1 and A2 β strands that sit at the base of the three fold axis, relative to native EV-F3. Both strands and the connecting loop seem to be shifted away from the bulk capsid, and towards the adjacent protomer, Figure 3.30.

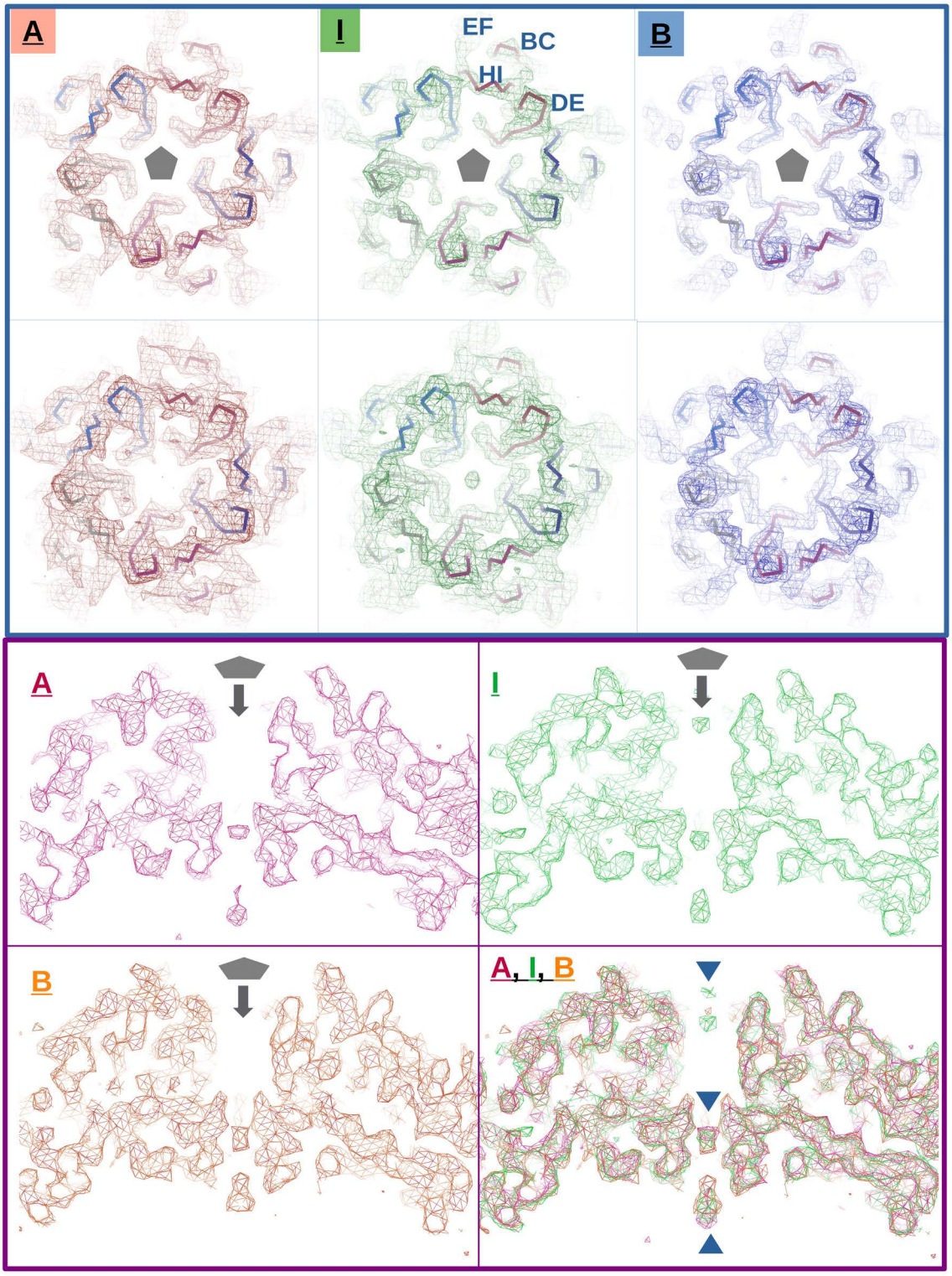


Figure 3.28: Blue box: Top and bottom panels show the top view of the 5-fold axis channel for A, I and B, as labelled at high (top) and low (bottom) contour levels. VP1 loops are labelled for I, but apply to all top views shown. Purple box: Slice through density at the 5-fold axis channel for particles A (pink), I (green) and B (orange), as labelled. The image on the bottom right shows an overlay of all three particles, with mysterious channel densities marked with blue arrows.

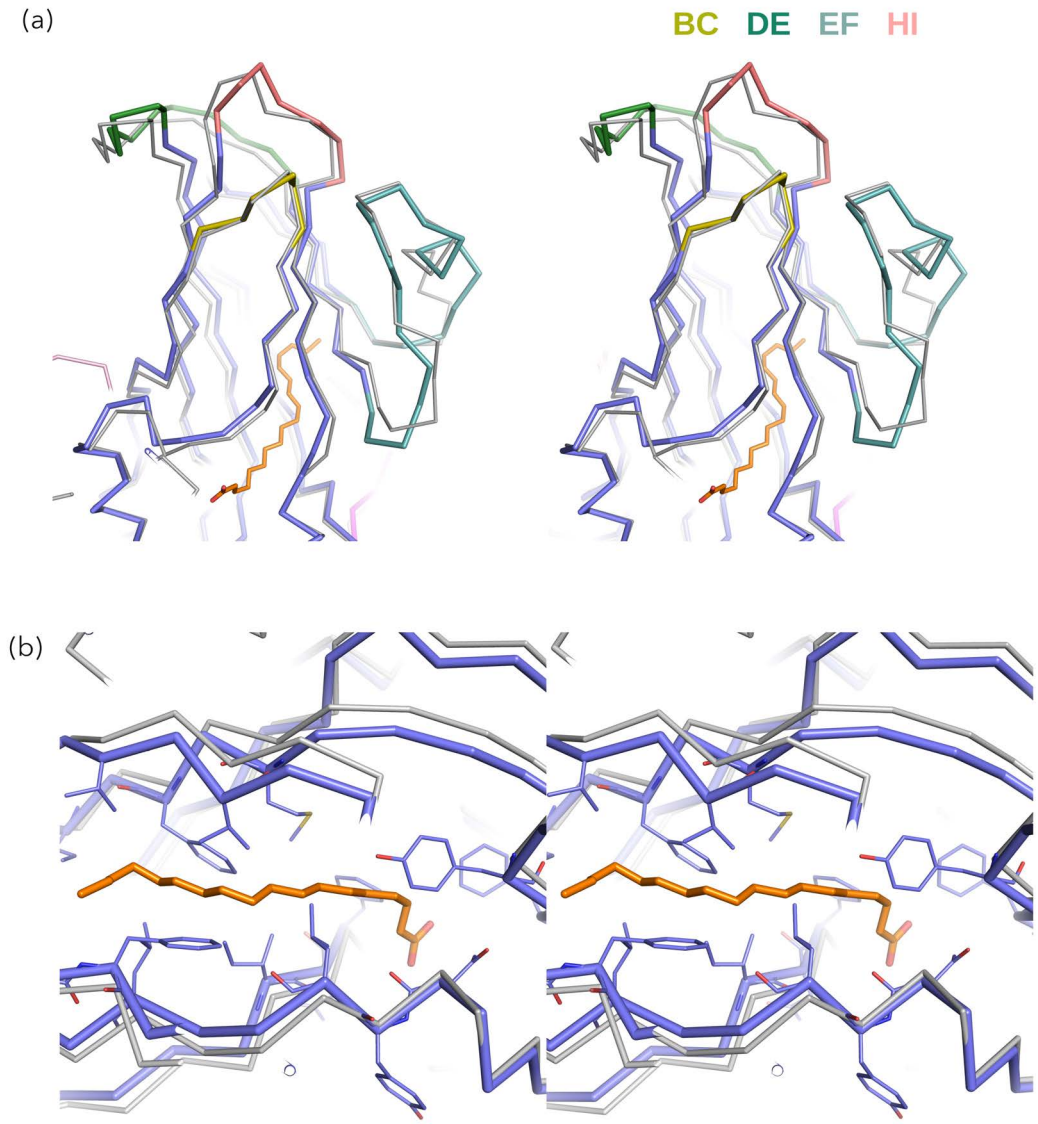


Figure 3.29: Wall-eyed stereo representation around the P-site. EV-F3 mature is shown in grey for comparison. The pocket factor, steric acid, is shown as orange sticks. (a) VP1 surface loops around the 5-fold axis for one protomer. Pertinent loops for EV-F3 are coloured according to the key at the top right; VP1 BC in yellow, DE in green, EF in teal and HI in salmon.The remainder of EV-F3 VP1 is shown in blue. The VP3 of the adjacent protomer is partly visible on the left-hand side, and is shown as a pale pink ribbon. The VP3 of the protomer of interest in this picture is shown in magenta. (b) Stereo view of the P-site for EV-F3 B (blue), with EV-F3 native (grey) overlaid.

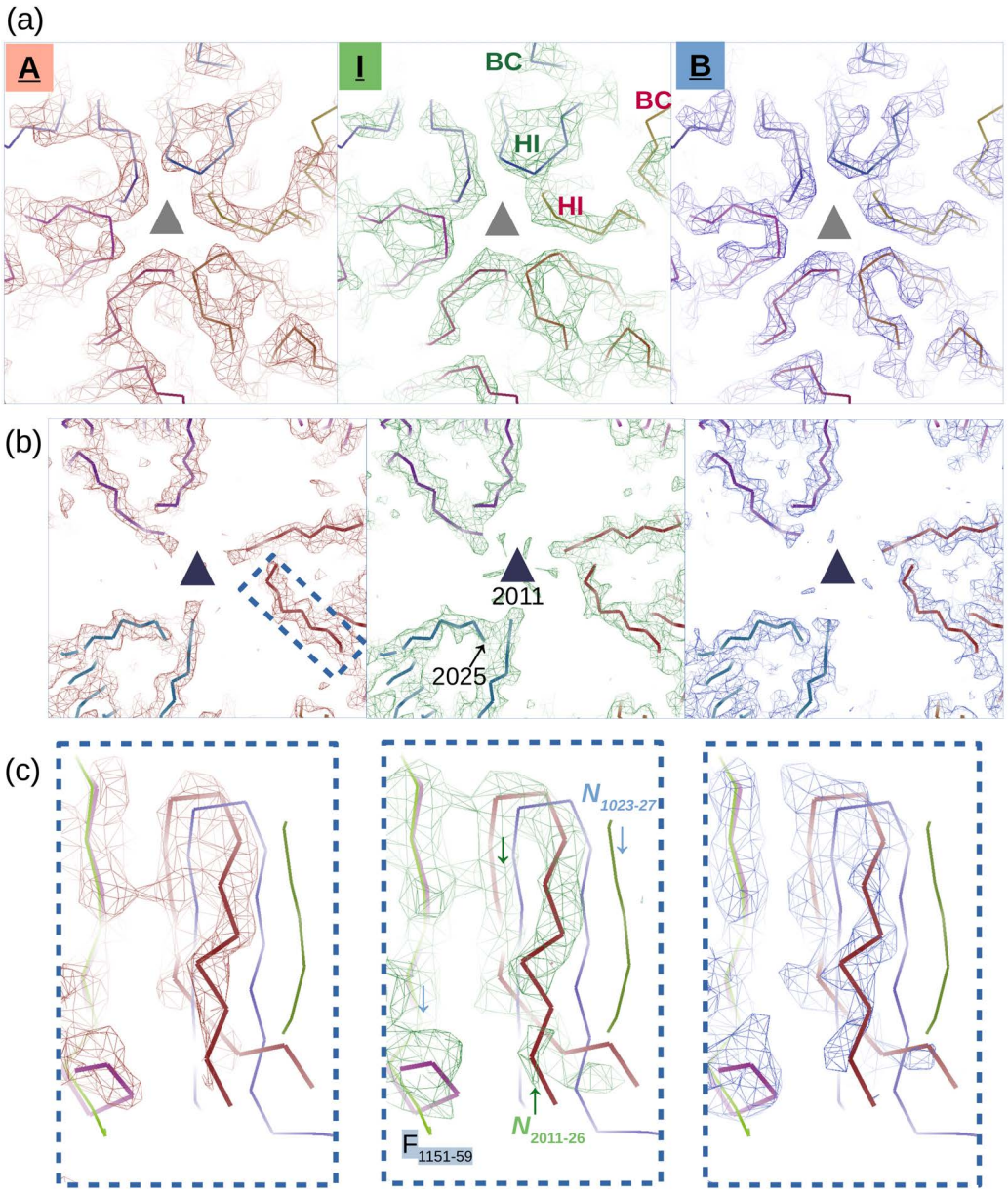


Figure 3.30: (a) View from the capsid surface of the icosahedral 3-fold channel for particles A, I and B, as labelled with models shown in ribbon representation. The VP2 and VP3 HI and BC loops are labelled in green and magenta text, respectively for I. (b) View of the base of the three-fold axis from the capsid interior, coloured as for (a). (c) Side view of the VP2 N-terminal region 2011-2026 (βA1-βA2 strands) boxed in (b) in thick, red ribbon, overlaid with native EV-F3 (thin blue ribbon, EV-F3-CG structure presented in Chapter 4 Section 4.1.3) to demonstrate the movement observed in our model, relative to native EV-F3. The βA1-βA2 strands is flanked by the N terminal residues 1023-1027 for native EV-F3 on the right hand side, and the VP1 F strand (residues 1151-1159 on the left) also for native EV-F3. Arrows point towards the C-terminal direction for each strand shown.

3.3 Discussion

Many parallels between EV-F3 and other enteroviruses may be drawn from the structures presented in this Chapter.

Foremost is the presence of a hydrophobic pocket with density consistent with the modelling of a long-chain fatty acid, such as stearic acid. The relatively extended P-site of EV-F3 capsids that enables accommodation of this elongated pocket factor, is similar to that seen for other enteroviruses, such as PV, SVDV, EV-A71 and CV-A16 structures (Chapter 1 Section 1.3.1 Table 1.5, for instance, Wien et al. (1997); Fry et al. (2003); Lyu et al. (2014); Ren et al. (2015); Zhu et al. (2018)). These observations are consistent with crystallisation soaks with lauric acid performed previously, where no density supporting the shorter chain lauric acid was observed (Roedig et al., 2017). It is conceivable that this extended P-site confers a greater capsid stabilising effect on EV-F3, compared to EV-E1. During uncoating, expulsion of the hydrophobic pocket factor appears to be a crucial step for most enteroviruses and thus the higher energy penalty associated with breaking more extended hydrophobic contacts at EV-F3's P-site would likely delay this process.

Marked differences at the antigenic surface loops between EV-E1 and EV-F3, due to sequence differences modifying interactions with neighbouring residues, is consistent with the lack of cross-neutralisation from EV-E1 antibodies (Smyth et al., 1992) and supports the taxonomic separation of the two BEVs.

EV-F3 and EV-E1's unusually extended VP3 GH-loop at the capsid surface that encircles the five-fold axis, partly filling the canyon, distinguishes them from other enteroviruses. Since the canyon is often a site for receptor engagement (see Chapter 1), its truncation may have substantial implications for receptor binding, if receptors engage in this region. Indeed, this smaller depression provides more limited contact points with potential receptors, which may instead contact more surface-exposed regions susceptible to antibody recognition. Structures of EV-F3 with a putative receptor are presented in the next Chapter.

Close similarities between EV-F3's natural empty particle and mature EV-F3 are quite consistent with observations for PV-1, excepting the location of the scissile bond ((Basavappa et al., 1994). By contrast to PV-1, the VP1 N-terminus does not appear to clash with the modelled VP0 chain, and it is possible that the absence of stabilising contacts with the VP1 N-terminus in PV-1 drives a major population with the scissile bond in place for cleavage. For EV-F3, disordering of this region

is most likely due to heterogeneity in VP0 positioning. It is very unlikely that we are seeing an averaging effect from co-crystallisation of empty virions and contaminating mature virions, since samples should have been well separated during sucrose gradient purification, and the presence of 5-fold NCS means that maps are relatively unbiased.

Also akin to other enteroviruses, such as EV-A71 and PV-1, EV-F3 is susceptible to global stimulation to uncoat by heat (Shingler et al., 2013; Levy et al., 2010; Shingler et al., 2013; Organtini et al., 2014). Careful classification and refinement of picked particles enabled the isolation of three uncoating intermediates. For the majority of reported enterovirus globally-stimulated uncoating intermediates, four key features have been identified: externalisation of the VP1 N-terminus; loss of VP4; exit of the pocket factor and expansion by ca. 4 % with channels at the 2-fold symmetry axes (Ren et al., 2013; Butan et al., 2014; Organtini et al., 2014; Shingler et al., 2013). All four of these features have been observed for our particles A-I-B, albeit with resolutions potentially too poor to resolve pocket factor density directly and expulsion instead must be inferred from the substantial movement of VP1 strands constricting the P-site area.

Differences in reconstructions for particles A-I-B with respect to the presence or absence of density in the capsid interior are consistent with a A or 135S, where the genome is present but the capsid expanded, B or 80S where the genomic material has been expelled and A/B intermediate form where the genomic material appears more condensed around the particle centre. We also see evidence of A, I, B representing a continuum along the uncoating pathway with respect to the VP3 GH-loop, which can be seen to slowly relax into a hairpin-like structure (similar to observations with CV-A16 (Ren et al., 2013)), while density around the region usually occupied by the VP2 hairpin that resides below the 2-fold helices is much clearer for particle A, despite its poorer resolution. Of course, in reality, these three classes likely represent averages from portions of a continuum of intermediate particle forms, with icosahedral averaging during refinement blurring out any asymmetry.

Chapter 4

EV-F3: Co-factor and ligand interactions

Understanding the basis of enterovirus capsid stability is key for vaccine and antiviral development, an important theme of the work described in this thesis. Within this Chapter, I explore ligands and co-factors that have been hypothesised to interact, and stabilise enterovirus capsids, using EV-F3 as a model.

Metal ions are important for picornavirus capsid integrity, but their exact role, identity, and position within viral capsids is ill-defined (Chapter 1, Section 1.1.5). Previous EV-F3 models have included K^+ ions; three per icosahedral asymmetric unit, distributed across the capsid interior (Roedig et al., 2017). However, the placement and selection of ions was speculative, and their role in capsid stability unknown.

As described in Chapter 1 Section 1.3.1, the importance of glutathione for the morphogenesis of some enteroviruses has been well documented (Smith and Dawson, 2006; Thibaut et al., 2014; Ma et al., 2014; Abdelnabi et al., 2017). A patch of density at the protomer interface has been observed for a number of enterovirus structures, including mammalian cell expressed PV VLPs (personal communication, Mohammad Bahar), EV-A71, and CV-A16 (personal communication, Jingshan Ren) and in the XFEL EV-F3 structure (PDB: 5OSN) (Roedig et al., 2017), a glutamine residue was modelled in this site. However, GSH inhibition studies show that some of these viruses, such as EV-A71, are GSH-insensitive (Thibaut et al., 2014). These results collectively led us to speculate that this apparently conserved binding surface, herein termed the G-site, might bind GSH during assembly.

Finally, EV-F3's receptor has remained a long-standing mystery, in part due to EV-F3's elusive

tissue tropism. Some enteroviruses have been shown to hold a preference for glycan-containing receptors (Zocher et al., 2014), and Stoner et al. (1973) showed that neuraminidase treatment altered cellular susceptibility to EV-F1 (BEV261). This led to the crystallographic visualisation of EV-F3 bound to 3' sialyllactose (Elizabeth Fry, Ren Jingshan, Thomas Walter, unpublished).

All EV-F3 material used for the following studies was grown and purified as outlined in Chapter 2, Section 2.2.2.

4.1 Understanding Capsid Stability: Metal ion Binding sites

4.1.1 Assessing the Thermostability of EV-F3 Capsids

As a simple test to see the effect of metal ions on EV-F3 thermostability, a PaSTRy assay (Walter et al., 2012), detecting release of genomic material with a SYTO9 (Invitrogen, see Chapter 2, Section 2.7) reporter dye, was conducted using a panel of metal ion additives (Mg^{2+}, K^{+}, Ca^{2+}, Ni^{2+}, Cu^{2+} and Mn^{2+}) in over 100-fold excess of putative binding sites (i.e. assuming four per ASU), in triplicate. EDTA was also tested to see if depletion of divalent metal ions by chelation could destabilise EV-F3's capsid. The hydrophobic dye, SYPRO Orange (Invitrogen) was not used in parallel to detect exposure of hydrophobic environments, because its signal is artificially enhanced by interactions with EDTA (Kroeger et al., 2017).

MES (2-(N-morpholino)ethanesulphonic acid) was also assayed as part of a separate line of enquiry. MOPS (3-(N-morpholino)propanesulphonic acid)), and its shorter analogue MES, comprise a morpholino group linked to sulphoxide moiety via a three or two carbon linker, respectively. $TCID_{50}$ measurements have shown that CV-B3 infectivity was reduced upon incubation with MOPS in a concentration-dependent manner, at pHs 7.0-7.5 (Carson et al., 2017). The authors hypothesised, based on computer-simulated docking experiments focussed on the P-site, that MOPS may confer thermostability by displacing the 'natural' P-factor and forming H-bonds with the side chains of residues R1095 and N1211 (EV-F3 has a Gly and an Asn in these positions) (Carson et al., 2017). We therefore investigated whether a similar such effect could be seen for EV-F3.

The results show a slight increase in calculated melting temperature with the addition of EDTA, MES, $MgCl_2$, $MnCl_2$ and KCl, by approximately 1 °C, and this appears to be statistically significant, based on upper and lower 95 % confidence intervals (Table 4.1, Figure 4.1). This observed increase

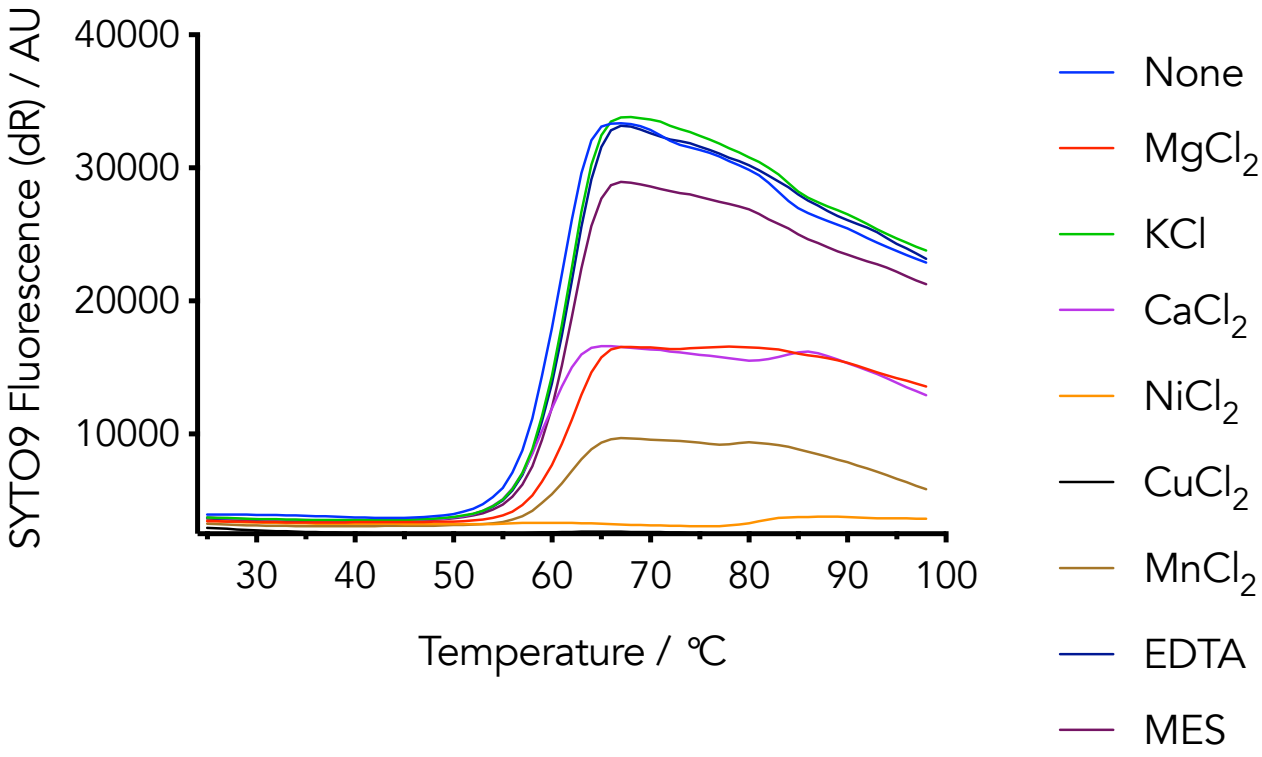


Figure 4.1: First derivative trace of metal ions thermofluor assay coloured by additives, see key on the right. Melting temperatures can be found in the accompanying table 4.1, below.

Table 4.1: 5-parameter Boltzmann fit of thermofluor data shown in Figure 4.1, with 95 % confidence intervals calculated by GraphPad Prism. The difference in measured melting temperature, relative to EV-F3 native with no additives (i.e. none) is shown in the final row.

Condition [10 μM]	None	$MgCl_2$	KCl	$CaCl_2$	$NiCl_2$	$CuCl_2$	$MnCl_2$	EDTA	MES
Tm / °C	60.2	61.4	61.1	59.0	~54.2	58.0	61.1	61.2	61.3
Upper	60.6	61.5	61.3	59.4	-	60.4	61.6	61.8	61.8
Lower	59.8	61.3	61.0	58.5	-	55.9	60.6	60.7	60.8
R^2	0.98	1.00	1.00	0.98	-	0.58	0.96	0.96	0.96
ΔTm / °C	0.0	1.3	1.0	-1.2	-	-2.2	1.0	1.1	1.2

in melting temperature implies a slight thermostabilisation of the EV-F3 capsid. Notably, a slight decrease in recorded melting temperature, by 1.2 °C, was observed on addition of $CaCl_2$, the opposite effect to that reported for aphthoviruses and PV A-particles (Tuthill et al., 2006, 2009). The $CuCl_2$ and $NiCl_2$ results should not be trusted, since both ions appear to quench the signal (this is reflected in the data shown in Table 4.1 and Figure 4.1; for instance, the poor R^2 for the $CuCl_2$ and inability to reliably fit the $NiCl_2$ data).

4.1.2 Anomalous Data Collection

Identification and location of metal ions in three-dimensional protein structures is arguably best achieved using anomalous data, as discussed in Chapter 2, Section 2.4.5. K^+ ions have been modelled for EV-F3 (Roedig et al., 2017), and are known to play a role in mediating RNA tertiary structural arrangement (Conn et al., 2002; Shiman and Draper, 2000; Rozov et al., 2019).

Calcium ions have been modelled in a number of picornavirus structures (see Chapter 1, Section 1.1.5), but, as shown above, may have a very slight destabilising effect on EV-F3 capsids during heating.

To determine (i) whether Ca^{2+} ions may occupy binding sites within EV-F3's capsid and (ii) unambiguously determine if the potassium sites previously modelled for EV-F3 are present, anomalous data were collected from crystals of EV-F3 soaked with $CaCl_2$.

Crystal Mounting

EV-F3 crystals were soaked with a glycerol cryoprotectant (see Chapter 2, Section 2.4.2) and, for some crystals, 2 mM $CaCl_2$ for ca. 5 min prior to cryo-cooling. It was necessary to mount crystals with bespoke loops suited for the system used at the I23 X-ray beamline (Diamond Light Source, UK, personal communication, Dr. Ramona Duman).

Emission Spectra

To first establish whether any K^+ or Ca^{2+} ions may be present in unsoaked EV-F3, emission was measured on a fluorescence detector following application of a 7 keV incoming beam to excite potassium for one EV-F3 crystal that had not been soaked with 2 mM $CaCl_2$.

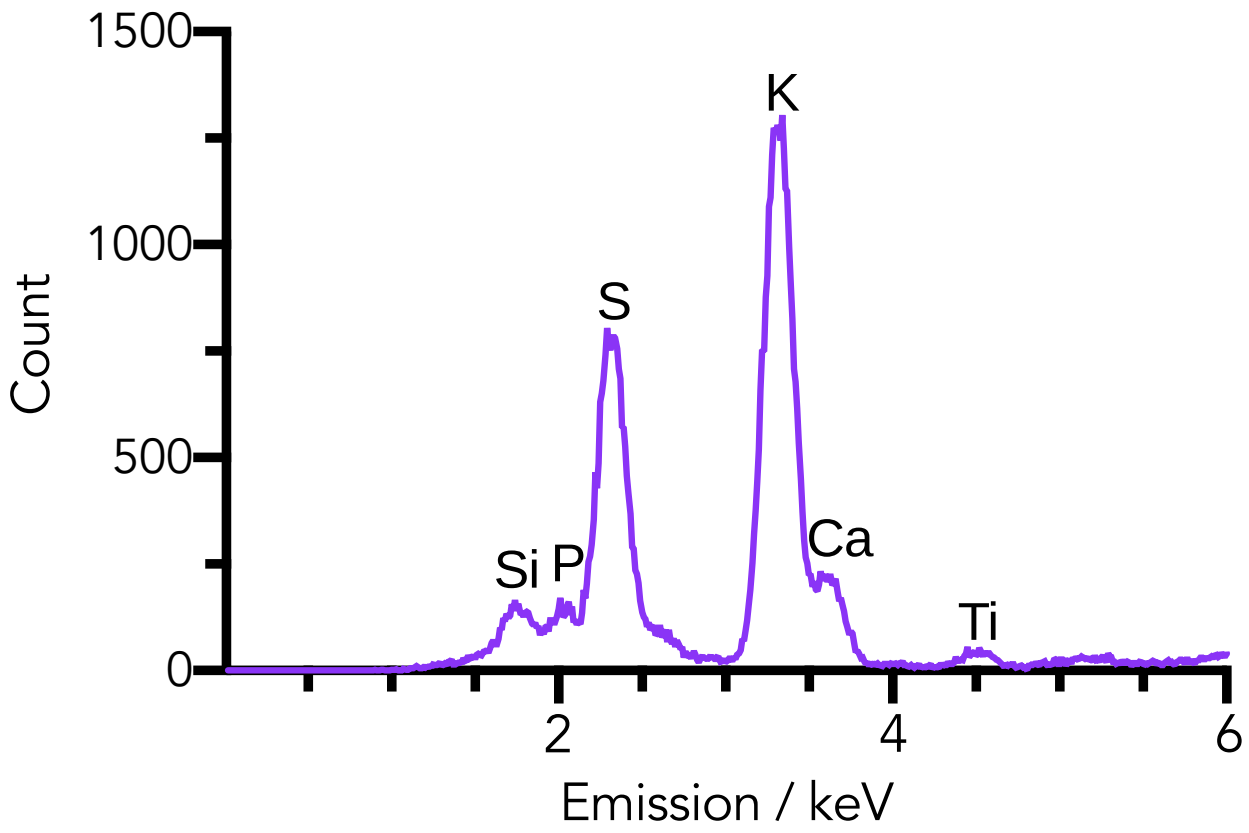


Figure 4.2: Fluorescence emission signal following excitation at 7 keV. Candidate elements for each peak are marked.

If calcium were present, we would anticipate a high emission peak at ca. 3.7 keV. In fact, the predominant emission peaks occurred for potassium (3.3 keV) and sulphur (2.3 keV), with a shoulder of the potassium peak that might be a very weak calcium signal (Figure 4.2). The experiment was not repeated on $CaCl_2$ soaked crystals, since the solvent signal would dominate, although back soaking could have been attempted (i.e. soak in $CaCl_2$ buffer for an allotted time, before soaking in a $CaCl_2$ free buffer), to allow transiently associated calcium ions to diffuse back into the bulk solvent.

X-ray Diffraction Data Collection at the Calcium and Potassium K-edges

Diffraction data were collected from EV-F3 crystals at four wavelengths: (i) Above the calcium K-edge, (ii) Below the calcium K-edge; (iii) Above the potassium K-edge; (iv) Below the potassium K-edge at the I23 X-ray beamline (Diamond Light Source). Associated K-edges and theoretical f"

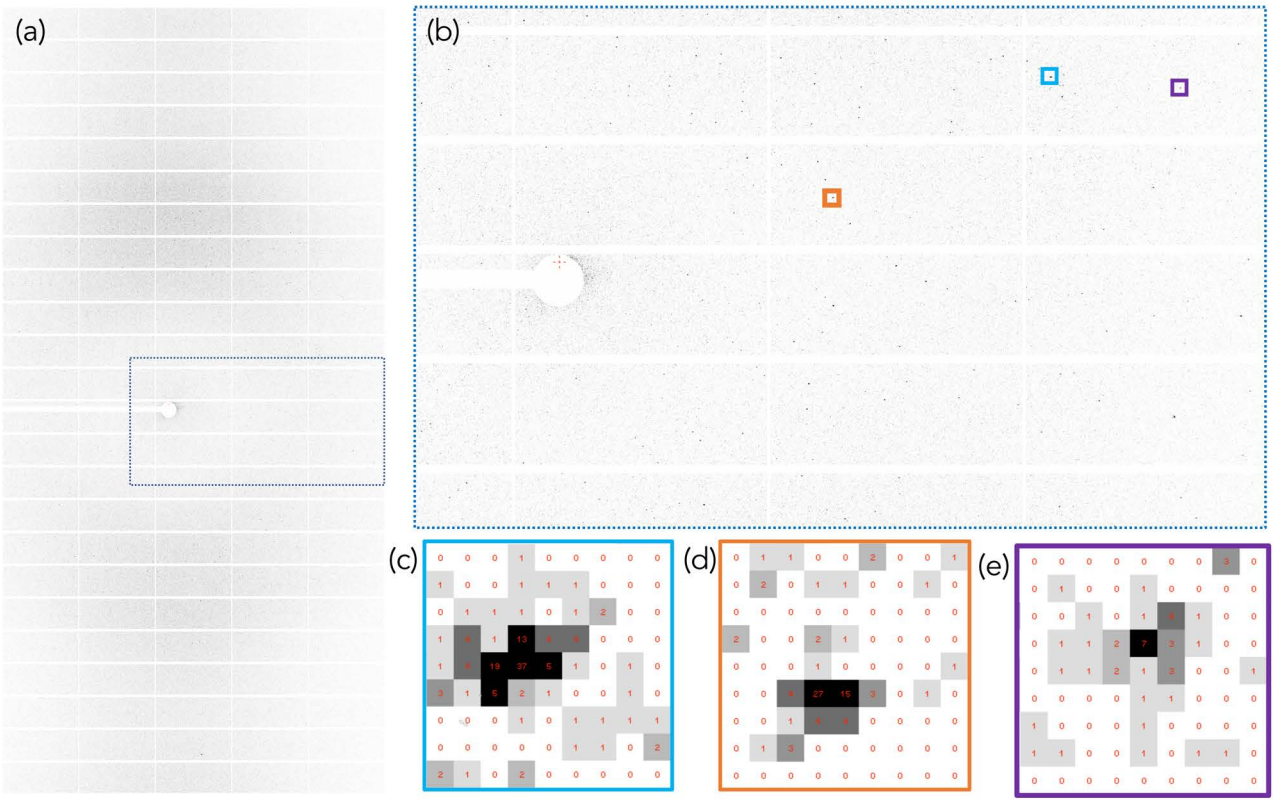


Figure 4.3: Example diffraction from data collected at 3.06 Å wavelength. (a) View of the entire detector (note that the ellipsoidal solvent ring is due to the curvature of the detector). The rotation axis is horizontal in this view. (b) Zoomed in region, where some faint lunes may be seen. (c-e) Example diffraction peaks, with counts, visualised with Adxv (*Adxv – A Program to Display X-ray Diffraction Images, http://www.scripps.edu/ arvai/adxv.html*, 2013)

values for potassium, calcium and sulphur are shown in Table 4.2 and Figure 4.4.

To ensure maximal signal (essential for weakly diffracting virus crystals), the beam was focussed. For the calcium K-edge data collection, the flux for an unfocussed beam was approximately 1x10^{10} for a 50 x 50 μm beam, whilst, when focussed, even at 10 % transmission, we had 3 times more photons, and, with 100 % transmission, a flux of 3x10^{11}. The only major caveat is that focussing can introduce significant divergence which can inflate spot size on the detector. Nevertheless, diffraction quality was good, with almost zero background due to the I23 vacuum endstation (necessary to avoid absorbance at the longer wavelengths used), Figure 4.3.

For data collected below and above the calcium K-edge (4.00 keV and 4.05 keV), crystals diffracted to ca. 2.5 Å resolution (Table 4.3). The datasets collected above and below the potassium K-edge, at 3.65 keV (λ = 3.35 Å) and 3.50 keV (λ = 3.54 Å), were from the same crystal, which

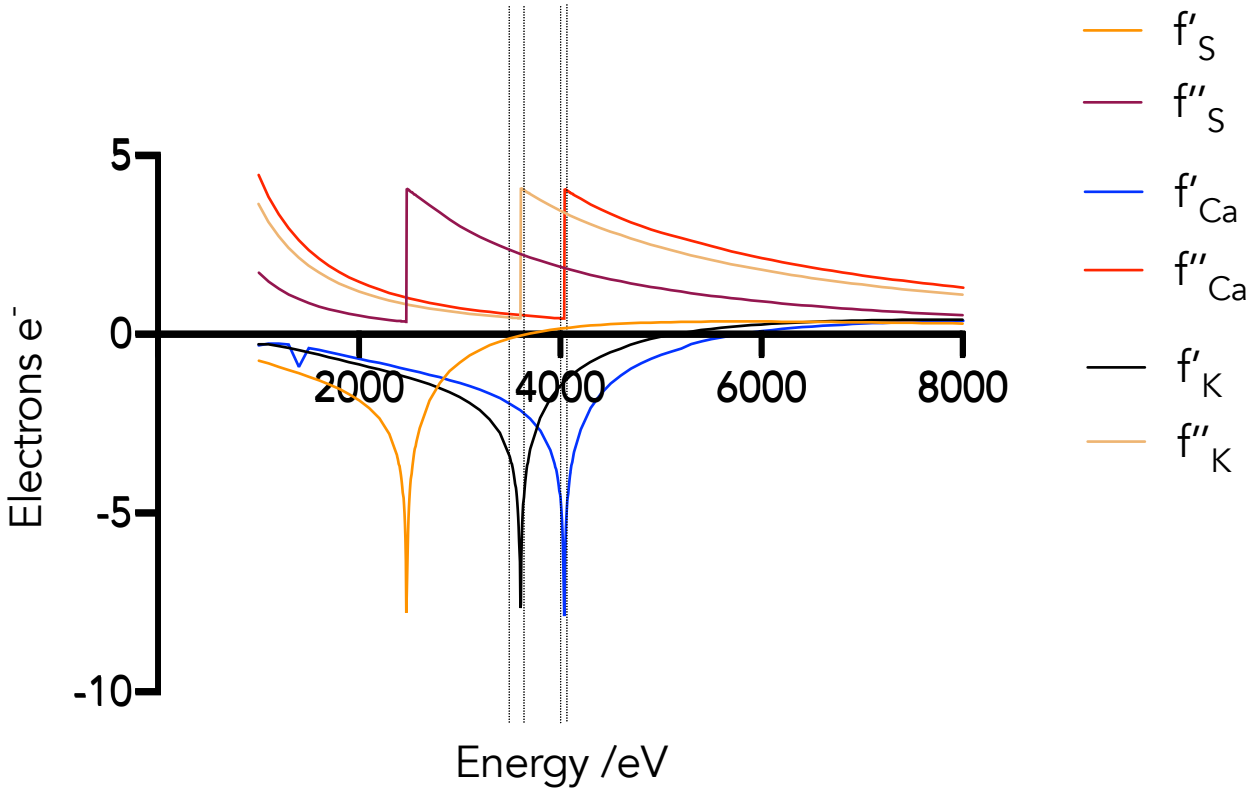


Figure 4.4: Sulphur, potassium and calcium absorption edges. Note that this is purely theoretical and does not take into account influences of neighbouring atoms (values are very sensitive to environmental changes around the absorption edge) (Brennan and Cowan, 1992). The four wavelengths probed for our experiments (3.50, 3.65, 4.00, 4.05 keV) are marked on each graph, as per the key.

diffracted to ca. 3.0 Å resolution. All datasets were found to be of *F*23 space group (see Chapter 2, Section 2.4.4 for processing details). An existing EV-F3 model from our sialyllactose-soak experiments (Section 4.1.5 below) was then rigid-body-fitted into each map with sequential resolution cut-offs before further positional and B-factor refinement with five-fold NCS constraints in CNS (Brunger, 2007).

Comparison of Anomalous Peaks

Refined models were used to determine anomalous densities with ANODE (Thorn and Sheldrick, 2011). Table 4.4 lists the top ten peaks in each of the four ANODE maps (note this map was not NCS averaged) (Thorn and Sheldrick, 2011). It is immediately apparent that, for all energies above the potassium K-edge, this list includes peaks ca. 2.2-2.8 Å away from oxygen atoms in the capsid

Table 4.2: Theoretical f′ and f″ for wavelengths used in data collections, taken from Brennan and Cowan (1992)

Element		Ca		K		S	
Wavelength / Å	Wavelength / eV	f′	f″	f′	f″	f′	f″
3.54	3500	-1.95	0.55	-3.44	0.44	-0.81	2.34
3.40	3650	-2.23	0.50	-3.98	4.01	-0.07	2.18
3.10	4000	-4.61	0.44	-1.47	3.39	0.15	1.91
3.06	4050	-5.75	4.03	-1.33	3.31	0.15	1.95

Table 4.3: Processing statistics from XSCALE, XDS, followed by CNS positional and B-factor refinement of data collected from I23 beamline (Diamond Light Source, UK) (Brunger, 2007; Kabsch, 2010*b*).

Wavelength / Å	3.06	3.10	3.40	3.54
Wavelength / eV	4050	4000	3650	3500
Data Collection				
Space group	*F*23	*F*23	*F*23	*F*23
Cell dimensions				
a = b = c / Å	431.8	432.2	435.9	436.5
$\alpha = \beta = \gamma$ / °	90	90	90	90
Resolution / Å	152-2.80 (2.85-2.80)	250-2.5 (2.54-2.50)	131.6-3.0 (3.05-3.00)	131.6-3.0 (3.05-3.00)
Rmeas	0.302 (–)	0.211 (–)	0.362 (–)	0.425 (–)
Rpim	0.067 (0.493)	0.097 (0.945)	0.096 (0.720)	0.156 (–)
I / σI	10.8 (1.5)	6.7 (0.8)	7.9 (1.3)	4.7 (0.7)
CC0.5 / %	99.2 (70.2)	98.0 (30.9)	98.9 (45.1)	97.4 (20.7)
Completeness / %	100 (100)	99.6 (93.5)	99.9 (97.1)	99.7 (94.7)
Multiplicity	21 (17.1)	4.8 (3.2)	15.6 (11.5)	7.5 (4.9)
Refinement				
Resolution / Å	50.0-3.0	50.0-3.0	50.0-3.0	50.0-3.0
No. reflections	154111 (8289)	216911 (11258)	129072 (6897)	129359 (6538)
No. of protein atoms	6331	6409	6138	6138
No. of ligand/ions	30	30	30	30
Rwork / Rfree	18.1/18.7	20.6/21.6	19.1/19.5	20.4/21.4
Ramachandran plot / %				
Favoured regions	91.1	91.9	91.9	92.3
Allowed regions	8.35	7.58	7.46	7.58
Outliers regions	0.51	0.51	0.64	0.13
R.m.s. deviations				
Bond lengths / Å	0.006	0.005	0.006	0.006
Bond angles / °	1.3	1.3	1.3	1.3

model consistent with signal from coordinating metal ions. Contrastingly, peaks detected in the dataset collected below the potassium K-edge (where only sulphur has a significant f″ signal, see Figure 4.4, Table 4.2), were exclusively at Met and Cys residues (Table 4.4).

To enhance the signal-to-noise, the ANODE derived anomalous density maps were 5-fold NCS averaged in Coot (Emsley and Cowtan, 2004; Emsley et al., 2010; Thorn and Sheldrick, 2011) and inspected for peaks distal from S atoms. Four notable peaks (labelled (a)-(d), see Figure 4.5) were found for all datasets except that collected below the potassium K-edge (3.5 keV), and were in agreement with the top four peak heights listed in the ANODE output (Thorn and Sheldrick, 2011).

Given that the sulphur f″ is reasonably similar between the wavelengths (1.91, 1.95, 2.18, 2.34 electrons at 4.00, 4.05, 3.65 and 3.50 keV, respectively, see Figure 4.4 and Table 4.2), to correct for variations in the map error levels and scaling (Figure 4.6), the 'base line' for contouring each NCS-averaged anomalous density map was set using the average of the peak heights for the two sulphur containing residues, M2231 and C2232 (which have similar B-factors across all datasets, Table 4.5), (see grey box, top left, Figure 4.5, Table 4.6). Note that a better approach, to account for environmental variations that can impact anomalous signal, would have been to use the average sigma level for all S atoms across the protomer. These baselines for data collected at 3.50, 3.65, 4.00, 4.05 keV were: 7.9, 9.4, 7.1 and 9.8 σ (Table 4.6). These variations do not reflect the size of the f″ signal, which is, for instance, 7 % higher at 3.50 keV than 3.65 keV (Table 4.2), but do indicate data quality; multiplicity was 4.8 at 3.5 keV vs 21 at 3.65 keV (Table 4.3).

The heights of each of the four major peaks were normalised using the appropriate sulphur values (Table 4.6). For all four sites, there is a major difference above and below the potassium edge, with no significant peaks below the edge for any of the putative metal ion sites. By contrast, there is strong density both above and below the calcium edge (Figure 4.5). Signal at site (c) (defined in Figure 4.5, and below) is significantly weaker than the other sites at all wavelengths tested, indicating partial occupancy (assuming that B-factors and occupancy are not conflated). Relative to the other datasets, the peak heights are a little lower than expected in the 4.00 keV dataset, which we cannot explain.

Table 4.4: Ten atoms with the highest unaveraged anomalous densities from ANODE (Thorn and Sheldrick, 2011). Note that residue numbers are shown without the VP number i.e. residue 120 of VP1 is shown as 120, and not 1120.

Wavelength / keV	Height / σ	Distance / Å	Atom:residue	VP
3.50	7.38	0.65	SG:CYS120	1
	7.10	0.96	SD:MET148	1
	7.08	0.25	SD:MET77	3
	6.73	0.31	SD:MET78	3
	6.63	0.75	SD:MET217	3
	6.60	0.84	CD:LYS193	2
	6.58	1.31	SD:MET148	1
	6.43	0.16	SD:MET61	2
	6.37	0.42	SD:MET203	3
3.65	15.98	2.76	O:ASN57	1
	13.50	2.68	O:SER42	1
	13.23	2.41	O:LEU33	1
	12.34	2.84	OE1:GLN222	3
	11.49	2.46	O:GLN53	2
	10.40	2.39	O:ALA65	4
	9.56	0.23	SD:MET217	3
	9.38	0.36	SG:CYS23	3
	9.13	0.36	SD:MET197	3
	8.99	0.41	SG:CYS158	1
	8.78	0.14	SG:CYS110	2
4.00	10.85	2.10	OD1:ASN17	1
	9.85	2.44	O:SER42	1
	9.19	1.93	OD1:ASN17	1
	9.12	2.39	O:ALA65	4
	9.07	0.76	SD:MET148	3
	8.98	0.55	SD:MET78	1
	8.90	0.23	SG:CYS110	2
	8.72	2.22	O:GLU63	4
	8.63	0.20	SD:MET197	3
	8.61	2.23	O:ALA65	4
	8.52	0.37	SD:MET137	1
4.05	13.73	2.43	O:ASN57	1
	11.54	2.94	OE1:GLN222	3
	11.18	2.64	O:LEU33	1
	10.74	0.59	CD:LYS193	2
	10.58	2.96	O:SER42	1
	10.47	0.55	SD:MET82	3
	10.45	0.101	SD:MET124	3
	10.11	0.95	SD:MET148	3
	9.91	0.17	SG:CYS265	1
	9.77	0.26	SD:MET197	1

Table 4.5: Average B-factors for modelled VP2 residues 231 and 232 for each K-edge dataset.

Wavelength / eV	Residue	Average Bfactor / Å^2
4050	M2231	30.1
	C2232	33.4
4000	M2231	30.1
	C2232	33.3
3650	M2231	30.2
	C2232	33.4
3500	M2231	30.3
	C2232	33.5

Table 4.6: Peak heights for K^+ ion sites. The σK / σS column shows the ratio between peak heights for each potassium site and the average peak height for two selected sulphur sites (residues M2231 and C2232), shown in the final row.

Wavelength / keV	4.05		4.00		3.65		3.50	
Site	σ	σK / σS	σ	σK / σS	σ	σK / σS	σ	σK / σS
(a)	16.4	1.7	11.8	1.7	19.9	2.1	2.7	0.3
(b)	14.5	1.5	8.3	1.2	18.1	1.9	–	–
(c)	9.8	1.0	7.8	1.1	14.4	1.5	–	–
(d)	14.9	1.5	8.4	1.2	19.8	2.1	2.5	0.3
Σ[M2231, C2232] /2	9.8		7.1		9.4		7.9	

Using the peak heights listed in Table 4.6, occupancies for each K^+ ion were determined using the following procedure (Table 4.7):

1. Estimate fully occupied potassium peak height (σ_{K100}), based on sulphur σ (σ_S, used as a 'lock') and f''(S).

$$\sigma_{K100} = (\frac{f''(K)}{f''(S)}) \times \sigma_S$$

2. Find ratio between 'measured' ($\sigma Kmeas$) peak height for a given potassium ion and that of a fully occupied potassium peak height:

$$\%occupancy(K) = (\frac{\sigma Kmeas}{\sigma K100}) \times 100$$

Overall, this analysis provides strong evidence that all four sites harbour K^+ ions, and at ca. 100 % occupancy for sites (a, b, d).

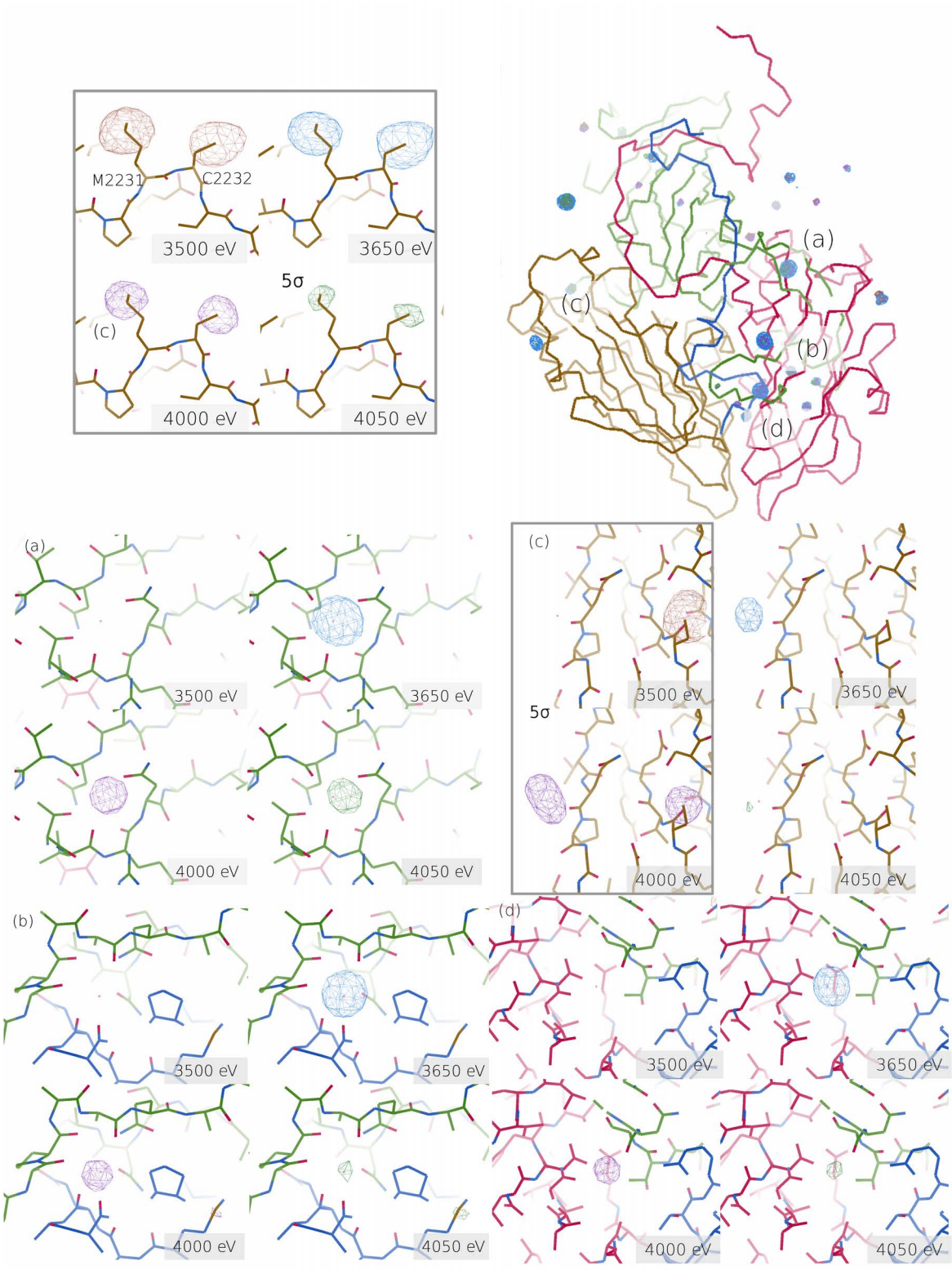


Figure 4.5: 5-fold averaged anomalous densities from EV-F3 crystals tested. Grey box, top: M and C densities set to 5 σ (labelled at the top left). Figure legend continued on next page.

Figure 4.5: Continued. Top right: anomalous sites normalised to the point at which no density could be seen for the residues shown on the left. Sites are labelled (a-d) and correspond to the figures below. Ribbon representation, coloured by chain. Bottom (a)-(d): four putative metal ion sites observed, based on anomalous density. All are contoured based on sigma levels described in the main text, and associated wavelengths are labelled. For (c), the density for 3.50 and 4.00 keV is at 5 σ (i.e. at a lower level than the other figures) to demonstrate that there is a small amount of signal in this region. Protein is in stick representation (colours: VP1 green, VP2 pink, VP3 brown, VP4 blue, and angled to correspond to the interaction sites shown in Figure 4.9.

Table 4.7: Estimated percentage occupancy for potassium sites (a-d), using the procedure described in the main text. Note that this assumes the B-factor of each potassium ion to be equivalent to that of M2231 and C2232 (Table 4.5), and that anomalous electrons are similarly distributed for both potassium and sulphur atoms, coming from the each respective atom's innermost shell. Occupancies were estimated as 0 % for all sites for the 3.5 keV and were thus omitted. Clearly, occupancies have been 'over-estimated' for the 3.65 keV sites.

	4.05 keV	4.00 keV	3.65 keV
f"(S)	1.95	1.91	2.18
f"(K)	3.31	3.39	4.01
(a)	98.6	93.6	115.7
(b)	87.2	65.9	105.2
(c)	58.9	61.9	83.7
(d)	89.6	66.7	115.1

Additional, weaker, peaks were observed at the icosahedral 5-fold axes ($> 4\ \sigma$ in the NCS-averaged anomalous density maps, see Figure 4.7). Signal at the top portion of the channel was very clear at 3.65 and 4.05 keV, probably mainly due to superior data quality. Density at the bottom portion of the 5-fold axis channel seemed to be strongest at the energy chosen to maximise the f" for calcium (Table 4.2 and Figure 4.7). Therefore, some of the soaked calcium might occupy the anionic five-fold channel. Since all crystals were soaked with calcium, it is plausible that some ions could transiently reside in the anionic 5-fold channel; the weak signal could indicate low occupancy. However, the calcium anomalous signal should be far greater at the highest energy (see Table 4.2, 4 electrons per calcium atom at 4.05 keV; less than 0.6 electrons for all other wavelengths tested), which is not what we see; there is also a stronger signal for 3.65 keV, by contrast to 3.5 keV, where the peak in this position is much weaker Figure 4.7). Instead, the appearance of these 5-fold channel peaks correlates with data quality metrics, such as multiplicity: 21, 4.8, 15.6, 7.5 at 4.05, 4.0, 3.65, 3.50 keV, respectively. These observations may be an artefact from averaging at the symmetry axis, where errors tend to build up.

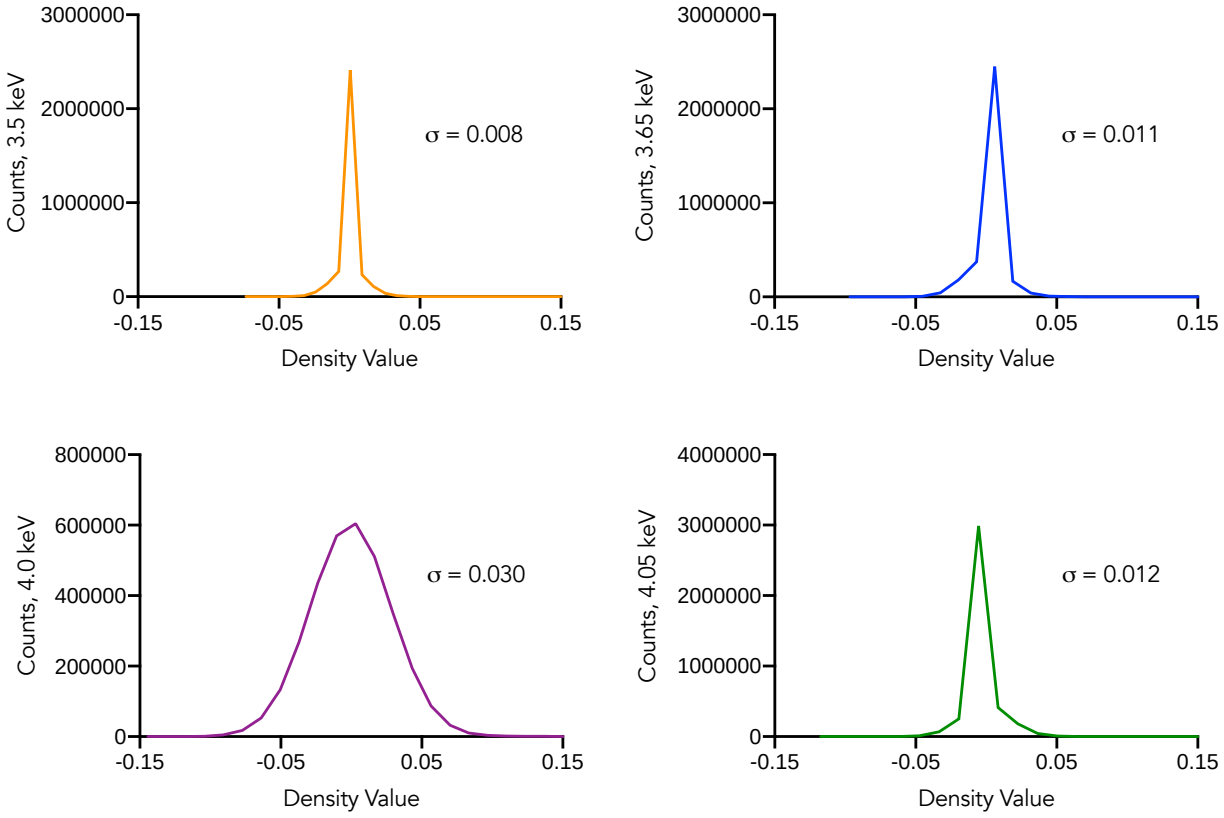


Figure 4.6: Histograms and associated standard deviations showing the scaling for each NCS-averaged anomalous density map, as determined by Coot (Emsley and Cowtan, 2004; Emsley et al., 2010).

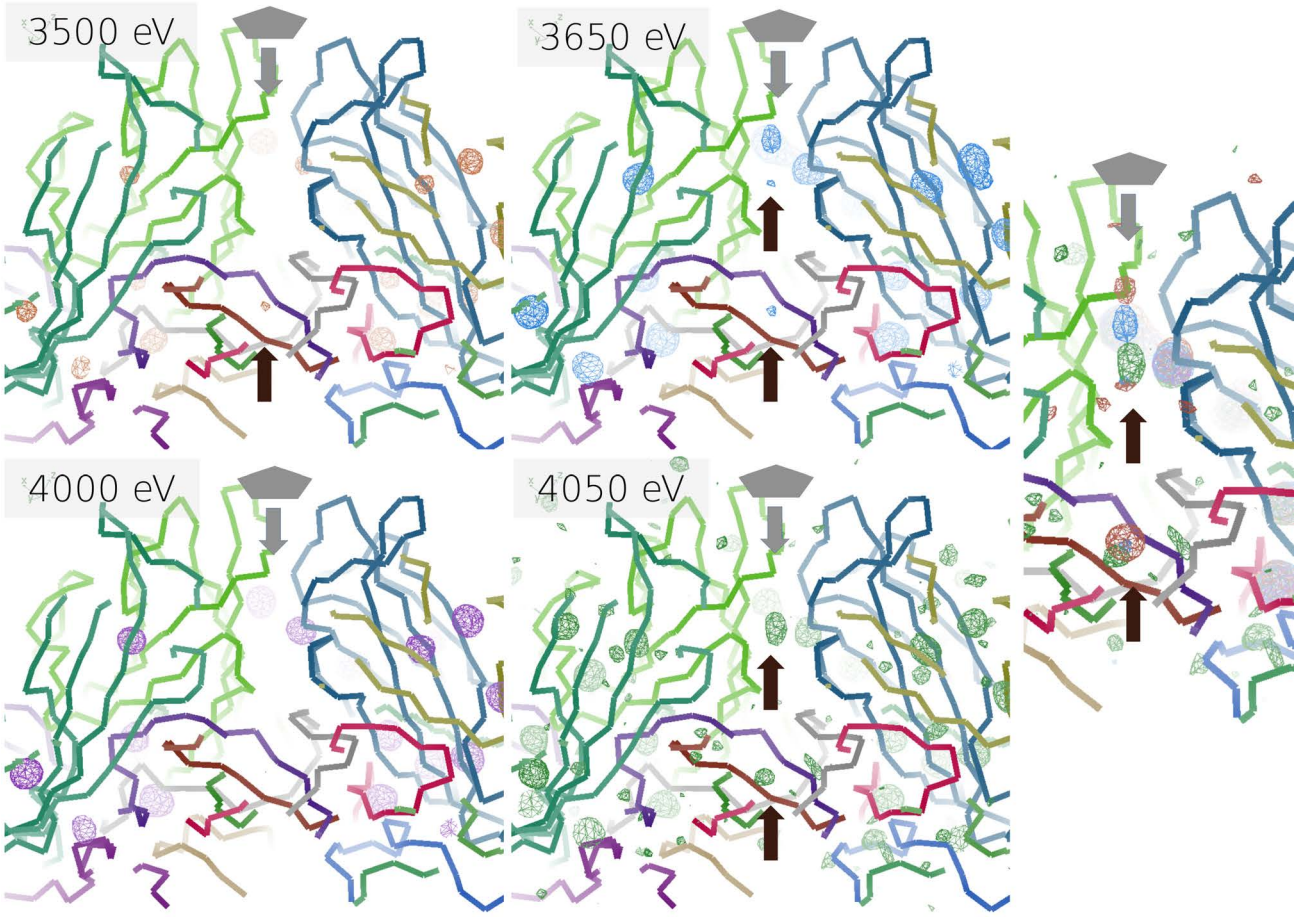


Figure 4.7: Slice through the five-fold channel (shown in ribbon representation), with 5-fold NCS-averaged anomalous peaks, contoured to 40 % of the lock sigma level (see main text). Black arrows indicate potential density of interest for each wavelength. The top of the 5-fold axis and position of the channel is indicated by a grey pentamer and arrow. The panel on the far right shows all the densities overlaid at 5 σ.

The absence of strong density at the lowermost portion of the 5-fold channel casts doubt on identity, or modelled occupancy, of the OSF molecule modelled at the 5-fold axis channel, whose sulphoxide moiety should show strong anomalous signal at all of the wavelengths tested (see Chapter 3, Section 3.2.2).

Significance of K-site Locations

All four identified K^+ sites reside on the interior of the capsid, and thus may interact directly with encapsulated RNA. Sites (b) and (d) radiate from the 3-fold axis, roughly following the interface between icosahedral protomers, parallel to VP4. Site (a) is closer to the off-axis channel from

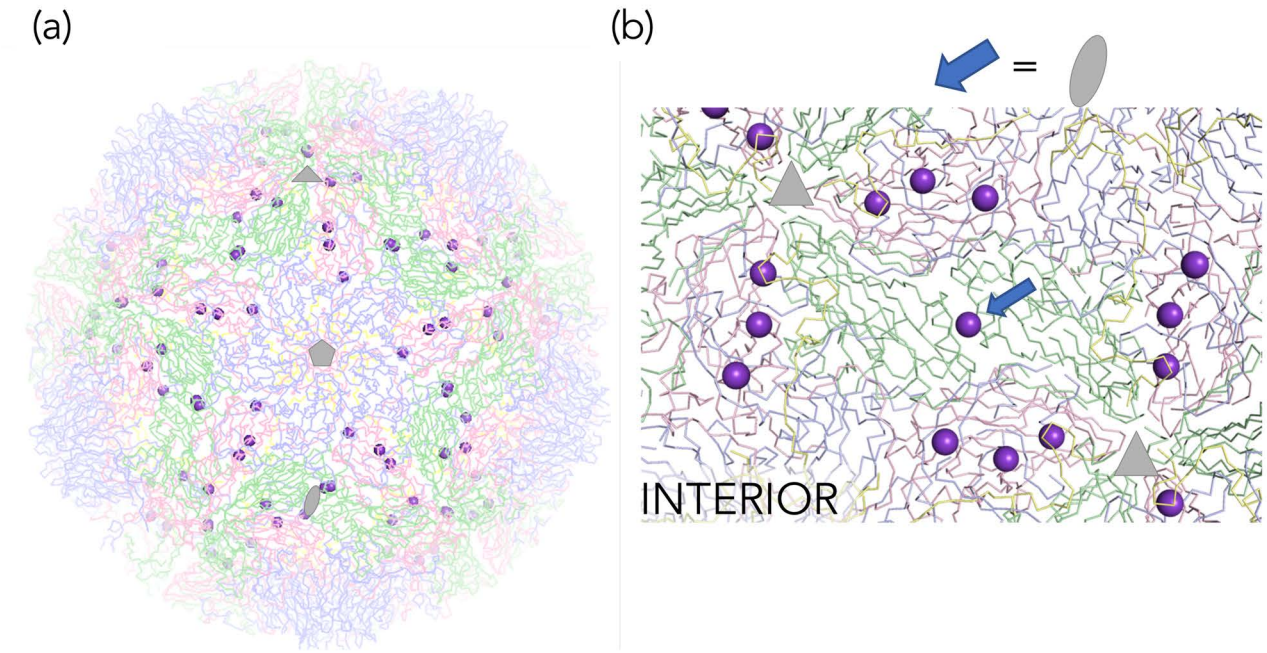


Figure 4.8: Overview of K^+ ion positions (purple spheres) in EV-F3's capsid. VPs are in ribbon representation, coloured by chain (blue, green, pink, yellow for VP1-4, respectively). Example 2-, 3- and 5-fold axes are marked with grey symbols. (a) in the context of the full virion, facing the icosahedral five-fold axis. (b) Interior of capsid, zoomed in on the two-fold axis (marked with a blue arrow). VPs and K^+ ions are represented as for (a) and the three-fold symmetry axes are marked with grey triangles. Ions act as a gate keeper in the vicinity of the icosahedral two-fold, where VP1 and possibly RNA egress occurs

which VP1 exits in CV-A16 (Ren et al., 2015) whilst site-(c) is located directly below the 2-fold interface (Figure 4.8(a,b)). For (a), the first liganding residue, T1014, is replaced by L1019 in EV-E1, suggesting that the site may not be conserved (Figure 4.8(a)).

For site (b), the modelled ion is contacted by residues E4063 and A4065 (see Figure 4.9(b)). Perhaps significantly, this modelled K^+ bridges the main chain carboxyl groups of residues T1030, L1033 and N1043, located within the highly conserved PAL-motif (see Chapter 3, Section 3.2.2), to VP4 C-terminal residues E4063 and A4064 (see Figure 4.9(b)).

At the 2-fold axes (site (c)), water molecules sit on either side of the modelled ion, while main chain Pro carbonyl atoms from the VP2 hairpins of two adjacent protomers that rest below the 2-fold helices, ligate the potassium on either side (Figure 4.9(c)). The lower number of coordinating residues holding this ion in place may account for the lower occupancy of this position, as evidenced by the reduced anomalous density, Table 4.6.

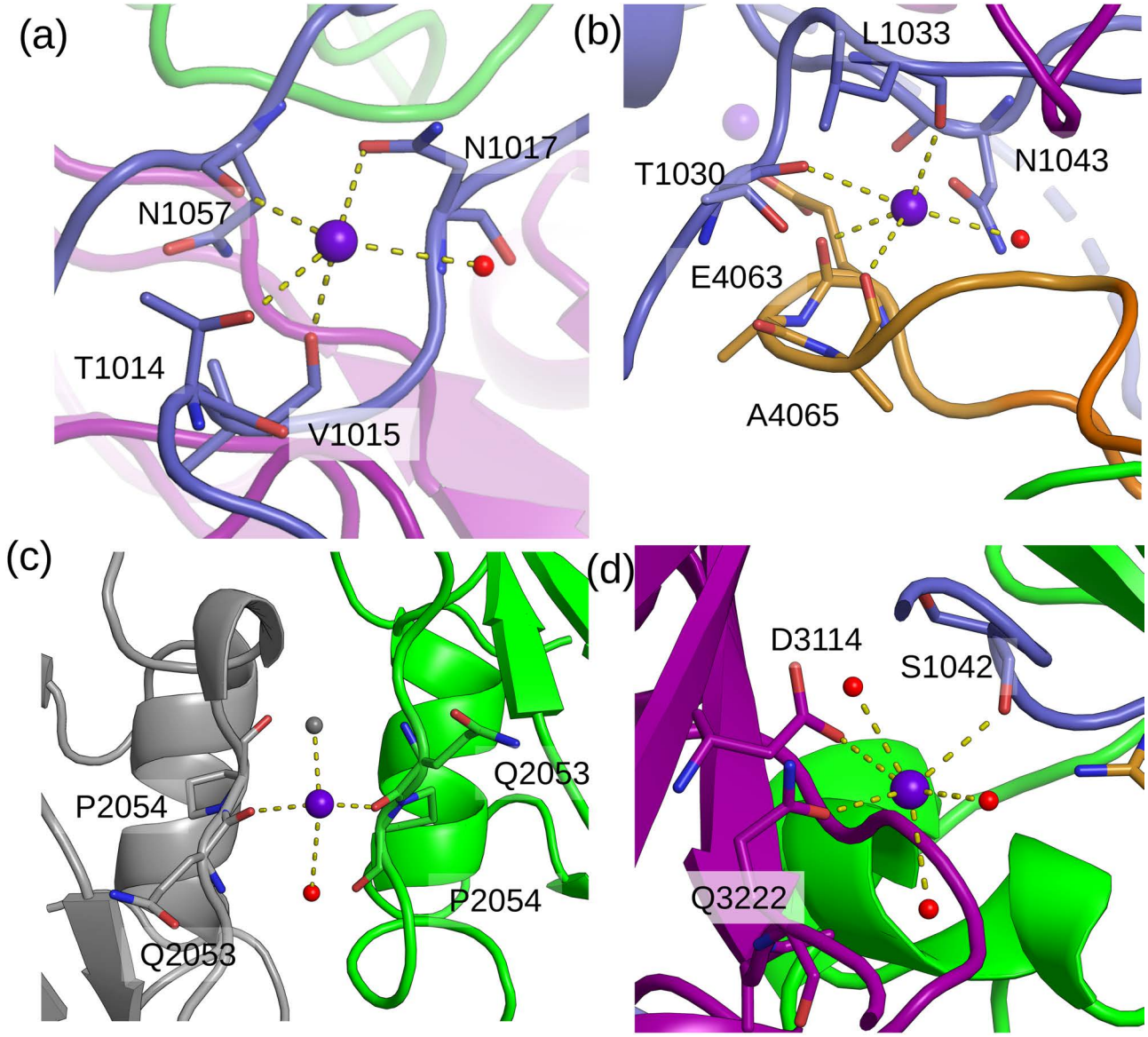


Figure 4.9: Cartoon of the K^+ ion interaction in the four strongest signal sites, identified in ANODE. K—O distances are all in the region of 2.5-3.6 Å. Residues are labelled by convention. For clarity, neighbouring protomers are coloured in grey. Chains are coloured by VP: VP1, blue; VP2, green; VP3, magenta; VP4, orange-yellow. K^+ ions are displayed as purple spheres, water molecules as red spheres. Respective closest coordination geometries (assuming no contribution by encapsidated genome and correct modelling of waters) are as follows: square pyramidal, trigonal bipyramidal, tetrahedral, trigonal prismatic.

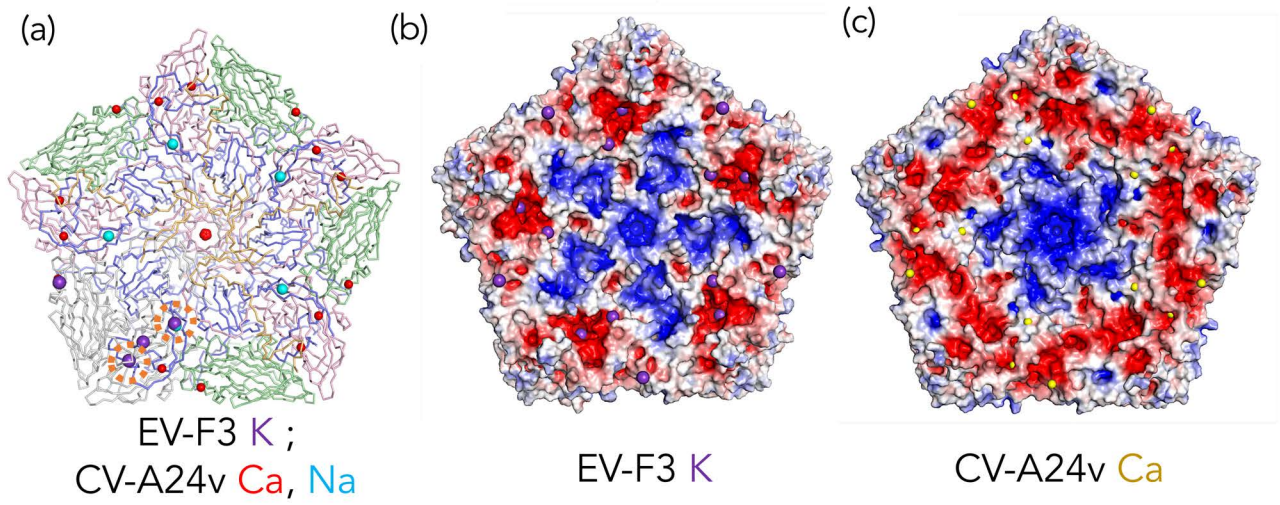


Figure 4.10: (a) Ribbon representation of a pentamer of CV-A24v PDB: 4Q4X (Zocher et al., 2014), with four of the protomers coloured by VP (blue, green, pink, yellow for VP1, 2, 3, 4, respectively). Calcium ions and sodium ions modelled for CV-A24v are shown as red and cyan spheres, respectively. EV-F3 potassium ion sites are mapped on to the bottom left protomer (grey), and sites that overlap with modelled CV-A24v calcium and sodium ions are circled in orange. (b) Electrostatic surface (+/- 5) of the interior of an EV-F3 pentamer, demonstrating the stabilisation of negatively charged portions of the capsid. K^+ ions are shown in purple and reside in the negative (red) patches. (c) Electrostatic surface of the interior of CV-A24v pentamer (PDB: 4Q4X, Zocher et al. (2014)), coloured as (c), with Ca^{2+} modelled in the PDB structure shown as yellow spheres.

Comparison with other Picornavirus structures

Sites (a-d) map onto some putative metal ion sites from other picornaviruses, most notably that of CV-A24v (Zocher et al., 2014), see Figure 4.10(a), and in both cases, reside exclusively in negatively charged patches Figure 4.10(b,c).

Sites (a) and (b) have also been modelled as K^+ for EV-A71 (PDB: 3VBH) (Wang et al., 2012), as shown in Figure 4.11.

Additional Features: A mis-modelled residue and the G-site

As an interesting aside to the metal ion story, clear density was observed at residue K2193 (see Figure 4.13) for all data collections strongly supporting a K2193M mutation i.e. a mutation at a single base (Met: ATG ; Lys: AAA/AAG).

The implications of the absence of any anomalous density, shown in Figure 4.13(a) at the glutathione binding site identified in section 4.2.8 will be discussed below.

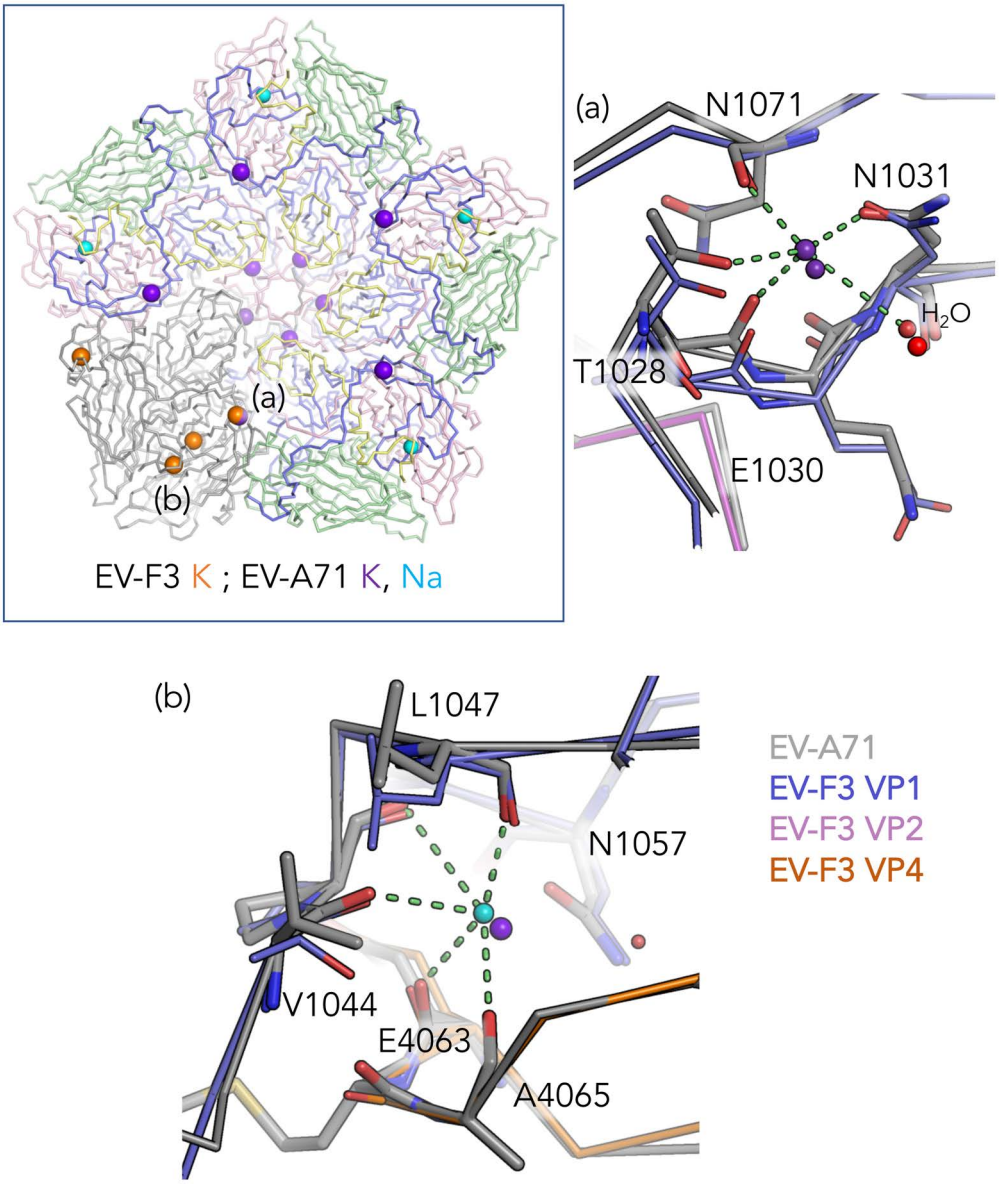


Figure 4.11: Blue box: Ribbon representation of a pentamer of EV-A71 PDB: 3VBH (Wang et al., 2012), with four of the protomers coloured by VP (blue, green, pink, yellow for VP1, 2, 3, 4, respectively). Potassium ions and sodium ions modelled for EV-A71 are shown as purple and cyan spheres, respectively. EV-F3 potassium ion sites are mapped on to the bottom left protomer (grey) as orange spheres for clarity. (a,b) Interactions between EV-A71 and its potassium ion sites '(a)' and '(b)', respectively (PDB: 3VBH (Wang et al., 2012)). EV-F3 is overlaid and coloured by VP (blue for VP1, light pink for VP2, orange for VP4), while EV-A71 is shown in grey. Potassium ions and sodium ions are represented as purple and cyan spheres, respectively. EV-A71 residues that interact its modelled potassium ion are labelled in black text, and interactions are shown as green dashed lines. Note that, an interacting H_2O molecule (red sphere) is modelled in a similar place to EV-F3 for site (a), but not for (b). The views for (a) and (b) should match that of Figure 4.9 above).

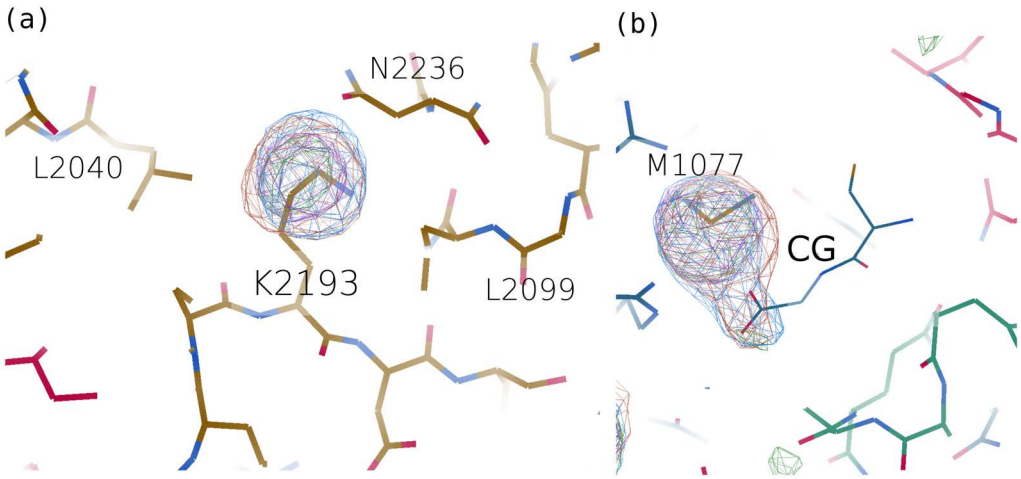


Figure 4.12: NCS averaged anomalous densities from all four data collections overlaid, contoured to 5 σ (a) Anomalous density at residue K2193, which might instead be a Met residue. (b) Anomalous densities at the G-site (see section 1.2.8). No signal can be seen for the S residue for the CG model presented in the next section. This could indicate that the position is incorrect, but the 2mFo-dFc and Fo-Fc density for GSH and CG was quite compelling (see below).

4.1.3 Glutathione mediates capsid stability at a druggable site

Does the G-site Exist?

To first establish whether GSH may have an affinity for EV-F3's capsid surface, EV-F3 crystals were soaked with 8 mM GSH (Sigma Aldrich) in addition to 5 mM DTT, with a glycerol cryoprotectant (see Chapter 2, Section 2.4.2).

Data were collected at I03 (Diamond Light Source, UK) using automated X-ray centring and data collection, and processed using the Xia2-Dials pipeline, as implemented on ISPyB (Delageniere et al., 2011; Waterman et al., 2013). By contrast to other experiments in this chapter, these crystals belonged to space group *I*222. Data from two crystals were merged to yield a 1.69 Å resolution structure. Following an initial rigid body fit of a previously refined EV-F3 structure, the model was subjected to positional and B-factor refinement, using CNS with strict 15-fold NCS constraints (Brunger, 2007). Model checking and rebuilding was performed using Coot, with 15-fold NCS averaged maps, as before (Emsley and Cowtan, 2004; Emsley et al., 2010). There was clear difference density in a depression towards the surface in the inter-protomer region consistent with a small molecule, Figure 4.13.

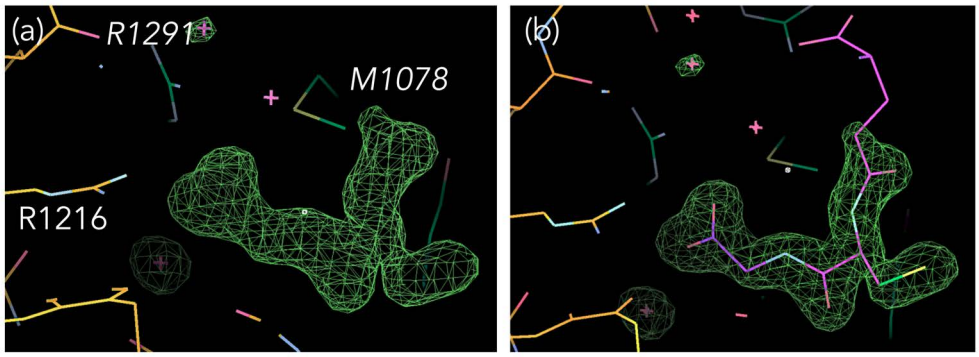


Figure 4.13: NCS averaged Fo-Fc map from the first cycle of refinement of the GSH-soaked EV-F3 dataset, i.e. prior to modelling GSH, (a) without modelled GSH and (b) with subseqeunetly modelled GSH (purple sticks). Residues from one protomer are shown as orange/yellow sticks, with a key Arg residue (see below) labelled in plain, white text. The other protomer is shown as green sticks and labelled in italic text.

Initial fitting to the difference density suggested that only the Cys-Gly portion of GSH was ordered (Figure 4.13). GSH is a tripeptide, comprising Glu-Cys-Gly (see Figure 4.14) and is catabolised to Cys-Gly (CG). To see if the binding site bound CG or GSH, and to establish which was the tighter binder, further experiments were performed; first, a competition soak with 8 mM GSH and 8 mM CG (Sigma Aldrich), and, second, 8 mM CG (Sigma Aldrich) alone. Diffraction data were collected with automated centring and collection and processed as before. The CG dataset collected from a single crystal extended to 1.67 Å resolution, and, at 99.7 % completeness, did not require merging with further data. The data for the GSH/CG, soak, however were less good and two crystals were merged to yield an overall completeness of 82.6 % to 2.20 Å resolution (one of the crystals diffracted to lower resolution (3.71 Å) and thus data were less complete at higher resolution).

Structural Analysis

As may be seen in Figure 4.16(a), the binding site (G-site) is nestled at the VP1/VP3 interface in each protomer. Interactions occur, as shown in Figures 4.16(b) and Figure 4.15 between VP1 residues 1075-1079, 1095-1098 and 1227-1231, C-terminal VP3 and VP1 residues from the adjacent protomer (1150-1155 and 1216). It is anchored through its carboxy terminus via salt-bridge contacts with a strong basic patch formed by R1216 of the adjacent protomer and R1229 at the site base. A further inter-protomer interaction is made with R1224 of the adjacent protomer through a H-

Table 4.8: Data processing statistics for glutathione investigation. Datasets marked with an asterisk contain two merged crystals.

	EV-F3 GSH*	EV-F3 CG	EV-F3 GSH/CG*
Data Collection			
Space group	*I*222	*I*222	*I*222
Cell dimensions			
a, b, c / Å	342.8, 348.2, 351.4	342.7, 348.3, 351.6	343.6, 349.2, 352.5
α, β, γ / °	90, 90, 90	90, 90, 90	90, 90, 90
Resolution / Å	20.0?1.80 (1.83-1.80)	20.0-1.67 (1.70-1.67)	110.0-2.17 (2.20-2.17)
Rmerge	0.394 (–)	0.216 (–)	0.374 (0.798)
Rpim	0.111 (–)	0.089 (–)	0.187 (0.787)
I / σI	6.2 (0.5)	7.2 (0.7)	5.7(3.6)
Completeness / %	99.9 (99.9)	99.7 (95.4)	82.6 (76.9)
Redundancy	13.4 (7.8)	6.9 (5.2)	3.7(1.4)
Refinement			
Resolution Å	20.0-1.80	20.0-1.67	20.0-2.17
No. reflections	179939/95446	2243186/117927	862708/45615
Rwork / Rfree	0.205/0.208	0.190/0.191	0.200/0.203
No. of atoms			
Protein	6291	6291	6288
Ligand/ion	86	79	94
Water	717	742	612
Ramachandran Plot / %			
Preferred regions	94.8	95.1	94.9
Allowed regions	4.4	4.1	4.3
Disallowed regions	0.8	0.8	0.9
R.m.s. deviations			
Bond lengths / Å	0.005	0.005	0.005
Bond angles / °	1.4	1.4	1.4

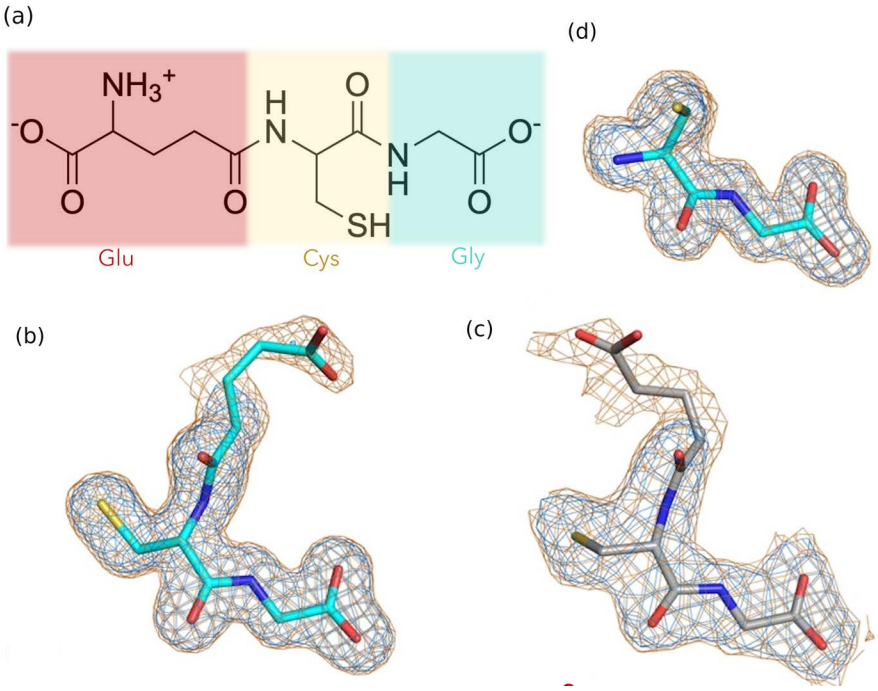


Figure 4.14: (a) Structure of glutathione. Glutathione is a tripeptide, comprising glutamic acid, cysteine and glycine, as marked. (b) 15-fold NCS averaged density for a glutathione only soak at two contour levels. (c) 15-fold NCS averaged density at the G-site for the competition soak, showing the alternative GSH pose, also at two contour levels, as before. (d) CG soak density at two contour levels.

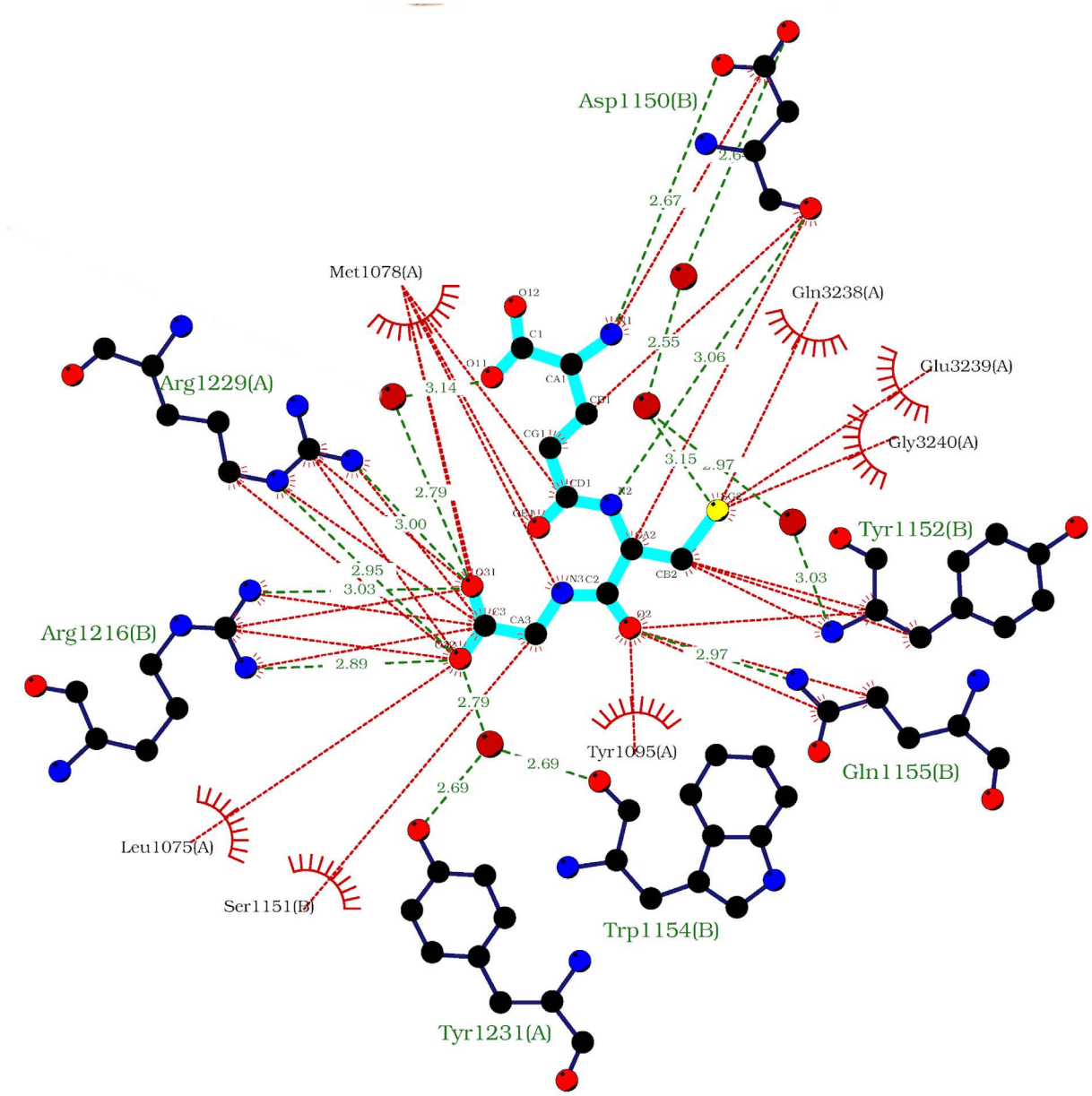


Figure 4.15: 2-dimensional plot of interactions with nearby residues, generated with Ligplot+ (Laskowski and Swindells, 2011). GSH is shown as cyan sticks, with hydrophobic interactions as brick red lines (with corresponding residues written in black text), and H-bonds and salt bridges as green dashes (connected to associated residues, shown in black) with distances in Å. The two protomers are distinguished as (A) and (B) for clarity.

bonding interaction via a bridging water molecule (see Figure 4.16(b)). The main chain carbonyl groups of the Cys and Gly portions of GSH form additional H-bond contacts with the side chain of Y1231 and the carbonyl oxygen of W1154 of the adjacent protomer via a bridging water molecule. Main chain amide and carbonyl moieties of Cys form additional H-bodning interactions with the side chain of D1150 and that of Q1155 from the adjacent protomer. The side chain of the final Glu moiety makes no contact with the capsid, but its main chain appears to form additional hydrogen bonding interactions with D1150 (adjacent protomer) and, indirectly, R1227 (Figures 4.16(b),4.15).

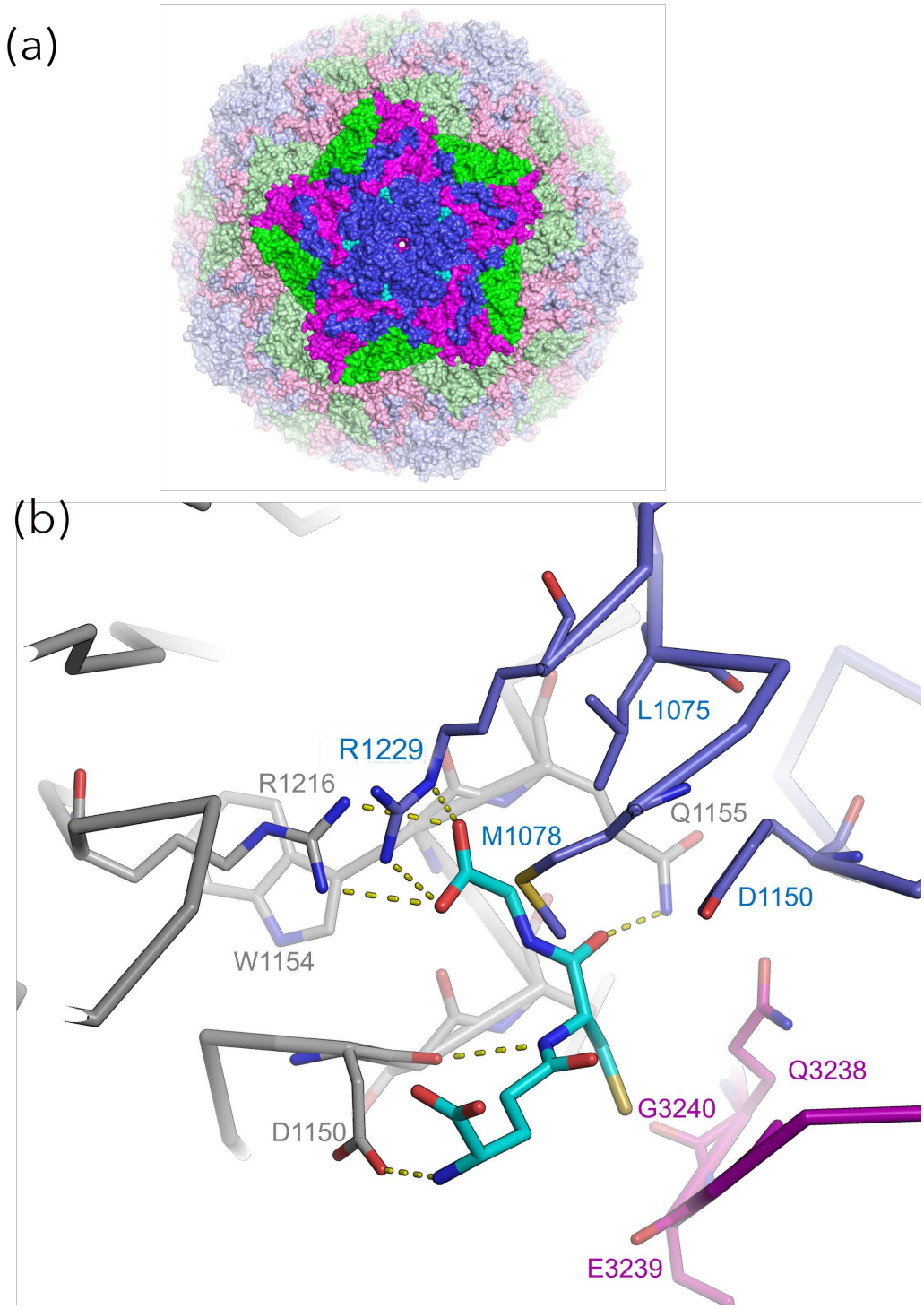


Figure 4.16: (a) Modelled glutathione in the context of the full EV-F3 virion. Modelled GSH is shown in cyan, and the virion is coloured by the usual convention, with a pentameric unit shown in bold colours. (b) GSH and interactions around the G-site. Residue labels, ribbon and sticks are coloured by VP for one protomer (blue for VP1; magenta for VP3), and should in grey for the other. GSH is represented as cyan sticks, and pertinent H-bonding interactions as yellow dashes. The main protein chain is shown in ribbon representation, with residues close to GSH shown as sticks.

Table 4.9: Infectivity assay with BSO treated EV-F3. As can be seen, even 12 hr p.i., there was a dramatic difference in CPE.

		% wells with CPE										
Time pi	Dilution	(-)	10^{-1}	$10^{-}2$	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
12hrs	-BSO	0.0	100.0	100.0	87.5	75.0	50.0	25.0	12.5	0.0	0.0	0.0
	+BSO	0.0	100.0	50.0	25.0	12.5	0.0	0.0	0.0	0.0	0.0	0.0
17hrs	-BSO	0.0	100.0	100.0	87.5	87.5	87.5	50.0	25.0	25.0	0.0	0.0
	+BSO	0.0	100.0	62.5	37.5	12.5	0.0	0.0	0.0	0.0	0.0	0.0

Table 4.10: Melting temperatures and confidence limits for GSH and CG thermostability assay. Note that 'CG100', for instance, refers to the condition where CG is in 100-fold excess of the anticipated 60 binding sites per EV-F3 capsid.

Sample	T_m / °C	Confidence ± °C	R^2
Virus only	60.4	0.2	0.99
CG100	60.1	0.2	0.99
CG1000	61.1	0.2	0.99
GSH100	61.1	0.1	1.00
GSH1000	62.4	0.1	1.00

Infectivity and Stability Assays

EV-F3 was tested for GSH sensitivity using an infectivity assay, whereby BSRT7 cells immobilised in a 96-well plate were treated with BSO 48 hr pre-infection. As may be seen in Table 4.9, EV-F3 growth diminishes by 3.5 log when treated with BSO (Table 4.9), indicative of a glutathione dependency.

The thermostabilising effect of GSH and its degradation product, CG, was then determined by thermofluor assay, where thermostability was assessed by signal from the genome-sensitive SYTO9 dye (Invitrogen, see Chapter 2, Section 2.7). For the two concentrations used (100-fold and 1000-fold excesses of GSH sites, assuming 60 per ASU), we see a delay in SYTO9 signal (Figure 4.17(a,b)), consistent with slight increases in calculated melting temperatures of 1.0 ± 0.1 and 2.3 ± 0.1 °C, respectively (Table 4.10, Figure 4.17(c)). This is by contrast to the degradation product CG, where we only see a 0.7 °C ± 0.2 °C increase when added in an 1000-fold excess of sites. The slight decrease seen with 100x excess of CG is below the confidence limits, and thus insignificant (Table 4.10).

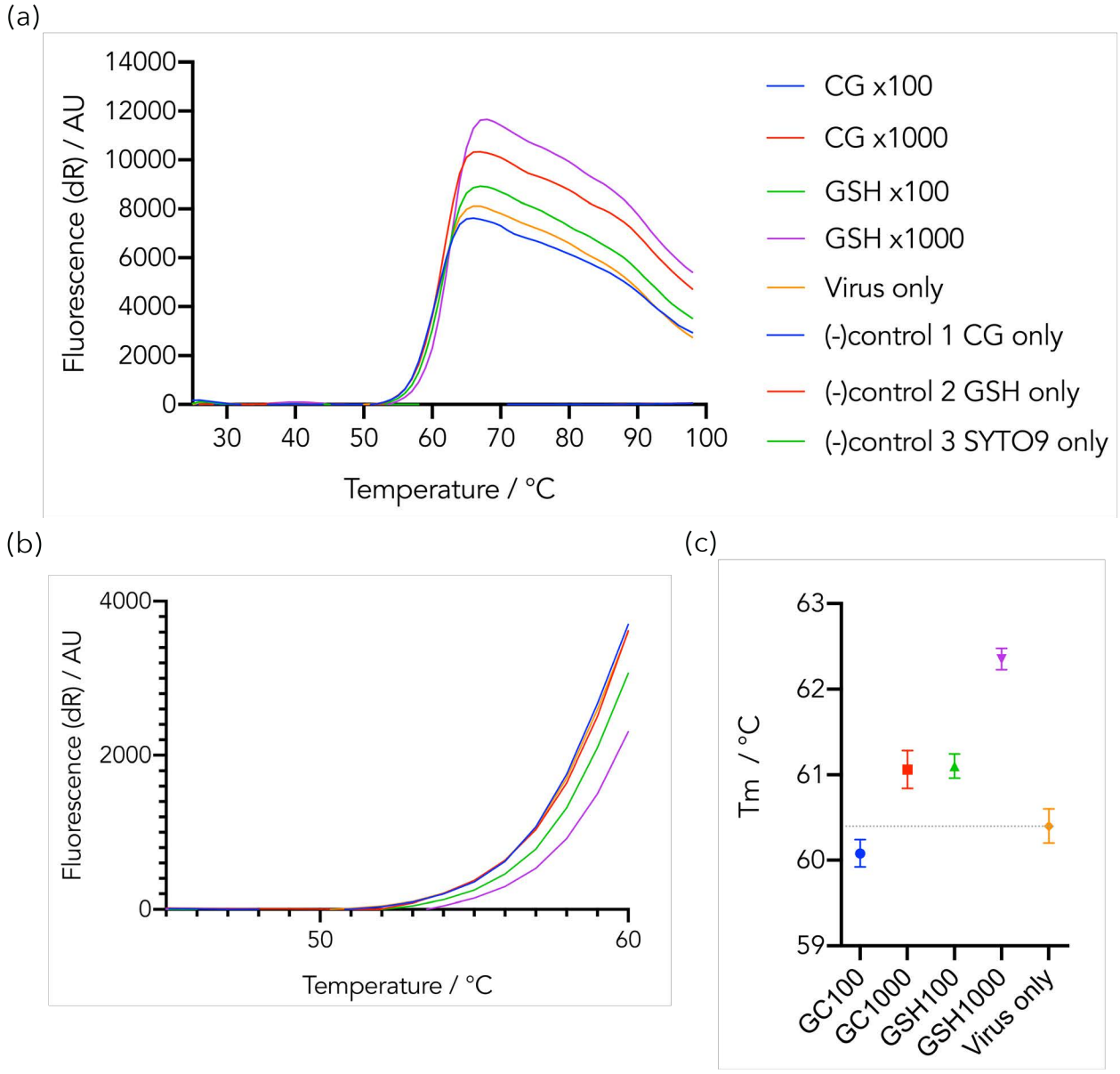


Figure 4.17: Thermofluor analysis of CG and GSH incubated with EV-F3 in 100, or 1000-fold excess of sites (x100 or x1000 as labelled), assuming 60 sites per asymmetric unit. Negative controls are shown too, corresponding to samples without virus. As may be seen, all of these controls are at baseline level. (b) Zoomed in portion of the melting curve to show differences between additives. A slight increase in stability can be seen with higher concentrations of GSH. (c) Plot of melting temperatures and associated error bars, as determined using a 5-parameter fit to each curve, the results from which are summarised in Table 4.10. Colours for points correspond to the key in (a).

Conservation of the glutathione binding site in other Enteroviruses

The G-site positive patch seems to be well conserved across enteroviruses (Figure 4.18), but is variable in breadth, and, for some enteroviruses, the modelled GSH Glu side chain clashes with the surrounding capsid. However, this flexible portion could rearrange. More broadly, there is remarkable sequence conservation across the G-site across a spectrum of enterovirus clusters.

As mentioned above, several other enteroviruses seem to have models with some occupation of the binding site. For instance, HRV14 (PDB: 4PDW) has a glycerol modelled in an analogous position, but the pocket is largely hydrophobic, with glycerol hydroxyl groups making hydrogen bond contacts with main chain atoms (Lacroix et al., 2014). Furthermore, in the empty PV structure (PDB: 1POV), we see Q3238 from VP3 occupying an analogous space (Basavappa et al., 1994).

Is the G-site Druggable?

A recent report showcased structural and functional evidence for a benzene sulphonamide derivative, so-called 'Compound-17', as a putative CV-B3 inhibitor (Abdelnabi et al., 2019). This was found to act in concert with the P-site binder plecornaril, and benzene sulphonamide variants showed activity against a number of enteroviruses (Abdelnabi et al., 2019).

Superimposition of this structure with EV-F3 bound GSH shows Compound-17 to bind at the G-site (see Figure 4.19(a)). Interestingly, elevated intracellular glutathione levels were found to not impact the antiviral activity of Compound-17, implying competition for the binding pocket.

Glutathione Resistant Mutations

A number of mutations have been identified that enable enteroviruses susceptible to GSH levels to overcome their dependence. Most notably, a mutation at T1077M. Mapping the mutations and the glutathione site onto the virus surface Figure 4.19(b), we see that the mutations generally occur at the protomer interfaces, but not directly with our G-site. Thus, resistance may be conferred by increasing inter-protomer interactions, relaxing the requirement for GSH binding. This means that, even with GSH-resistant mutations, GSH or analogues might provide an additional glue to impede, for instance, uncoating processes. Such compounds might have value as antivirals and also for the stabilisation of vaccine antigen.

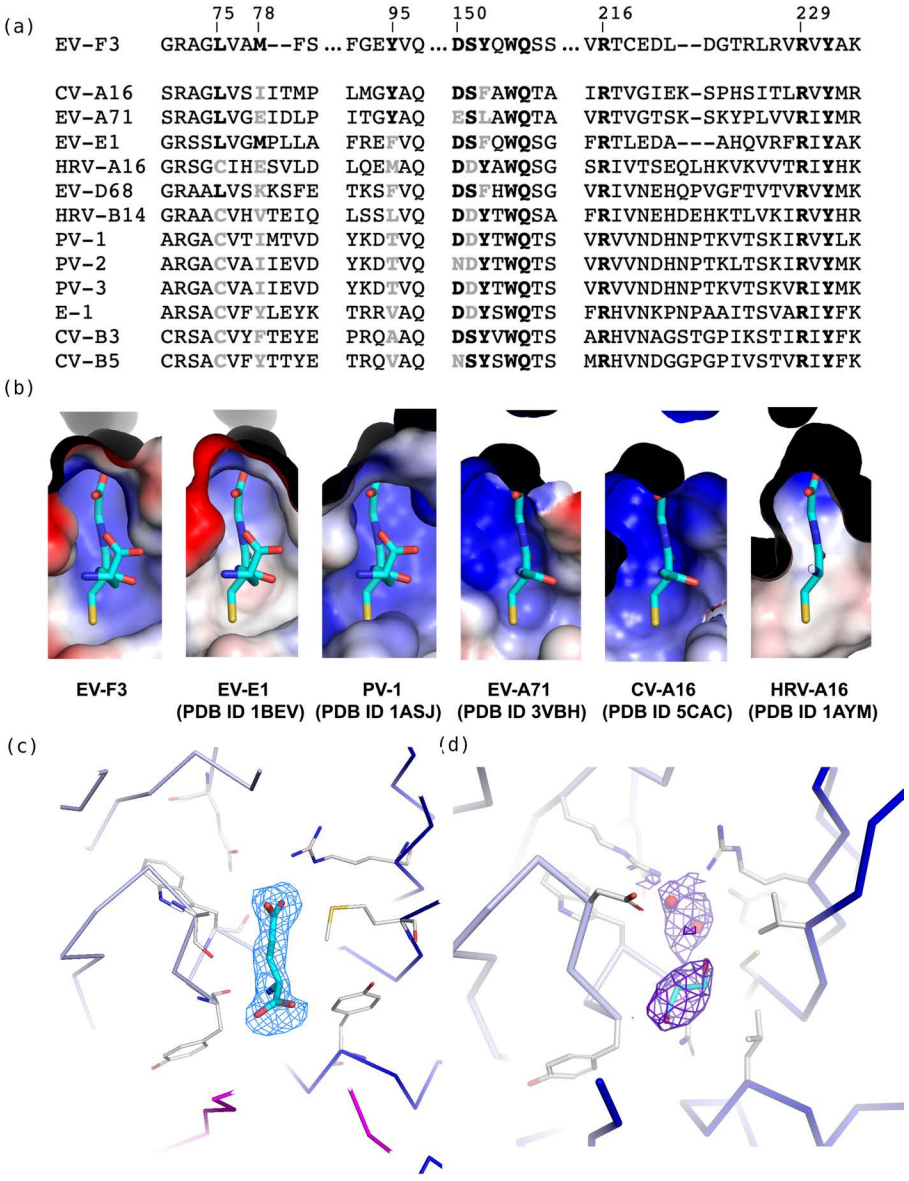


Figure 4.18: (a) Sequence alignment for a collection of enteroviruses, as labelled. Conserved GSH-interacting residues are shown in bold, and, where mismatches occur, residues are coloured grey. (b) Electrostatic charge (+/-5 kT e^{-1})) mapped onto the viral surface and coloured conventionally, for selected enteroviruses, as labelled, PDB codes in parenthesis, as reported in, from left to right, the following publications: Smyth et al. (1995); Wien et al. (1997); Wang et al. (2012); Oliveira et al. (1993). (c) Example density at the G-site for the EV-F3m room temperature structure. (d) Density at the G-site for HRV-14, PDB: 4PDW (Lacroix et al., 2014). Main chains are shown in ribbon representation, coloured by convention, with interacting residues shown as white (for C atoms) sticks, and the adjacent VP1 protomer a lighter shade of blue.

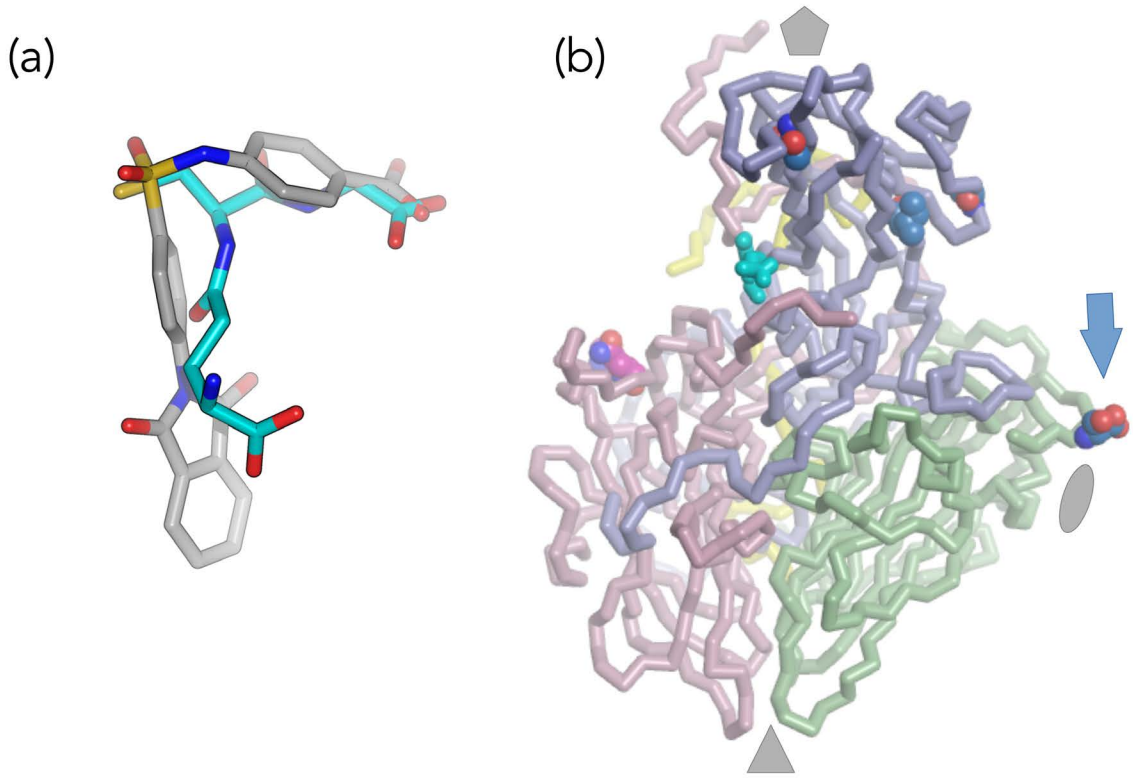


Figure 4.19: (a) Comparison of the pose of our GSH, cyan sticks with the binding site of compound 17 (grey sticks), reported as a potent enterovirus inhibitor (Abdelnabi et al., 2019). (b) Mutations observed in GLT resistant PV-1, as reported in Thibaut et al. (2014) (thick spheres), mapped EV-F3 (shown as thick ribbons coloured by VP, VP1 blue, VP2 green and VP3 pink, VP4 yellow). The one mutation that does not map, based on structural alignment, onto EV-F3 is indicated by a blue arrow. Our modelled GSH is shown in cyan, occupying a site some distance away from the Thibaut et al. (2014) mutations.

Table 4.11: Melting temperatures for EV-F3 incubated with different concentrations of 6'SL (6' sialyllactose) with SYTO9 dye.

SL	6 μM	10 μM	0 μM
Tm °C	60.4	60.3	60.3
+/-	0.5	0.2	0.1
R^2	0.94	0.99	1.00

4.1.4 Investigating Sialic Acid as an EV-F3 Receptor

Sialyllactose Assays

To investigate EV-F3's reliance on sialic acids for cell entry, a plaque assay (see Chapter 2, Section 2.2.1) with sialidase (neuraminidase from *Clostridium perfringens*, Roche) treated cells was compared against untreated controls.

Removal of BHK21 cell surface sialic acids appeared to reduce plaque formation, consistent with a reduction in EV-F3 infection (Figure 4.20), but significant deterioration at the central part of each well made interpretation difficult. The same problem was encountered after repetition of the assay, and was initially thought to be failure of growth of a cellular monolayer, careless washes, or cell disruption due to pre-incubation in the absence of FBS. However, assays without pre-incubation worked well (for instance, see Figure 1.1(a) in Chapter 3, Section 3.2.1), though additional wash steps may have been to blame.

For some picornaviruses, receptor binding triggers substantial conformational changes to prime particles for genomic egress. By contrast, sialyllactose has been found to inhibit infection by some enteroviruses, presumably either due to occlusion of the receptor binding site, or by conferring a stabilising effect that inhibits uncoating. To see if sialyllactose impacts on EV-F3 capsid stability, the virus was incubated with sialyllactose and analysed by thermofluor assay (Chapter 2, Section 2.7). Sialyllactose concentrations of 6 μM and 10 μM were used, corresponding to 100-fold and 167-fold excess of binding sites per virion, assuming 60 sites per virion. Despite this massive excess of binding sites, very little change in genomic release signal and therefore capsid thermal stability was observed upon incubation of sialyllactose (Figure 4.21).

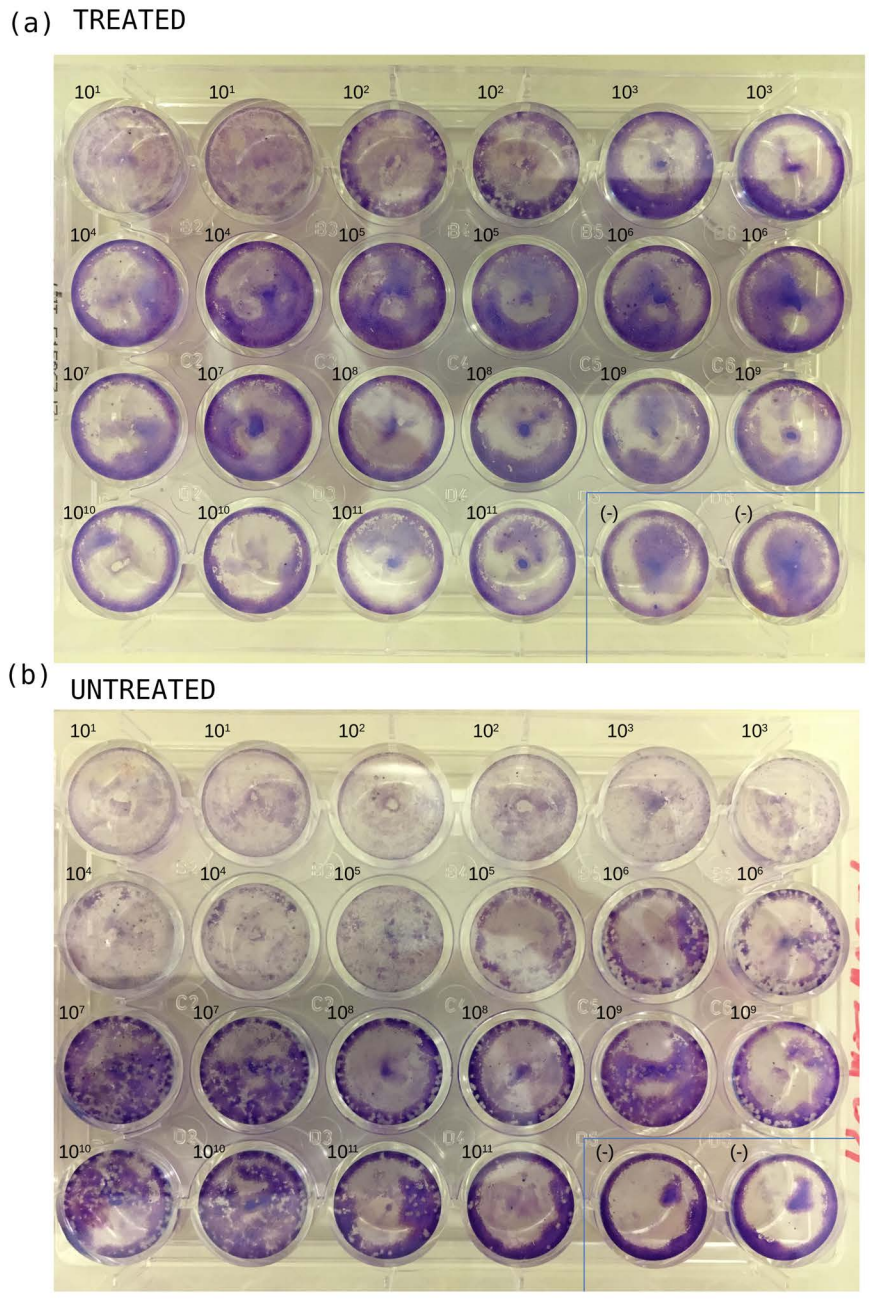


Figure 4.20: *(a) Example wells from plaque assay with BHK21 cells treated with sialidase for 1 hr before washing and infection at EV-F3 dilutions as listed on the figure. (b) Control plaque assay without sialidase treatment, instead left in serum free GMEM for 1 hr prior to washing and infection.*

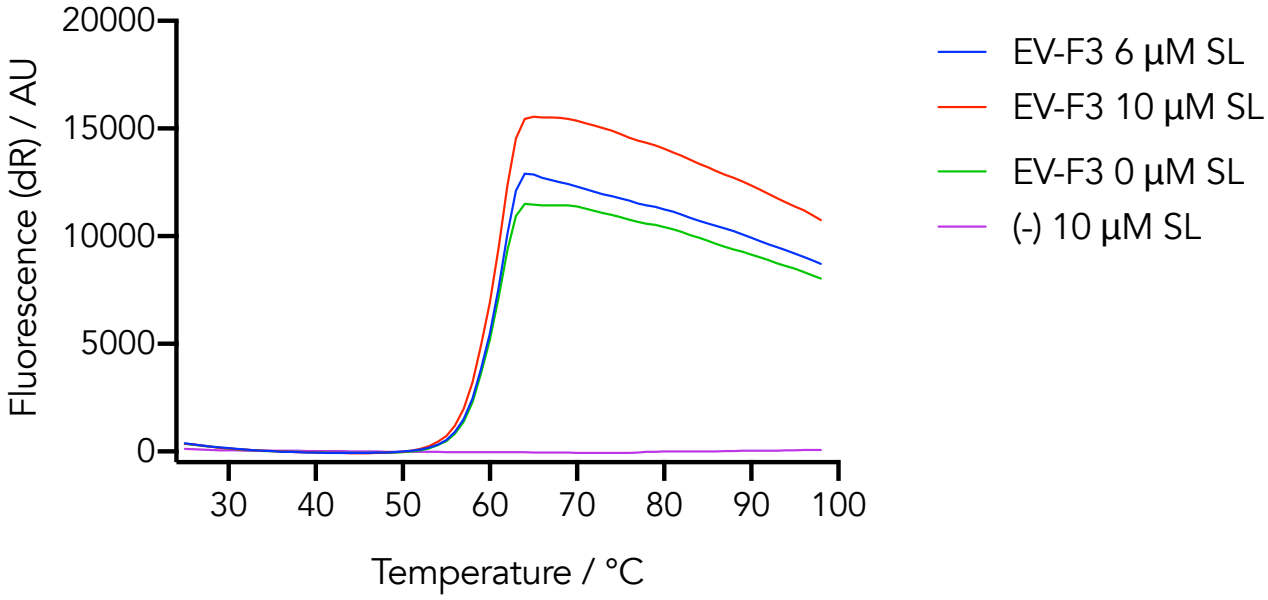


Figure 4.21: First derivative of SYTO9 melt curves with EV-F3 and siallylactose (SL), as stated in the key. Melting temperatures can be found in the accompanying Table 4.11.

Sialyllactose binding from a crystallographic prospective

Two possible sialyl linkages (3' and 6') have been modelled for picornaviruses (Chapter 1, Section 1.1.4), whilst influenza viruses only recognise one linkage. To determine whether this is the case for EV-F3, crystals were soaked with 6' sialyllactose, and both 3' and 6' sialyllactose as a competition soak, overnight. Structures will be referred to as EV-F3-3SL6SL and EV-F3-6SL in the proceeding text.

Data Collection

Data were collected at the I03 beamline, Diamond Light Source (as for section 4.1.3 above). For each condition, data were 93 % complete, and processed to 2.32 Å and 1.92 Å resolution for EV-F3-6SL and EV-F3-3SL6SL, respectively (Table 4.12). Initial models for the two sialyllactoses were derived from PDB depositions, ligand IDs SLT (used in PDB ID: 5B2D, Kubota et al. (2016)) and 4U0 (PDB ID: 5BNN, Liu et al. (2015)) for 3' and 6' linkages, respectively. These fitted well into the respective densities. As before, Coot was used for model building, but with 15-fold NCS averaged maps (Emsley and Cowtan, 2004; Emsley et al., 2010).

Can Sialyllactose Bind EV-F3?

To establish whether sialyllactose was bound, a difference map [Fo(SL)- Fo(native)] was determined using our EV-F3-CG structure as a 'native' model (this had comparable unit cell dimensions, see section 4.1.3 above), and each of the sialyllactose soaked structures (Figure 4.22). Reflection files were first combined with Sftools (B. Hazes, unpublished), before structure factors for EV-F3-SL and EV-F3 native were scaled on a resolution shell basis using Shellscale (D. Stuart, R. Esnouf, unpublished). [Fo(SL)-Fo(native)] was then calculated using Sftools (B. Hazes, unpublished), and the resulting map 15-fold averaged in Coot as before (Emsley and Cowtan, 2004; Emsley et al., 2010).

For both the competition and 6' sialyllactose only soaks, difference density supported the modelling of the sialic acid portion of sialyllactose on the capsid surface.

3' versus 6' Sialyllactose

To establish whether a noticeable preference for one of the two sialyl linkages can be discerned from the competition soak data, we turn again to the difference maps shown in Figure 4.22. Sialyllactose comprises a sialic acid (Sia) linked to galactose (Gal), via a 6' (Figure 4.22(a)) or 3' linkage (Figure 4.22(a)), followed by glucose (Glc). For the competition soak, difference density (Figure 4.22) accounts for the Sia and part of the Gal portion, making it hard to distinguish between 3' sialyllactose and 6' sialyllactose. However, as may be seen in Figure 4.22(b), the difference density accounts for the Gal portion of the 3' sialyllactose better. Note that the 6' sialyllactose initial model needed some adjustment to fit into the 6' only sialyllactose difference map.

Sialyllactose Location and Interactions

Difference density for siallylactose occurs at the south rim of the canyon (Figure 4.23)(a), close to the VP2 EF loop and VP1 C-terminus (Figure 4.23(b)).

Sialyllactose binding displaces three H_2O molecules and a sulphate ion that interacts with the main chain S2132 (EF loop) and Y2258 residues. The carbonyl moiety of Sia sits close to the displaced sulphate ion position, but closer to the capsid surface, making favourable contacts (see Figures 4.22(d), 4.23(c)). For maximal contact with sialyllactose, at EV-F3's surface, the VP2 EF loop

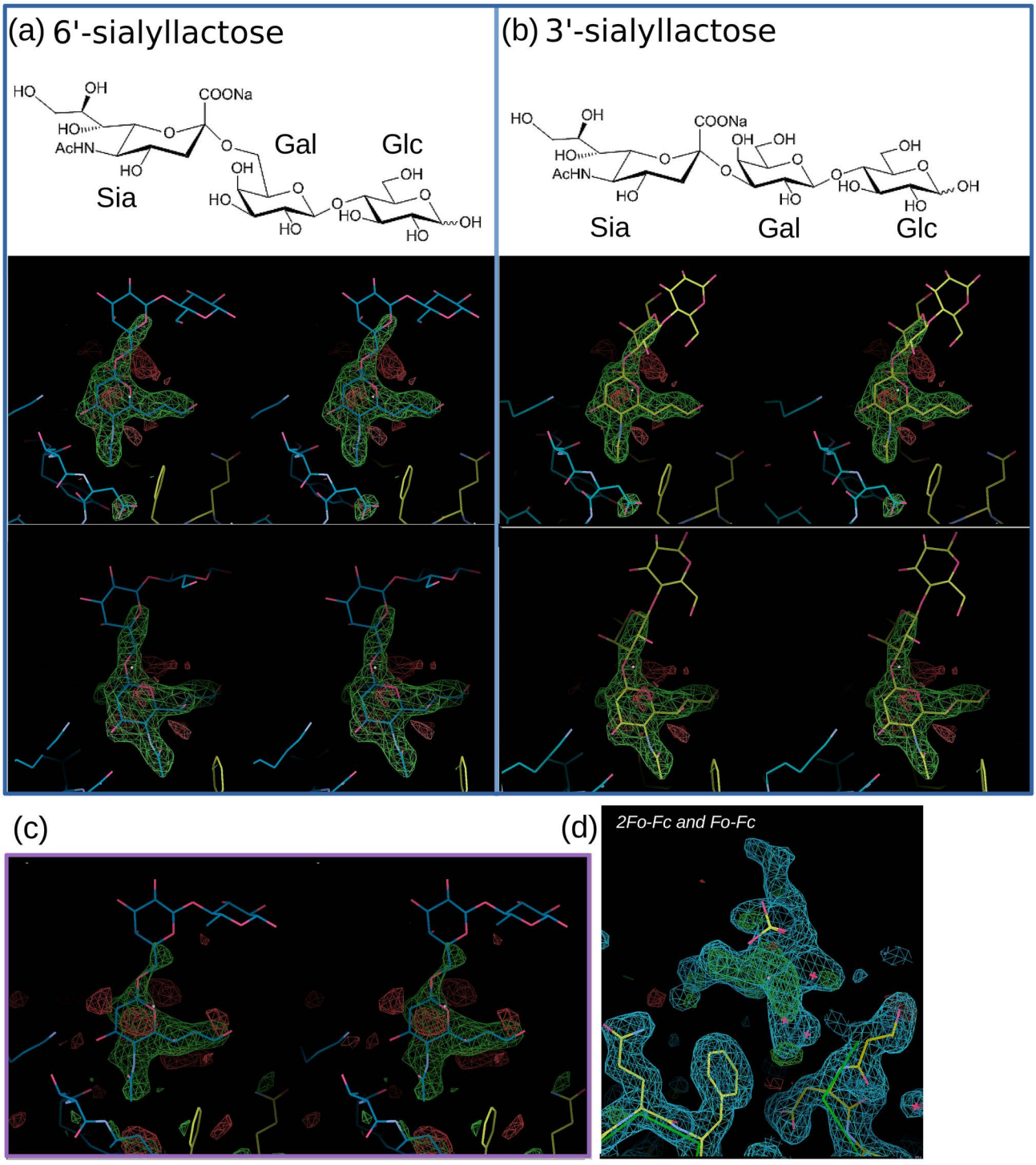


Figure 4.22: Blue box, (a,b) Top: structures of 6′ and 3′ sialyllactose, Sia = sialic acid; Gal = galactose, Glc = glucose. The lower two images are wall-eyed stereo representations of [Fo(SL)-Fo(native)] for the 3′ sialyllactose/6′ sialyllactose competition soak with (a) 6′ sialyllactose and (b) 3′ sialyllactose models. (c) Wall-eyed stereo representation of difference map for the 6′ sialyllactose only soak. (d) 2mFo-dFc and Fo-Fc map from the competition soak with the native EV-F3 model overlaid (main chain in ribbon representation, sulphate ion and waters in stick representation). All maps are 15-fold NCS averaged.

Table 4.12: Data processing statistics for sialyllactose crystallography experiments.

Crystal	EV-F3-6SL	EV-F3-3SL6SL
Data Collection		
Space group	*I*222	*I*222
Cell dimensions		
a, b, c / Å	342.6, 348.3, 351.2	342.8, 348.0, 351.4
$\alpha, \beta, \gamma/°$	90,90,90	90,90,90
Resolution / Å	123.6-2.32	123.63-1.92
Rmeas	0.325 (1.00)	0.249 (0.863)
I / σI	3.5 (1.3)	3.9 (1.0)
CC0.5 / %	0.80 (0.30)	1.00 (0.50)
Completeness %	93.2 (91.8)	93.1 (78.8)
Redundancy	3.6 (3.5)	3.3 (2.0)
Refinement		
Resolution / Å	20.0-2.32	20.0-1.92
No. reflections	788276/41839	1459354/131026
Rwork / Rfree	23.4/23.7	20.9/21.1
No. of atoms		
Protein	6291	6291
Ligand/ion	92	92
Water	612	623
Ramachandran Plot / %		
Residues in preferred regions	93.1	94.6
Residues in allowed regions	6.1	4.6
Residues in disallowed regions	0.75	0.75
R.m.s. deviations		
Bond lengths / Å	0.006	0.006
Bond angles / °	1.4	1.4

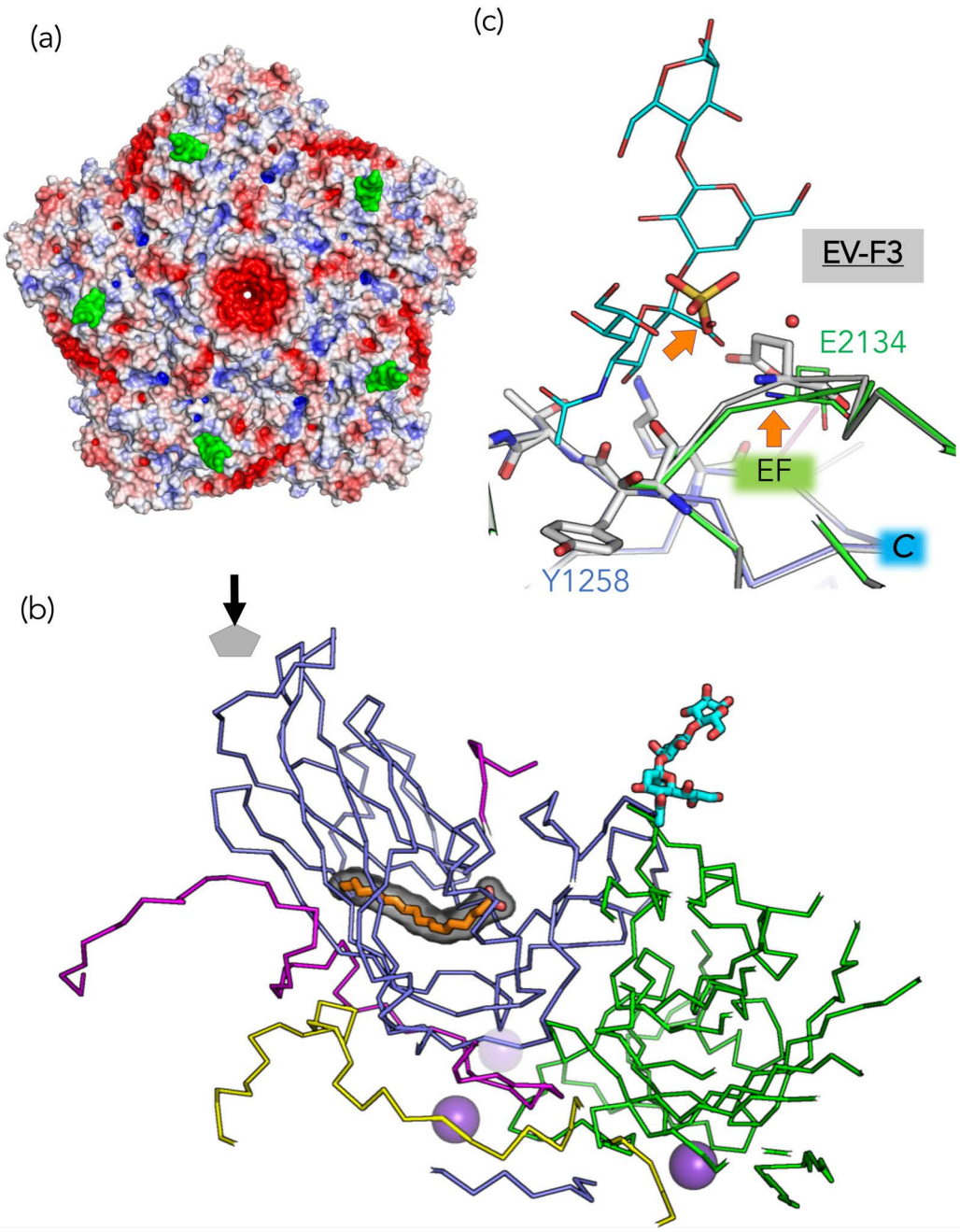


Figure 4.23: (a) Isosurface of EV-F3 pentamer, coloured from -5 (red) to +5 (blue) kT e^{-1}. Modelled 3′ sialyllactose is shown in green. (b) Side view of capsid protomer (5 fold shown as a grey pentamer), with modelled 3′ sialyllactose in cyan. The pocket factor is shown as orange sticks, surrounding by a grey cavity. VPs are shown in ribbon representation with VP4 coloured yellow; VP3 magenta; VP2 green and VP1 in blue. Potassium ions are shown as purple spheres. (c) Overlay of the EV-F3-3SL6SL model (coloured by VP as in (b)) with unsoaked EV-F3 (grey, model from CG soak shown above). Pertinent changes are indicated by orange arrows; first, the modelled SO_4^{2-} ion for unsoaked EV-F3 is displaced by the Sia portion of sialyllactose. Second, the EF loop is pushed away from the VP1 N terminus.

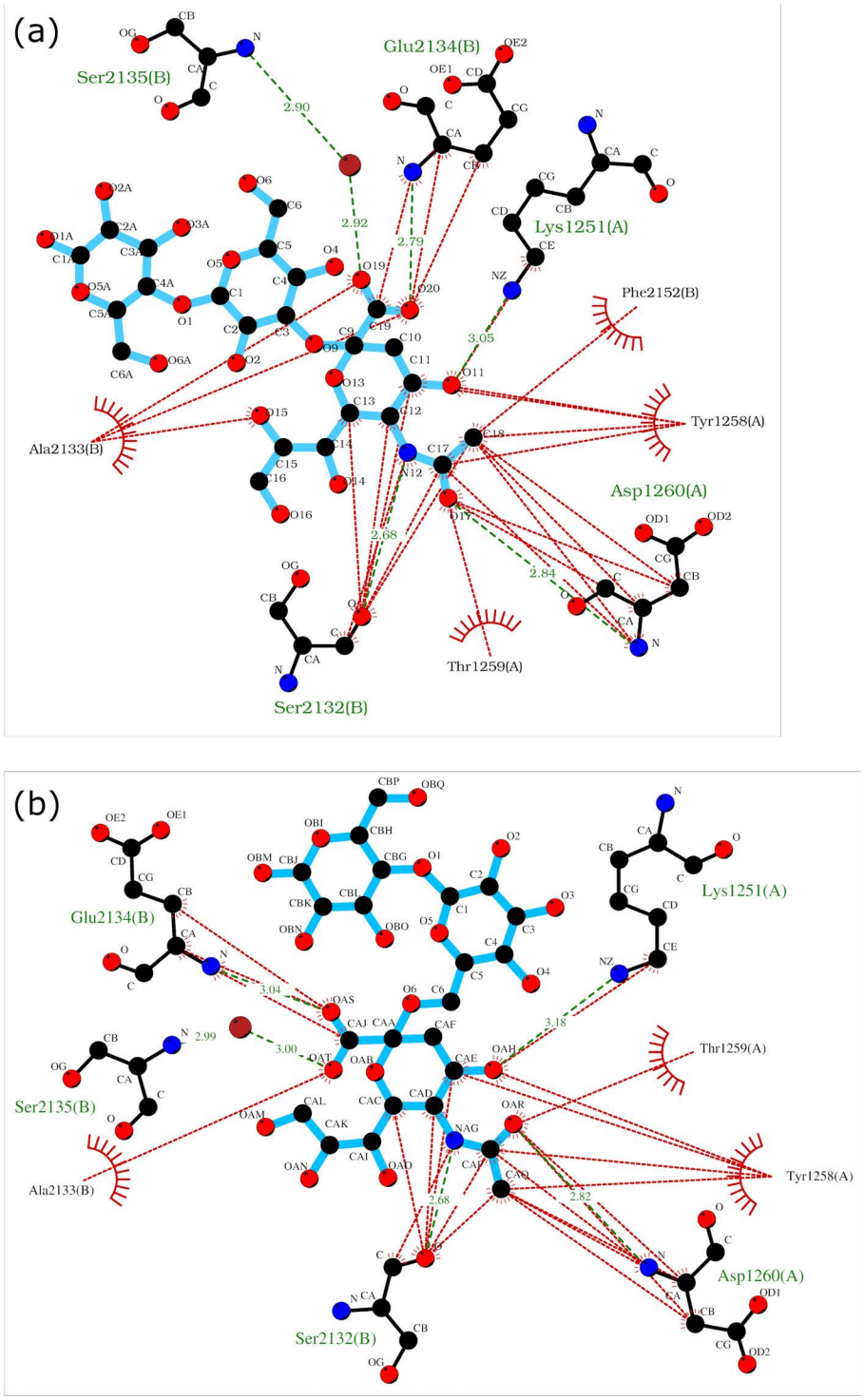


Figure 4.24: 2D representation of the binding surface between (a) 3′ sialyllactose (pale cyan) model from the competition soak dataset, and (b) 6′ sialyllactose (sky blue). Hydrophobic interactions are shown in red-brick coloured lines, and hydrogen bonds as green dotted lines. Atoms are coloured by conventional colouring, with water molecules shown as a deep red colour. Residue numbering adheres to the convention stated for this thesis, but, for further clarity, corresponding chains are in parenthesis, where (A) corresponds to VP1 and (B) VP2. Generated by Ligplot+ (Laskowski and Swindells, 2011)

appears to shift, relative to unsoaked EV-F3 structures (see Figure 4.23(c)). Residue E2134 (VP2 EF loop) is pushed out of the way (Figure 4.23(c)).

Despite compelling difference map density supporting the galactose moiety for the EV-F3-3SL6SL data (Figure 4.22(a,b,d)), there are no obvious galactose contacts with EV-F3. Indeed, for both 3' and 6' sialyllactose, Sia is anchored in place by via four hydrogen bonding interactions with VP1 C-terminal and VP2 EF-loop residues (Figure 4.24) via its N-acetyl contacting D1280's main chain N atom, and S2132. An additional contact can be seen between Sia and Y1258's carbonyl O atom, the side chain of K1251, and, via a bridging H_2O, S2135 (Figure 4.24).

Comparison with EV-E1

EV-E1's VP2 EF loop is notably more extended than EV-F3 such that it clashes (residues 2134-2136) with our modelled sialyllactose Figure 4.25. Of course, rearrangement of this loop could occur upon receptor binding, but, to accommodate the extended loop of EV-E1, additional, compensatory movements would be required amongst neighbouring surface residues, and, given the limited number of contacts made between sialyllactose and the capsid surface, it seems unlikely that this energetic penalty could be overcome upon binding to EV-E1. Therefore, sialyllactose probably engages with EV-E1's capsid at a different location (as mentioned earlier, EV-E1 has been shown to be dependent on sialic acids for cellular entry (Stoner et al., 1973)).

Comparison with other Enterovirus Sialic acid binders

Interacting residues for EV-F3, mapped onto a sequence alignment are shown as blue stars in the Appendix, and a list of known sialic acid binding picornaviruses can be found in Chapter 1, Section 1.1.4.

Our modelled sialyllactose clashes with the A2134-S2140 VP2 EF loop of EV-D68. Rather, EV-D68's sialic acids are modelled some 30 Å away in the 'eastern end' of the canyon formed at the intersection between VP1, 2 and 3 (Liu et al., 2015) (Figure 4.26(a)). CV-A24v's sialic acid binding site differs again, instead occupying a position at VP1 that interferes with the VP1 N1122-V1127 loop of EV-F3 (Zocher et al., 2014), Figure 4.26(b).

Notably, however, EV-F3's sialyllactose is situated in a similar site to residues 190-194 of SCARB2 bound to EV-A71 (PDB: 6I2K, Zhou et al. (2019)), except for the latter being slightly more raised

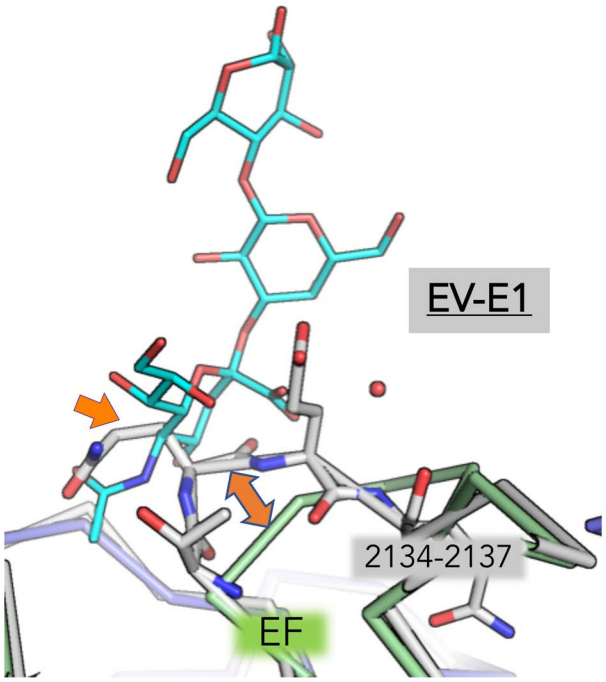


Figure 4.25: Overlay of EV-F3-3SL6SL model (coloured by VP: VP2 green and VP1 in purple blue) with EV-E1 in grey (PDB: 1BEV) (Smyth et al., 1995). Residues 2134-2137 of EV-E1 (VP2 EF-loop) are shown as sticks, and clash with the modelled Sia portion of sialyllactose (indicated by orange arrows). Part of the reason for this is the more extended EF-loop in EV-E1 (the gap between EV-F3 and EV-E1's EF-loops is shown as a double arrow).

from the VP2 surface owing to EV-A71's more extended VP2 loop (T2139-G2137), Figure 4.26(c). EV-A71 entry has been shown to be inhibited by soaking with sialic acids, and it is conceivable that sialic acids bind at the same place on the EV-A71 capsid (Yang et al., 2009).

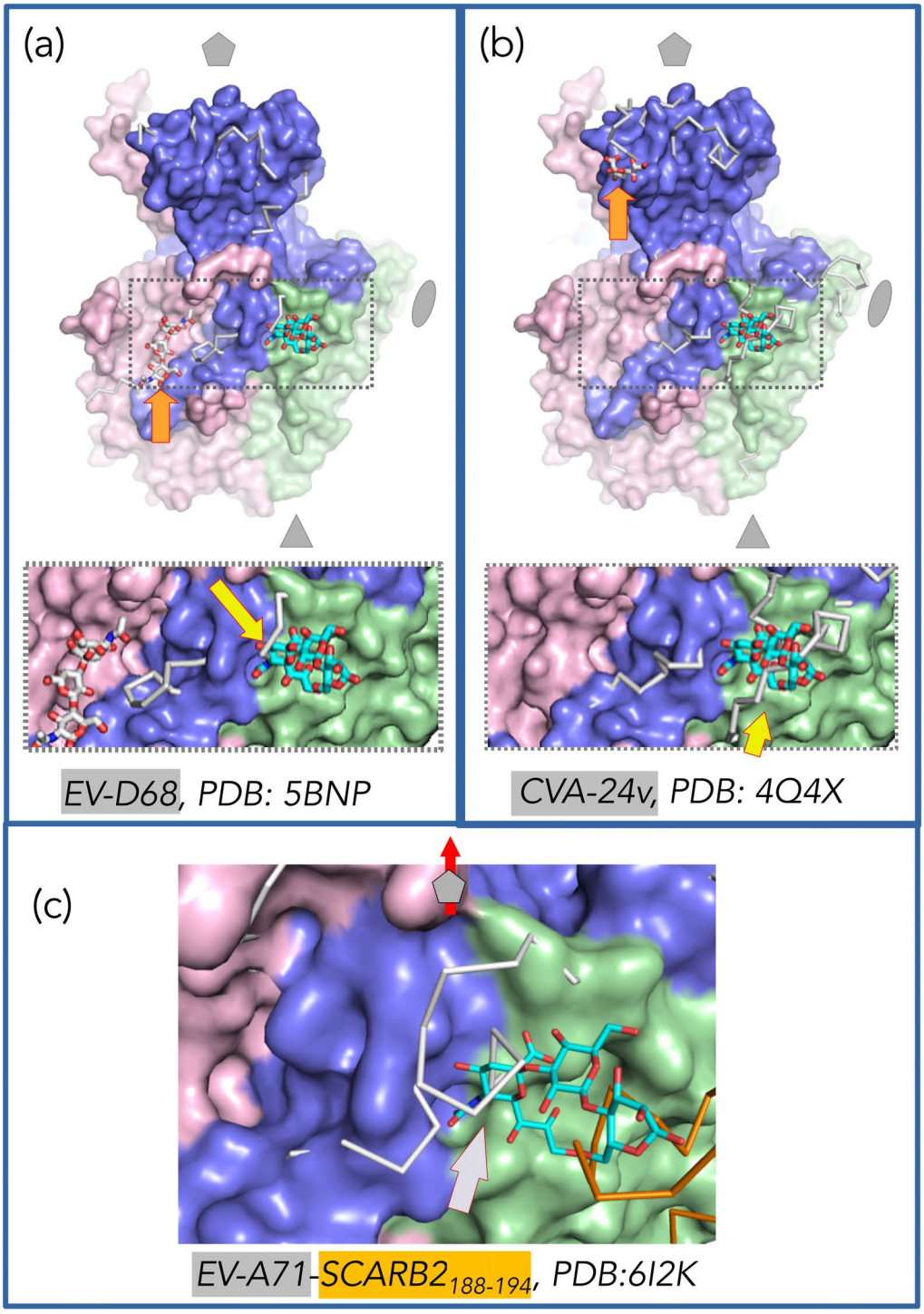


Figure 4.26: Overlay of EV-F3 protomer in surface representation (coloured blue, green and pink for VP1, 2 and 3) with modelled 3′ sialyllactose shown as cyan sticks with (a) EV-D68 displayed in grey ribbon with 6′ sialyllactose as grey spheres (PDB: 5BNP, Liu et al. (2015)); (b) CV-A24v in grey ribbon with Sia as grey sticks (PDB: 4Q4X, Zocher et al. (2014)). Modelled sialyllactoses for EV-D68 and CV-A24v, respectively, are indicated with orange arrows in the top panel. Clashing loops for EV-D68 and CV-A24v are shown as yellow arrows. (c) EV-A71-SCARB2, (PDB 6I2K, Zhou et al. (2019)) in grey ribbon, with SCARB2 residues 188-194 in orange ribbon.

4.2 Discussion: Implications for Drug Development

None of the additives covered in this Chapter, whether this be GSH, metal ions, putative receptor, were shown to confer a substantial thermostabilising effect on EV-F3's capsid, based on thermofluor assays alone.

Metal ions may have a minor influence on capsid thermostability, assuming minimal interference between SYTO9 dye and probe metal ions.

The slight increase in melting temperature determined for EV-F3 incubated with MES at pH 7.1 (1.2 °C $\pm$ 0.5 °C) is consistent with results found previously for CV-B3 with the longer MOPS (Carson et al., 2017). For CV-B3, authors hypothesised that MOPS bound to the P-site entrance, making salt bridge interactions with Arg and Asn residues: R1095 and N1211. This modelling was focussed on the P-site alone and one of the P-site residues identified as key for interacting with the sulphoxide group of MOPs is not conserved in EV-F3, where we instead find G1093. The way to resolve this would be by structural analysis.

The four K^+ ion binding sites validate some previous crystallographic models. A more rigorous quantification of the anomalous peaks might use the integrated peak volume rather than peak heights. K^+ ions were liganded by five residues, when the coordination sphere was set to 3.5 Å, but it is possible that some internal RNA may occupy additional coordination sites, especially for the site located at the 2-fold axis. More broadly, the ease with which these sites were identified unequivocally by X-ray analysis demonstrates the power of long wavelength crystallography for the analysis of lighter elements in virus particles.

The G-site may hold potential for as a binding site for antivirals, but it is likely that resistance would occur rapidly. Nevertheless, the close proximity of the glutathione site to some of the key interacting residues with our receptor binding site (see Appendix for sequence overlay, green stars) implies that glutathione may have the potential to inhibit entry. Given its occurrence at the interface between two pentameric units, and the knowledge that these pentameric units remain intact during uncoating but alter their configuration, it is likely that only some mutations can be tolerated in this region.

Structures demonstrate a propensity of EV-F3 for sialic acids, and plaque assays show some dependence on their presence for entry, with a preference, demonstrated by competition soaks, for 3' sialyllactose, contrasting studies on CV-A24v, where authors reported a preference for 2,6

linkages (Zocher et al., 2014). The sialidase used for plaque assays was non-specific for 3', or 6' linkages, both of which have been reported as putative receptors for other picornaviruses (see Chapter 1, Section 1.1.4). However, given that crystal soaks indicate EV-F3's capability to bind to both 3' and 6' sialyllactose (this may be rationalised by the contact point of the sialyllactose), it seems likely that EV-F3 may utilise both for entry.

However, the exact role of sialic acids in EV-F3 entry is unclear; it may serve as a secondary receptor, as is the case for EV-A70, for instance (Nokhbeh et al., 2005; Zocher et al., 2014). In addition, the EV-F3 used in this study had been heavily passaged, meaning that sialyllactose binding may be a product of tissue culture adaptation, as was the case for FMDV binding to HS (Sa-Carvalho et al., 1997), although this does not rule out the virus being predisposed to binding a glycoprotein or ganglioside. Sialic acids tend to occur in intestinal epithelial cells, and thus this result may provide some insight into EV-F3's tissue tropism.

Collectively, the structures presented in this Chapter provide compelling evidence for capsid binding, but additional functional studies are required to fully understand the complex capsid dynamics.

Chapter 5

Characterising FMDV VLP Vaccine Candidates

5.1 Introduction

FMDV remains endemic in impoverished regions globally, and, despite stringent trade control programs, culling and vaccination, major outbreaks have occurred in several developed nations in the past 20 years, including European countries (2000, 2001) and Japan (2000, 2010), Chapters 1, 7 (Knight-Jones and Rushton, 2013; Sinkala et al., 2014). Efforts to control the spread of this highly infectious disease have historically involved the culling of all infected animals and those in close proximity to a site of infection (the net efficacy of this approach remains dubious), resulting in a disease free area, but without vaccination status. For other regions, FMDV spread is controlled by vaccination and considered disease free with vaccination, while others, particularly more disadvantaged continents such as Africa, have no OIE status (the OIE is The World Organisation for Animal Health), Figure 5.1.

Seven serotypes of FMDV exist, and current vaccine antigens are effective against only a single serotype, sometimes failing to provide even cross-subtype protection (Brooksby, 1982; Cartwright et al., 1982), thereby necessitating the use of multivalent preparations for adequate protection (Asfor et al., 2014).

Upwards of one billion doses of inactivated FMDV vaccine are produced worldwide annually, but stocks have a limited shelf life due to the acute pH and temperature lability of FMDV (Brown

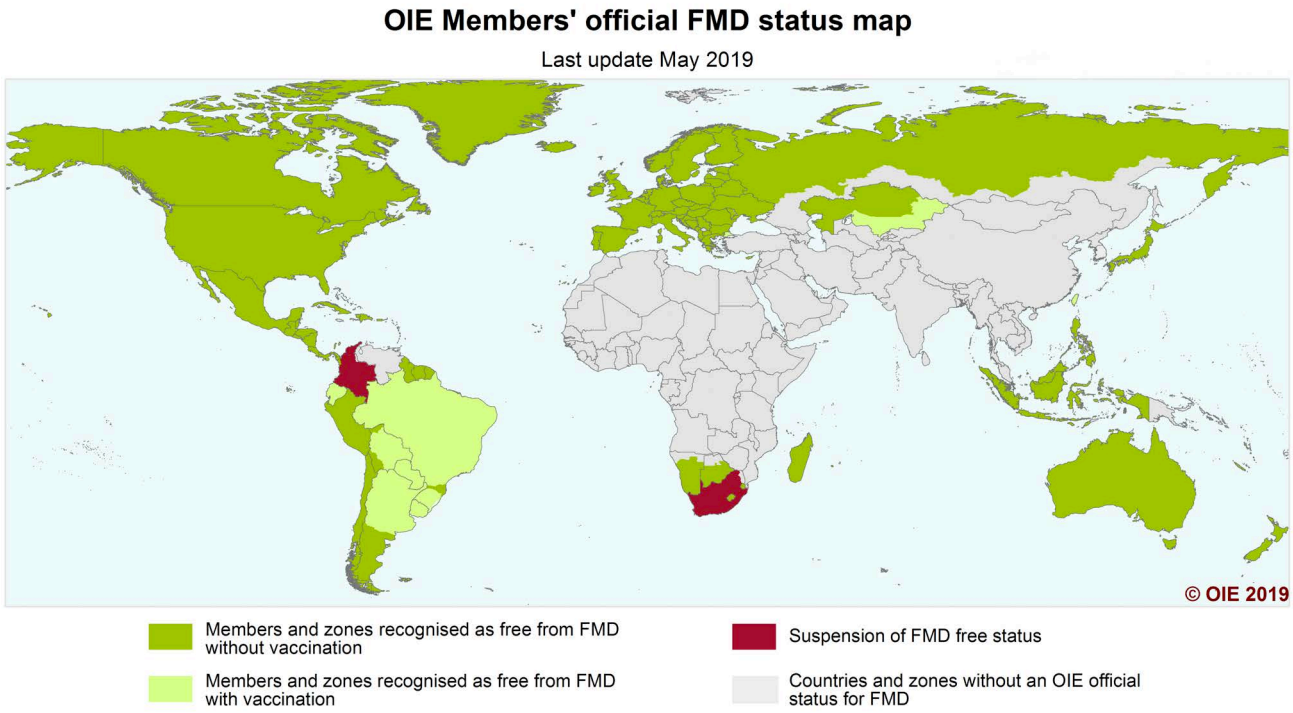


Figure 5.1: OIE world map coloured by FMDV 'status', adapted from the World Organisation for Animal Health (OIE), https://www.oie.int/en/animal-health-in-the-world/official-disease-status/fmd/en-fmd-carte/.

and Cartwright, 1961; Doel and Baccarini, 1981; Knight-Jones and Rushton, 2013; Porta et al., 2013a). This necessitates the storage of vaccines in a cold chain, but, even with such precautions, preparations only last for a short time. Furthermore, the majority of outbreaks occur in rural communities lacking the infrastructure to maintain chilled storage (Sinkala et al., 2014; Yano et al., 2018).

Current FMDV vaccine preparations contain binary ethyleneimine (BEI) inactivated FMDV, and thus require costly containment facilities for safe production of live virus (Cottam et al., 2008; Aarthi et al., 2004). Outbreaks associated with vaccine production centres have heightened the barrier to creating such containment facilities (Sinkala et al., 2014). BEI inactivation further reduces particle stability, such that they rapidly deteriorate into pentameric subunits of differing antigenicity to native virus, at temperatures above 30 °C (Cartwright et al., 1980; Doel and Baccarini, 1981; Doel and Chong, 1982; Mattion et al., 2004; Paton et al., 2005). The extent of thermo- and pH-sensitivity varies between serotypes, with A and Asia1 being the most stable and O and SAT2 the least stable (Doel and Baccarini, 1981). Traces of non-structural proteins (NSPs, see Chapter 1, Section 1.1.1) within vaccine preparations also interfere with the serological tests used to distinguish between infected and vaccinated animals, although this may be mitigated by more rigorous purification (Dill et al., 2019). Virus yields are also inconsistent, in part due to variations in cell culture media, making it often challenging to scale up quickly in times of crisis (Dill et al., 2019). Vaccines also fail to provide sterile immunity, meaning that vaccinated animals can still be carriers of the disease as viral replication at the epithelial cell surface is still possible (Alexandersen et al., 2002).

Numerous alternative methods of vaccine development have been reported (Cao et al., 2016), but are beyond the scope of this introduction. I will focus on one such method, recombinant empty capsids (VLPs, see Chapter 1), which obviate the need for live virus handling, but still maintain the antigenic properties required for effective protection, thereby providing sterile immunity. Despite much research into recombinant FMDV vaccines, as yet, no candidates have reached the market (Grubman and Baxt, 2004).

Many expression systems have been employed for picornavirus VLP production, including *Escherichia coli* (Guo et al., 2013), baculovirus driven expression in insect cells (Li et al., 2008; Cao et al., 2009; Li et al., 2011; Subramanian et al., 2012; Li et al., 2012b; Porta et al., 2013a; Bhat et al., 2013), poxvirus to drive expression in mammalian cells (Zheng et al., 2006), and replication de-

ficient human adenovirus driven expression (Moraes et al., 2002; Lu et al., 2008; Pacheco et al., 2005; Barrera et al., 2018). FMDV VLP vaccine design strategies are often hampered by its poor thermo- and acid-stability (37 °C, pH < 6.5) (Brown and Cartwright, 1961; Porta et al., 2013a).

Empty capsids are even less thermostable than virus particles (Rowlands et al., 1975; Doel and Baccarini, 1981; Basavappa et al., 1994; Curry et al., 1997) (note, however that empty capsids have been found to be more acid resistant than full virions (Curry et al., 1995)), and the Stuart lab, in collaboration with the Pirbright Institute, have rationally engineered virus-like particles to yield VLPs with greater acid and thermo-stability traits (Porta et al., 2013a; Kotecha et al., 2015). One particularly successful design includes a cysteine disulphide bond engineered between the two VP2 helices from adjacent monomers at the two-fold axis following the observation that the relatively stable A serotypes form a stacking interaction with between histidine residues at position 2093 of adjacent VP2s, whereas more labile O1M, SAT2 contain serine or lysine (King et al., 2009; Kotecha et al., 2015). This resulted in VLPs with impressive heat and pH stability that have been shown by western blot analysis to withstand pHs of 5.2 for 15 min and temperatures of 56 °C for 2 hr (Porta et al., 2013a). Similarly, a F/S2093Y mutation to introduce a $\pi - \pi$ stacking interaction also works well in infectious clone viruses, but, for VLPs, this has tended to express more poorly (personal correspondence, Claudine Porta and Elizabeth Fry). The 2093C mutation has been most extensively applied, with fairly consistent success (personal correspondence, Elizabeth Fry).

Our understanding of the nature, antigenicity and long-term stability of particles with an engineered disulphide bridge at the 2-fold axis (residue 2093C, here termed 'cc'), relative to their native counterparts, remains poor. Thus far, a crystal structure of A22 serotype with the engineered disulphide bridge (hereafter termed A22cc) has been published (PDB: 4IV3) (Porta et al., 2013a), but in general the relative proportion of particles for which a S-S bridge has been formed is unclear. In this chapter I will describe further studies on capsids with this mutation.

As described in Chapter 1 and shown in Chapter 7, divalent metal ions, Ca^{2+} in particular, appear to be important for FMDV capsid stability, and FMDV O1M has shown greater thermostability (1.5 °C) when incubated with over 0.5 M calcium (Kotecha et al., 2016). To see if Ca^{2+} ions are bound to VLPs, experiments were conducted at the soft-wavelength beamline I23 (Diamond Light Source, UK).

In addition, efforts to develop a rapid means of characterising VLP stability, building on the

PASTrY assay (see Chapter 2 and Walter et al. (2012)) that circumvents the less sensitive hydrophobic SYPRO dyes (Thermo Fisher), since the more sensitive genomic release dyes are inappropriate, are also described.

5.2 Results

5.2.1 Investigations into S-S bridging in FMDV VLPs

To assess particle quality following medium term cold storage of vaccinia-produced VLPs, cryoEM experiments were conducted. To this end, vaccinia derived S93C-engineered A22 (A22cc) and AsiaShamir (AsiaShcc) that had been stored in 40 mM Sodium Phosphate pH 7.6 buffer with 150 mM NaCl at 4 °C for 5 months were immobilised on lacey carbon grids (Agar Scientific) and vitrified in liquid ethane (Chapter 2).

In addition to an engineered S-S bridge at the 2-fold axis, 'natural' S-S bridges have also been observed between the N-terminal VP3 residues (C3007 for AsiaShcc) that surround the the 5-fold axis. To assess the effects of S-S disulphide bridge reduction on capsid stability, AsiaShcc and A22cc VLPs were incubated with the reducing agent DTT. AsiaShcc was treated with 1 mM DTT for 3 hr, while A22cc aliquots were treated with 10 mM DTT for 3 hr or 1 mM DTT for 30 min. The 10 mM DTT results will be presented later (Section 5.2.3). For each experiment, fresh DTT solutions were prepared using a 50 mM HEPES pH 8.0, 200 mM NaCl buffer, and treated VLP samples were kept slowly rotating at 4 °C. Unfortunately, due to sample availability, it was not possible to repeat the 3 hr 1mM DTT condition for A22cc. Samples were then applied to lacey carbon grids (Agar Scientific) before vitrification in liquid ethane using the standard protocols described in Chapter 2.

For all samples, data were collected with a Titan Krios microscope (Thermo Fisher).

Grid Screening

Micrographs of both A22cc and AsiaShcc 1 mM DTT treated and untreated samples showed particles devoid of clear interior density (Figure 5.2), consistent with genome-free VLPs. A number of flower-like particles could also be seen for all VLP samples studied, and these are currently under investigation (Figure 5.2).

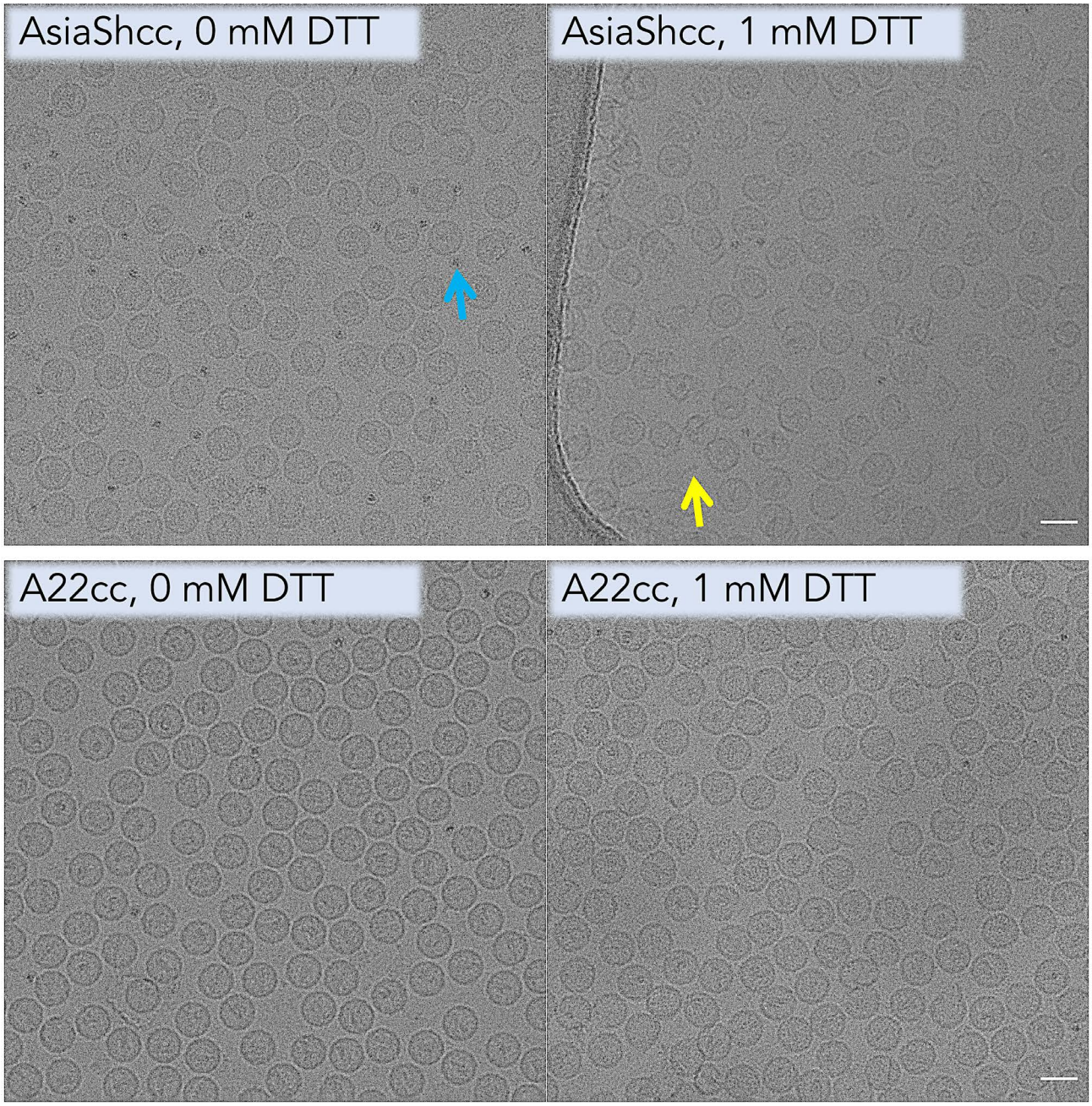


Figure 5.2: Exemplar micrographs from AsiaShcc and A22cc treated with and without DTT, as per labels. Scale bar = 300 Å. The blue arrow corresponds to an example mysterious flower-like particle, the yellow arrow points to an example pentamer side view.

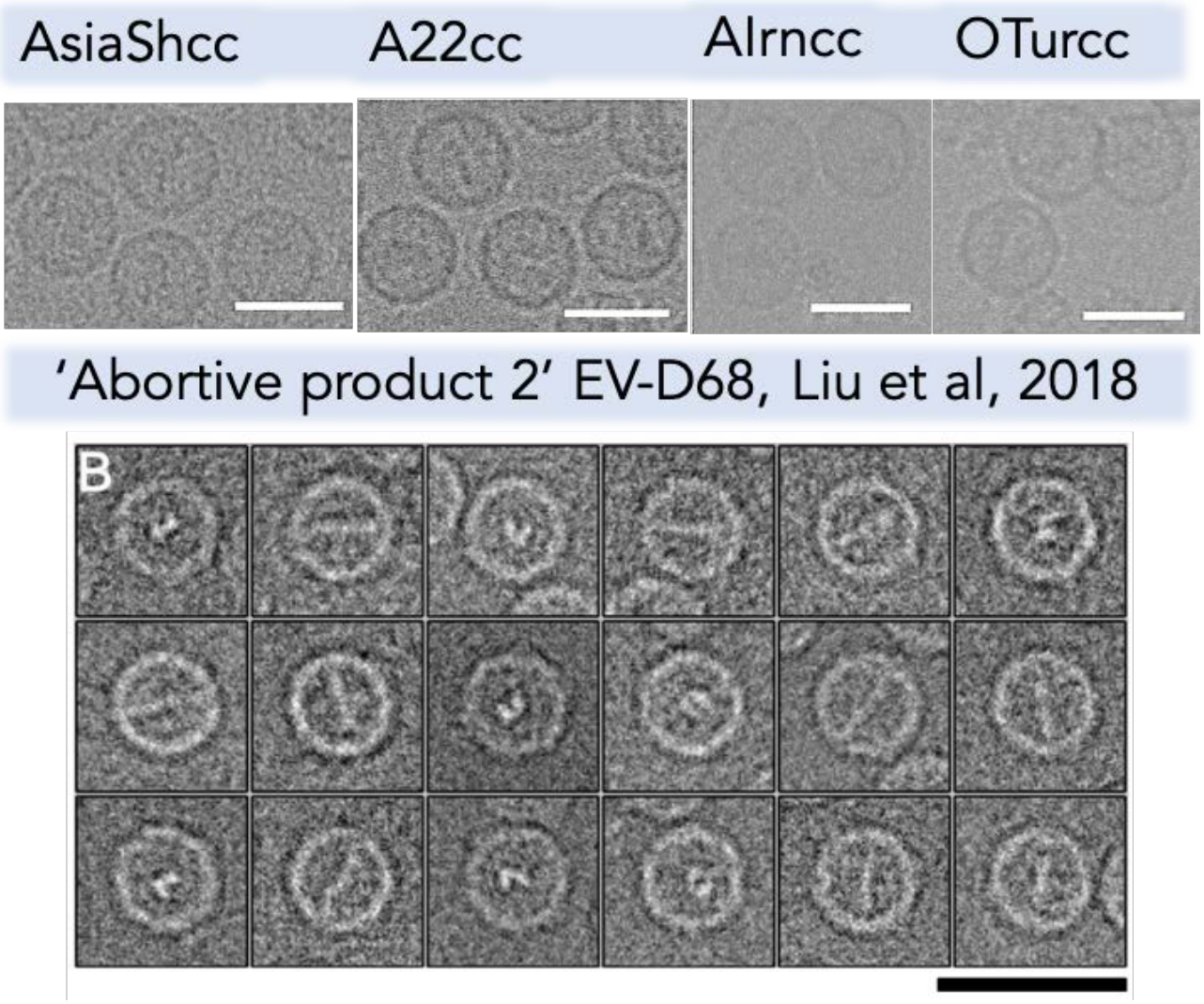


Figure 5.3: Top row: zoomed in regions of each micrograph shown in Figure 5.2, showing particles with bars. Sample labels are shown in the blue-grey boxes above each micrograph set. Bottom: extract from supplementary S11 figure B shown in Liu et al. (2018), used here under the 'Creative Commons' license. Note that for the 'Abortive product 2' dataset, particle contrast is inverted. Scale bars: white = 300 Å, black = 500 Å.

Also for all four samples, particles appeared to have a strange bar across them (Figure 5.2), consistent with observations from previously studied baculovirus-derived VLPs, Figure 5.3. Notably, this bar is somewhat reminiscent of recently reported data from a particularly acid-labile EV-D68 strain US/MO/1418947 virion at neutral pH, characterised using cryo-EM, where, amongst the many intermediate states reported, 7 % of selected particles showed 'rod-like densities' (Figure 5.3) (Liu et al., 2018). Thus, this bar is not expression system-dependent, and can be observed in wild-type virions. It has been hypothesised that this bar is an extra protein, perhaps from partly formed particles, appended to the outer capsid surface, consistent with mass imaging measurements (Young and Kukura, 2019), which showed a gain in mass corresponding to one pentamer per particle (personal communication, Elizabeth Fry). To try to discern what this may be, the particle-picking program Warp (Tegunov and Cramer, 2018) was trained to select particles with bars, no thorough checks were conducted to ensure preferential selection of this subpopulation, and future processing should address this.

Clear differences can be seen between DTT treated and untreated particles, with more distorted capsids and a higher population of edge-on pentamer structures in micrographs for 1 mM DTT treated sample, relative to untreated controls (Figure 5.2). Ice thicknesses appeared, from visual inspection, to be comparable between micrographs collected with and without DTT treatment, meaning that differences in capsid morphologies are more likely to be due to DTT treatment than disruption during vitrification.

Data Collection and Processing

Data processing and model interpretation for all VLP datasets is still ongoing, and, for this thesis, I will focus on a brief overview of the DTT untreated ('U') and 1 mM DTT treated ('D') AsiaShcc reconstructions.

AsiaShcc 0 mM DTT, U

A total of 1363 micrographs were collected in counting mode using a FIII camera at a nominal magnification of 75 kX. Particles were selected using Warp (Tegunov and Cramer, 2018) (see Chapter 2), and processed in Relion (Zivanov et al., 2018), see Figure 5.4. Reconstructions from the four highest resolution classes, below 4.3 Å (Figure 5.4), labelled U-(i-iv), will be discussed below.

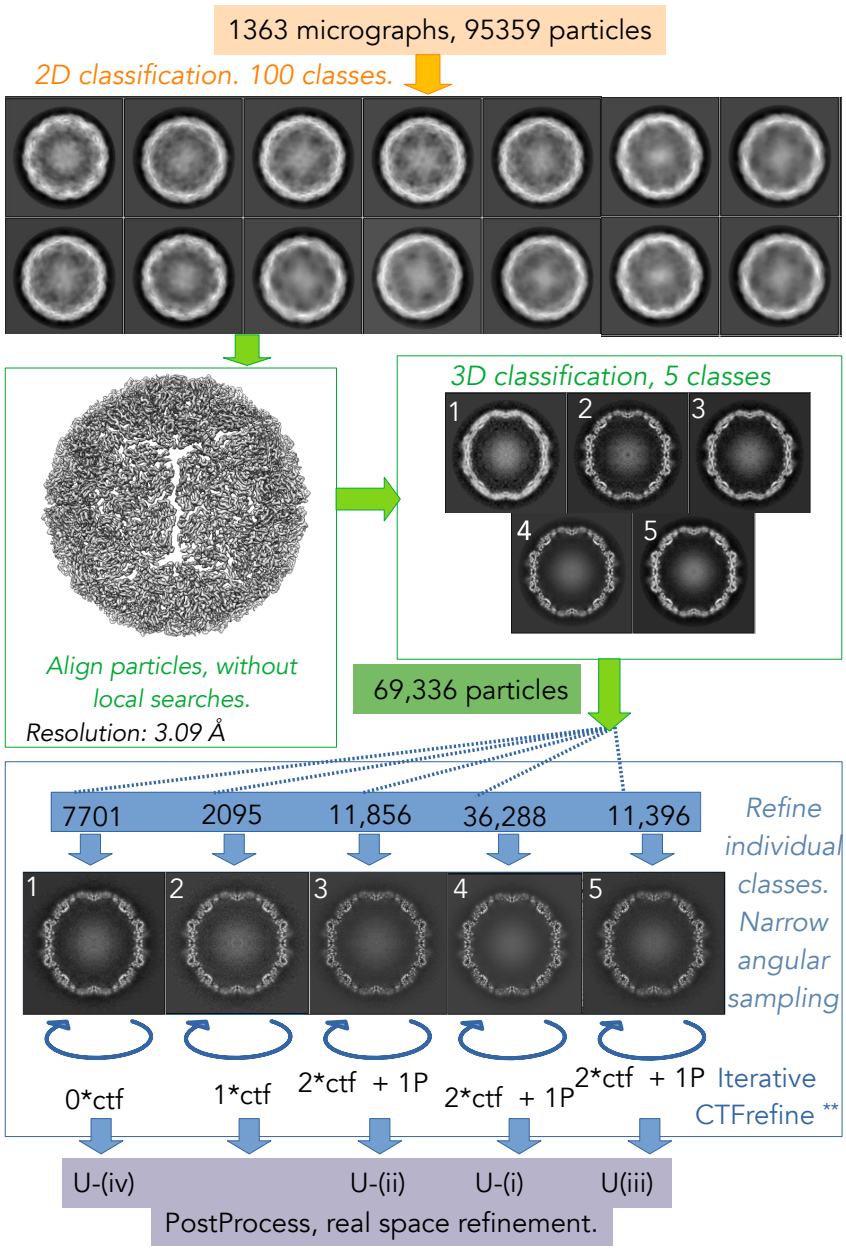


Figure 5.4: Processing workflow with Relion, following processing of movies with MotionCor2 and CtfFind4, and particle picking using Warp, see Chapter 2 (Zivanov et al., 2018; Zheng et al., 2017; Rohou and Grigorieff, 2015; Tegunov and Cramer, 2018). Briefly, extracted particles were classified into 100 2D classes, and classes with apparent secondary structure and with promising rotational and translational accuracies (around 1) were selected. Selected particles were then aligned before 3D classification into five classes. ** In this case, each particle set was then refined, and subjected to 0 (0*CTF) 1 (1*CTF) or two cycles of CTF refinement/refinement (2*CTF), and, where marked, a final round of polishing/refinement (1P). Typically, the decision was made to stop the CTF refinement/polishing cycles when the reported resolution and apparent map quality failed to improve. Icosahedral symmetry was imposed throughout. Resolution estimates and associated FSC curves, determined using gold-standard procedures, are shown in Figure 5.6(a) below (Scheres and Chen, 2012).

Table 5.1: Data processing statistics for VLP data. Reported resolution is based on the gold-standard FSC = 0.143 criterion, and associated FSC plots may be found in Figure 5.6 below.

Sample	AsiaShcc, 0 mM DTT				AsiaShcc, 1 mM DTT		
State	U-(i)	U-(ii)	U-(iii)	U-(iv)	D-(i)	D-(ii)	D-(iii)
Resolution/ Å	2.56	2.90	3.09	3.50	3.13	4.35	4.26
RMSD							
Bonds / Å	0.006	0.006	0.007	0.011	0.007	0.010	0.009
Angles / °	0.78	0.83	0.69	1.30	0.79	0.90	0.93
Protein atoms	4458	4448	4229	4231	4455	4766	4231
Ramachandran / %							
Favored	96.0	94.0	94.3	93.1	93.6	91.1	92.4
Allowed	4.00	6.00	5.70	6.55	6.30	8.80	7.50
Outliers	0.00	0.00	0.00	0.36	1.00	1.00	0.00
C-beta dev	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Rota-outliers	0.10	0.10	0.10	0.10	0.30	0.60	0.40
Model-to-map fit							
Phenix CC	0.856	0.838	0.827	0.820	0.800	0.781	0.828

AsiaShcc 1 mM DTT, D

AsiaShcc D data were collected and processed using identical conditions (75 kX nominal magnification, FIII detector) to the AsiaShcc U dataset above, as illustrated in Figure 5.5. Of the five separately refined classes, the three highest resolution datasets, below 4.3 Å (Figure 5.5), labelled (i-iii), will be described below.

Modelling

In the absence of a model for Asia1 in the PDB, the A22cc crystal structure, PDB: 4IV3 (Porta et al., 2013a), was initially rigid body fitted in Coot, before correction of the amino acid sequence and a second round of rigid body fitting (Emsley and Cowtan, 2004; Emsley et al., 2010). Real space refinement was conducted in Phenix with surrounding protomers, as described in Chapter 2, Section 2.5.5 (Adams et al., 2010, 2011). The central protomer was then rebuilt in Coot (Emsley and Cowtan, 2004; Emsley et al., 2010), followed by additional refinement cycles (Table 5.1). This model was then used as a starting point for the similar refinement of the DTT treated AsiaShcc structures (see below).

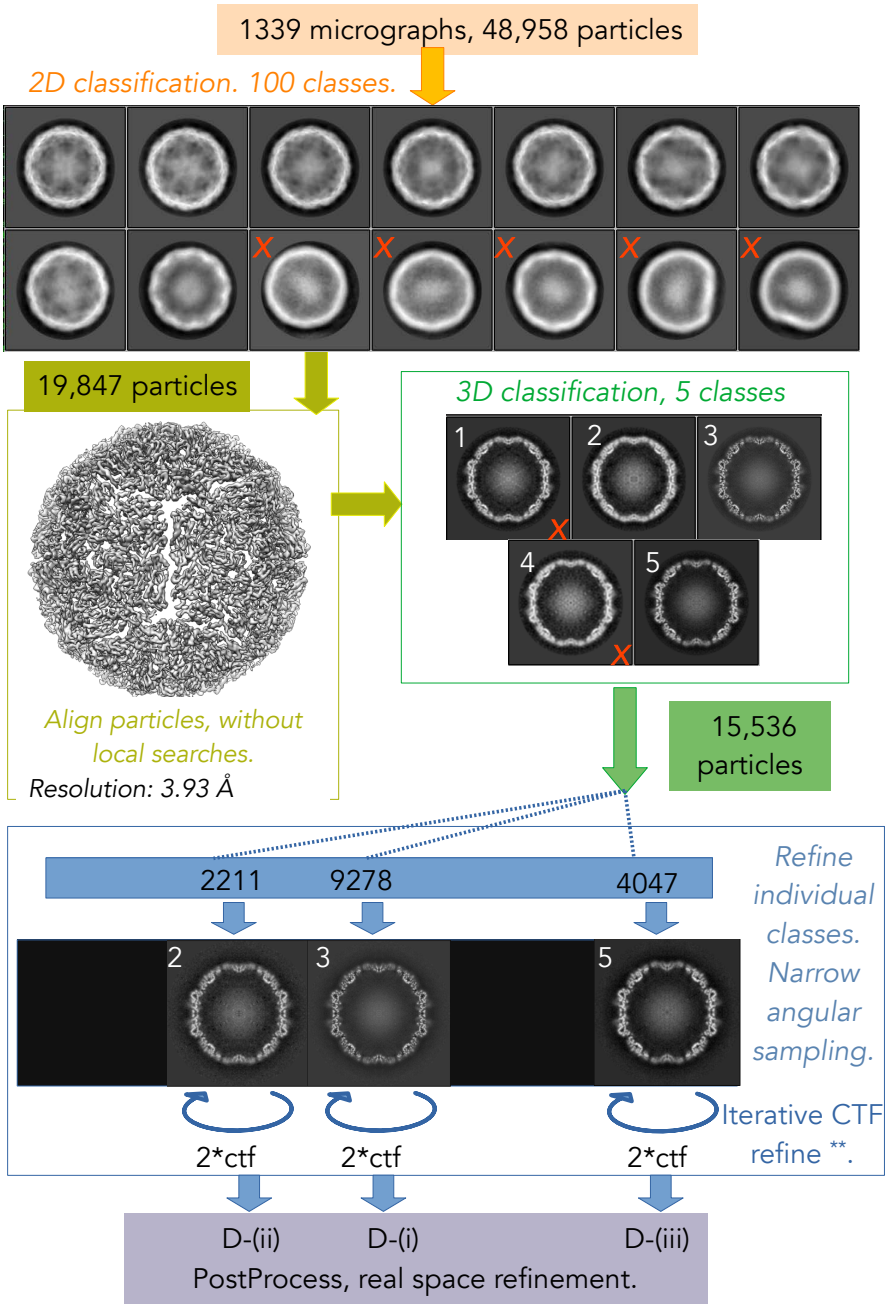


Figure 5.5: Processing workflow with Relion3, following processing of movies with MotionCor2 and CtfFind4, and Warp particle picking (Zivanov et al., 2018; Zheng et al., 2017; Rohou and Grigorieff, 2015; Tegunov and Cramer, 2018). As for the DTT untreated AsiaShcc particles above, extracted particles were classified into 100 2D classes, and classes showing signs of secondary structure, and with promising rotational and translational accuracies (around 1) were selected. Selected particles were then aligned before 3D classification into five classes. ** in this case, each particle set was then refined, and subjected to 1 (1*CTF) or two cycles of CTF refinement/refinement (2*CTF). Polishing was found to reduce map quality in the case of this dataset. Icosahedral symmetry was imposed throughout. FSC curves and associated resolution estimates are shown in Figure 5.6(b) below.

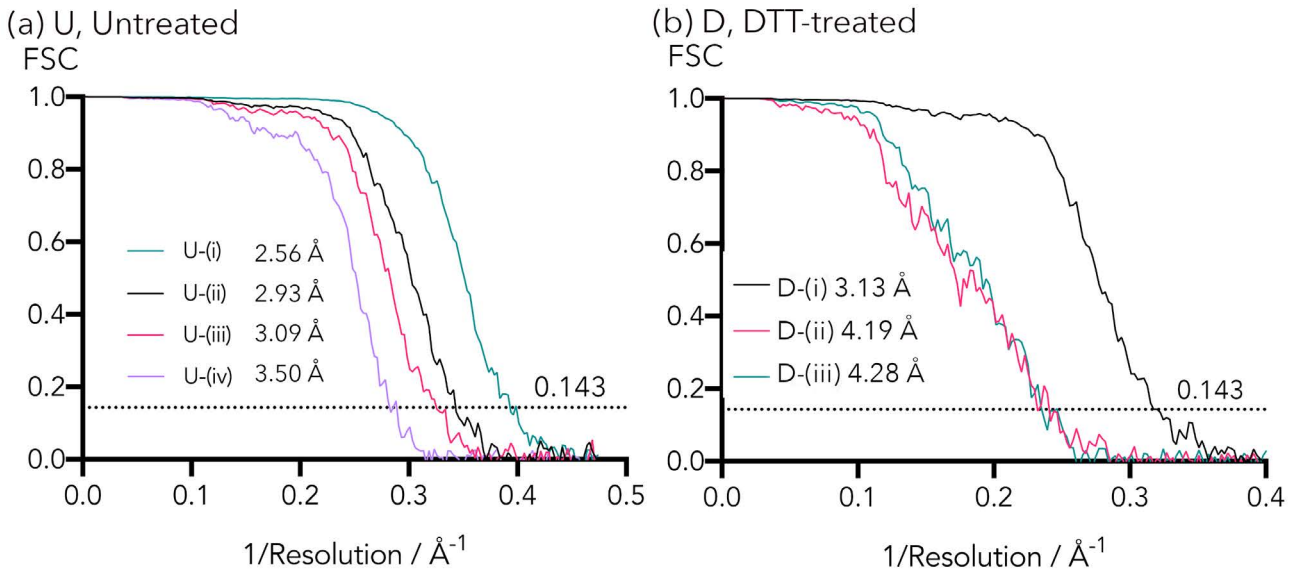


Figure 5.6: Gold-standard FSCs for (a) for the DTT untreated 'U' AsiaShcc classes; (b) for the 1 mM DTT treated AsiaShcc, 'D' classes (computed by Relion) (Scheres and Chen, 2012; Zivanov et al., 2018). Particle labels (i-iv) correspond to classes defined in Figures 5.4 and 5.5 for U and D datasets, respectively. Resolution estimates for each particle, evaluated using the threshold = 0.143 criterion, are listed next to the key. The dashed line represents FSC = 0.143.

Comparison of AsiaShcc Subsets

The four highest resolution 3D classes from AsiaShcc untreated with DTT will be used in this discussion, labelled U-(i-iv). The pixel size had been calibrated previously on a high-resolution small molecule structure collected shortly before our data collection (personal communication, Dimple Kotecha and Simonas Masiulis) as 1.055 Å/pix and this was used in processing.

For subsets U-(i-iv), particles appear to be expanded, with a gap between the VP2 2-fold axis helices, the site where we would expect the disulphide bridge to be formed (Figure 5.7). Although seemingly identical in terms of radius and appearance (Figures 5.7, 5.8), subtle differences between subsets U-(i-iv) can be seen at the 2-fold axes, with, for some classes, a bridge connecting VP2 helices from opposing protomers, which might be consistent with a S-S bridge between opposing C2093 residues (Figure 5.9, 5.10). For the most populated class, U-(i), this bridge cannot be seen, even at lower contour levels (Figure 5.9). Although this could be due to averaging effects of the relatively large number of probably heterogeneous particles in the U-(i) class, clear density can be seen for an alternative position of the C2093 side chain for U-(i) only (blue arrow, Figure 5.9). Contrastingly, the remaining classes, U-(ii) in particular, show evidence of at least a partially

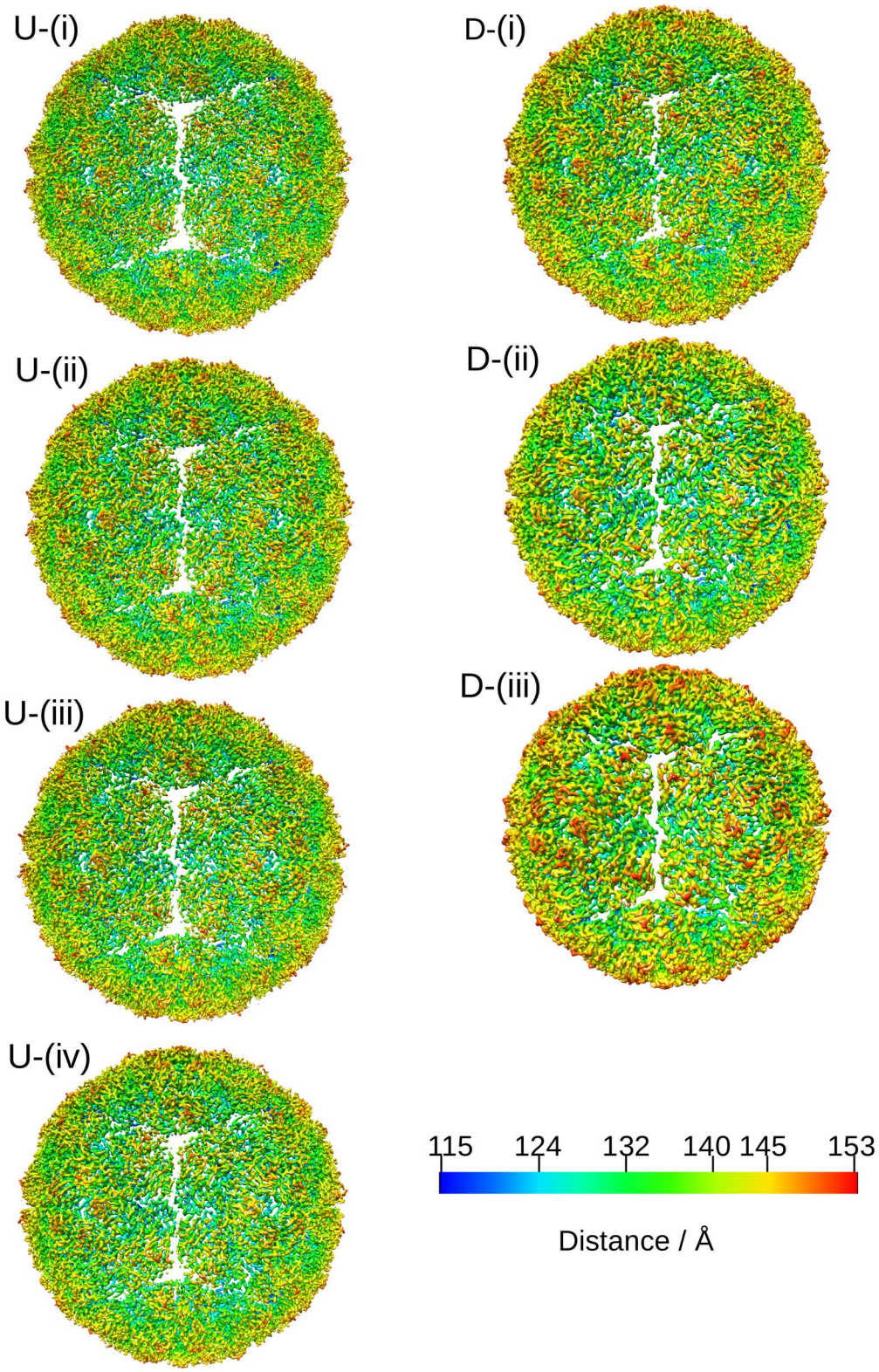


Figure 5.7: Reconstructed volumes from AsiaShcc with and without 1 mM DTT treatment, coloured by radius from blue to red, according to the scale bar shown. Particles are labelled according to the classes defined in the text.

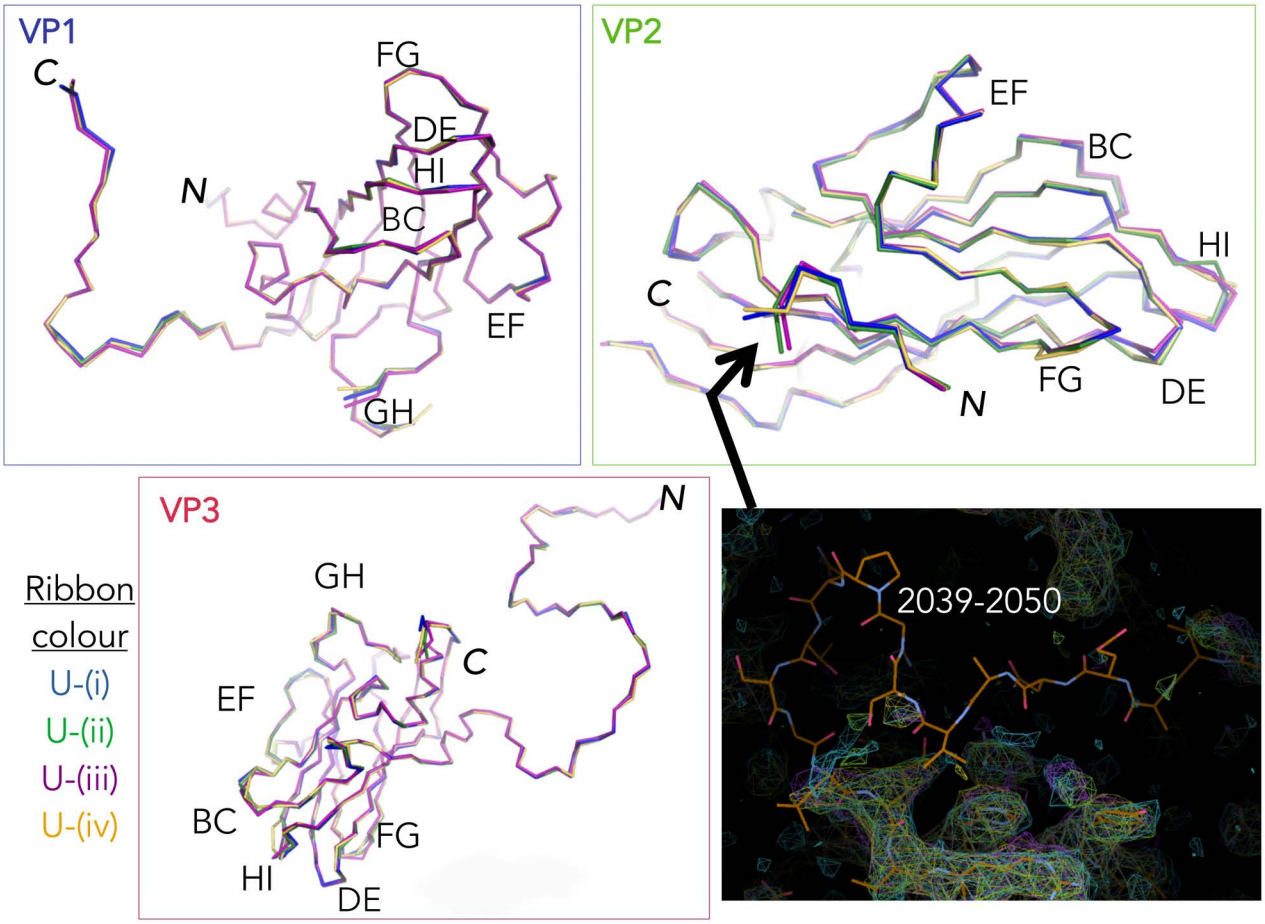


Figure 5.8: Overlaid models for all four classes, U-(i)-(iv), coloured according to the key. The black bottom right image shows overlaid density from all classes to demonstrate that, in every case, there is no density to support modelling of residues 2039-2050.

Table 5.2: Summary of the four AsiaShcc classes and associated particle numbers from processing. 'C-C C_α Separation' refers to the measured distance between two C2093 residues from two opposing protomers.

State	U-(i)	U-(ii)	U-(iii)	U-(iv)
Particles in final reconstruction	36288	11856	11396	7701
% of particles selected	52.3	17.1	16.4	11.1
Evidence of 2-fold S-S?	No	Yes	Yes	Yes
C-C C_α Separation / Å	6.6	6.8	7.0	6.4
Resolution (FSC = 0.143) / Å	2.56	2.93	3.09	3.50

formed disulphide bridge between adjacent 2093C residues, as summarised in Table 5.2 (Figure 5.9, 5.10).

Another interesting feature at the 2-fold axis is at residue H3140, which is highly conserved across many FMDV serotypes (Appendix, Figure 5.21(a)). For U-(i-iv), additional density close to the modelled H3140 may evidence an alternative His position (Figure 5.9, see later). The map quality at this potentially alternative position varies greatly between U-(i-iv), and does not correlate with reported resolution, nor particle numbers used in each reconstruction (see Table 5.2, Figure 5.9). Indeed, maps for classes U-(ii) and U-(iv) have the strongest density for this secondary position (2.93 Å, 3.50 Å respective reported resolutions and 11856, 7701 particles), contrasting U-(iii) (11396 particles, 3.09 Å resolution).

All four classes are similar at the 5-fold axis (Figure 5.11), where we would anticipate a ring of disulphide bridges (Acharya et al., 1989). For U-(i) and U-(ii), density close to C3007 may be consistent with a partially occupied S-S bridge (Figure 5.11). This is expected since not all Cys residues can engage in a S-S bond at any one 5-fold.

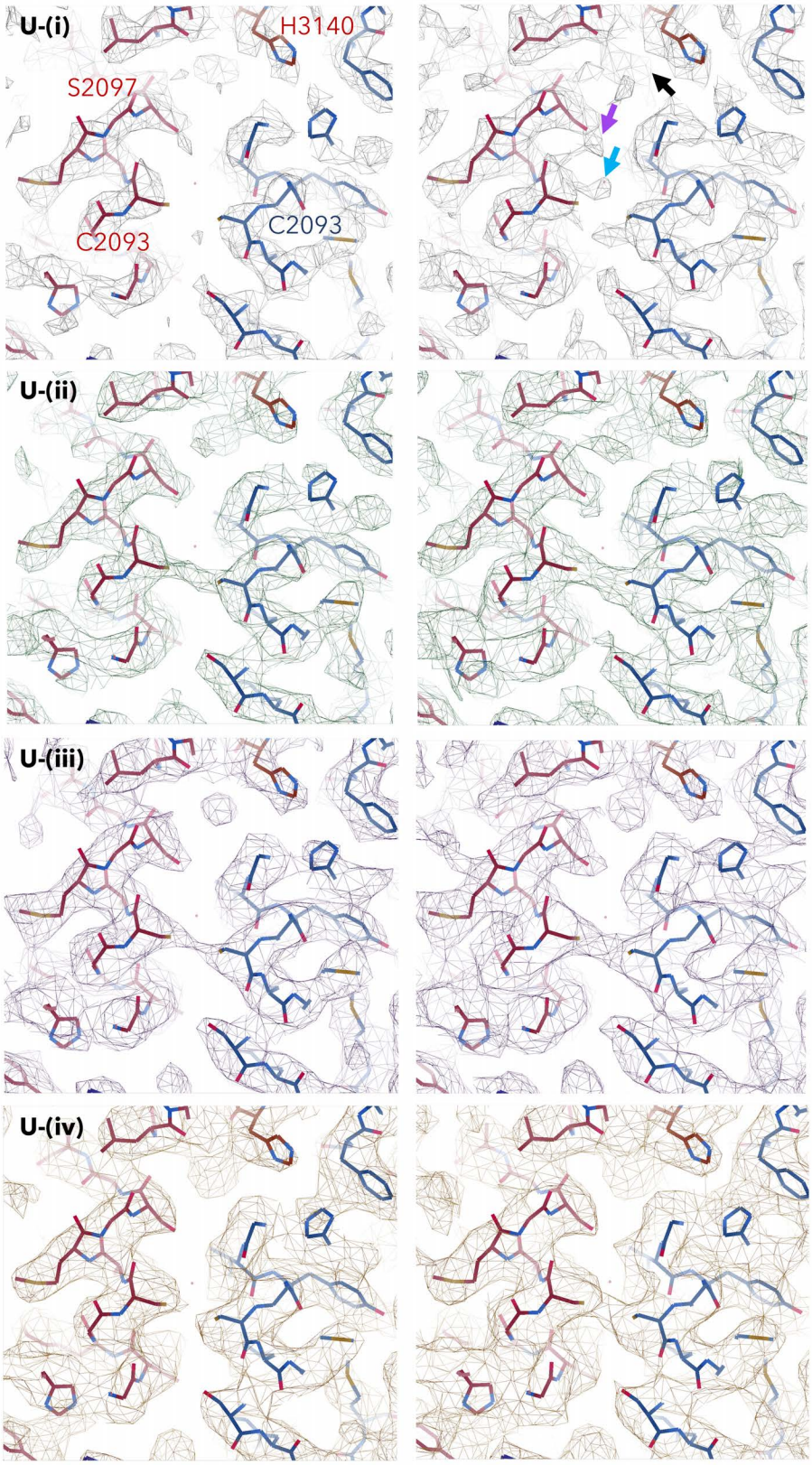


Figure 5.9: Density at the 2-fold axis for classes U-(i)-(iv) of AsiaShcc at two different contour levels (left high, right low). Pertinent residues are labelled only for the first panel for clarity, but apply to all images. Note that the model for U-(i) is shown in all images to aid comparison.

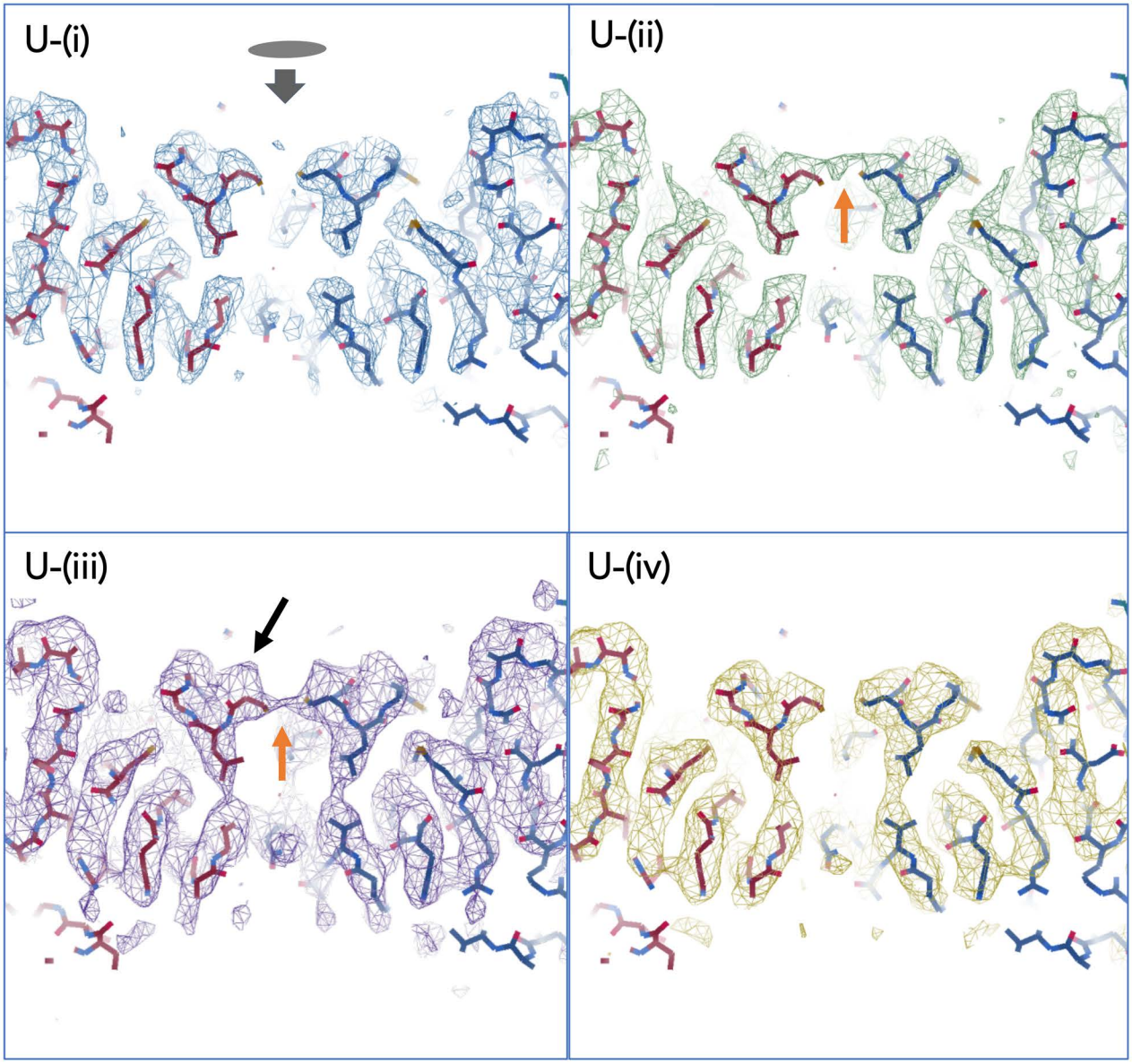


Figure 5.10: Side view of the 2-fold axis for U-(i-iv), with an ellipse and arrow demarcating the 2-fold axis channel marked in U-(i). The black arrow indicates a potential secondary position (raised towards the capsid surface) for the side chain of C2093, but this may simply be an artefact from the low contour level. The orange arrow indicates the S-S bridge, which is only resolved at extremely low contour levels for U-(iv).

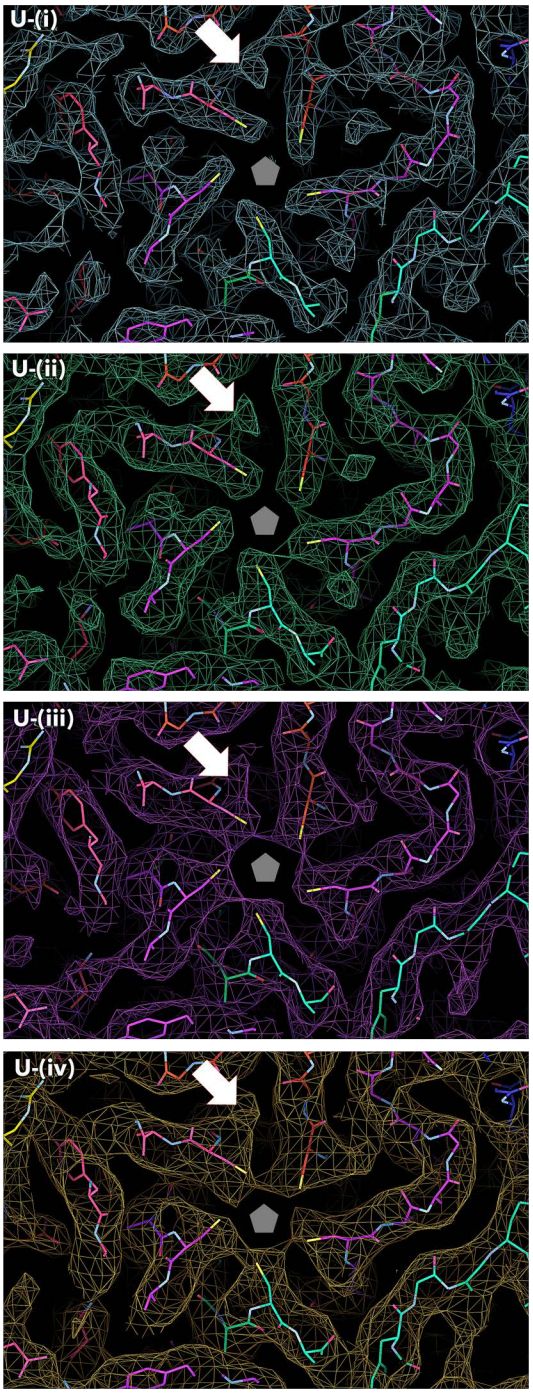


Figure 5.11: Maps of classes U-(i)-(iv) of AsiaShcc, as per labels viewed from the top of the 5-fold axis, level with residue C3007, which points to the centre of the channel. The white arrow indicates density unaccounted for by the model, and may represent and alternative position for C3007.

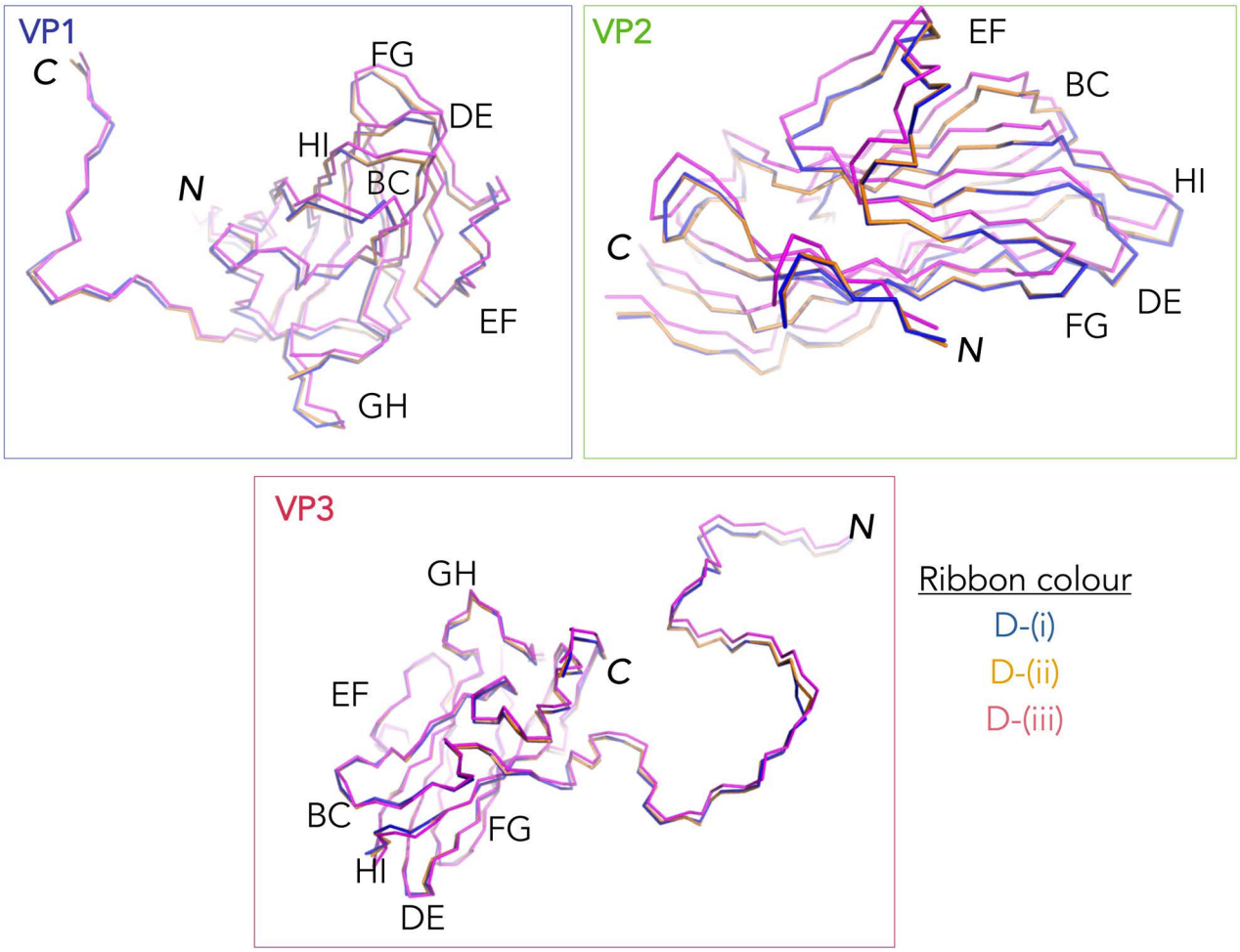


Figure 5.12: Ribbon representation of classes D-(i)-(iii) for each VP unit, with D-(i) in blue, D-(ii) in orange and D-(iii) in magenta.

DTT Treated Particles

For the 1 mM DTT treated AsiaShcc dataset, three distinct, well populated classes were identified (Table 5.3). These will be termed D-(i-iii) for the purposes of this discussion.

Differences between D-(i-iii) are greater than for U(i-iv). As seen in Figure 5.12, D-(i) and D-(ii) share a similar structure, whereas D-(iii), although comparable in terms of structure, is clearly shifted outwards by approximately 2 Å for every structural protein. This is consistent with the radially coloured reconstructions shown earlier in Figure 5.7, which show an expansion of ca. 2 %.

The most populated class, D-(i), containing 59.7 % of refined particles, shows clear density supporting a disulphide bridge spanning the two-fold axis (Figure 5.13). The approximately 50:50 split between broken and unbroken disulphide bridged particles selected from micrographs of the

Table 5.3: Data processing summary for 1mM DTT treated AsiaShcc

State	D-(i)	D-(ii)	D-(iii)
Particles in final reconstruction	9278	2211	4047
% of particles selected	59.7	14.2	26.0
Disulfide bridge at 2-fold?	Y	N	N
C-C C_α Separation / Å	7.2	7.4	7.7
Resolution (FSC) / Å	3.13	4.10	4.26

untreated sample may simply arise from selection bias during 2D classification, where particles showing significant distortions for the idealised icosahedral capsid shape were filtered out (see rejected 2D classes in Figure 5.5). Indeed, particles were noticeably more distorted in the DTT treated sample (Figure 5.2). As for U-(i-iv), D-(i-iii) represent a continuum; starting from a full set of intact S-S bridge at the 2-fold, in the case of D-(i), which is broken in D-(ii), leading to capsid expansion in state D-(iii). D-(iii), the lowest resolution of the classes (4.26 Å reported resolution, see Table 5.3), shows little helix structure, and opposing 2-fold helices appear to have slidden past each other, with one side moving further outwards (Figure 5.13). This expansion can be seen clearly with a slice at the 5-fold axis, shown in Figure 5.14.

There is little difference at the 5-fold axis between classes D-(i-iii) and U-(i-iv), excepting an appreciable expansion in D-(iii), Figures 5.14-5.15.

Another difference occurs at both the N and C-terminal ends of the VP1 GH-loop. Here, a distancing of 4 Å was measured between C_α atoms of residue A1157 (C-terminal end) for D-(i) versus D-(iii), and 2.5 Å at K1133 (N-terminal end), Figure 5.16, with D-(i)'s loop noticeably more flattened relative to that of D-(iii). However, map quality is poorer in this region, especially for the lower resolution D-(ii) and D-(iii) structures.

Comparison with Untreated Sample

Overall, as judged by RMS distances, models for classes U-(i-iv) and D-(i-iii) are very similar (Table 5.4), consistent with visual inspection (Figure 5.7). The positioning of the VP2 helices w.r.t. the 2-fold axis differs slightly, however, as seen in the separation between C-C residues shown in Tables 5.2 and 5.3. To quantify the positioning of the 2-fold helices, the angle (here called θ) between C_αs of residues S2097-*G2088*-G2088 (where italic text denotes the opposing protomer) was measured, as shown in Figure 5.17. The measurement of 84 ° for D-(iii) versus 65-73 ° for U-(i-iv) and D-(i, ii), for

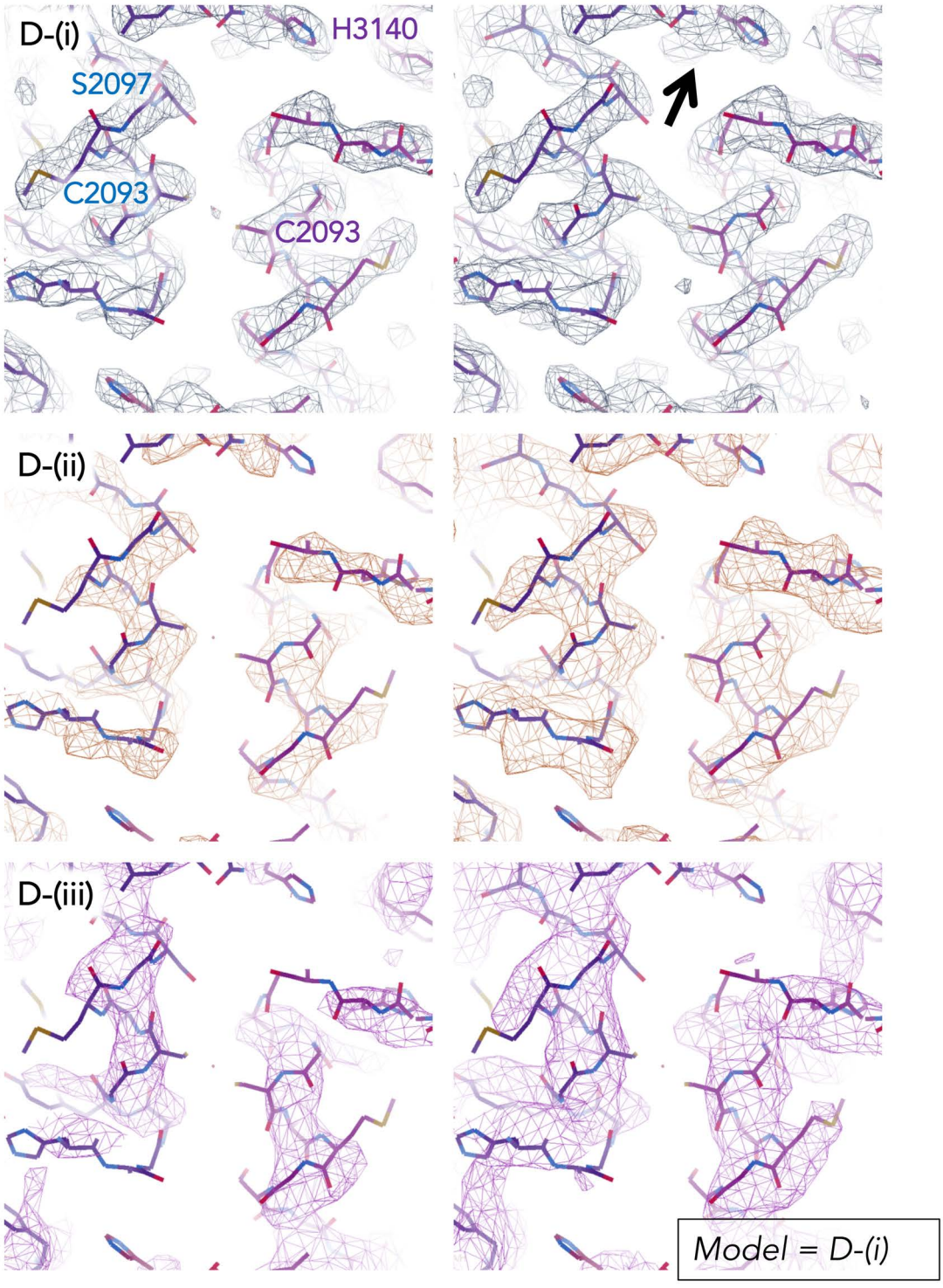


Figure 5.13: Density at 2-fold protomer interface for 1 mM DTT treated AsiaSh particles. Each row corresponds to one of the classes D-(i-iii), as per label and density is shown at high (left column) and low (right) contour levels. Density associated with movement of H3140 is shown with a black arrow. For clarity, the model derived for state D-(i) is shown in stick representation.

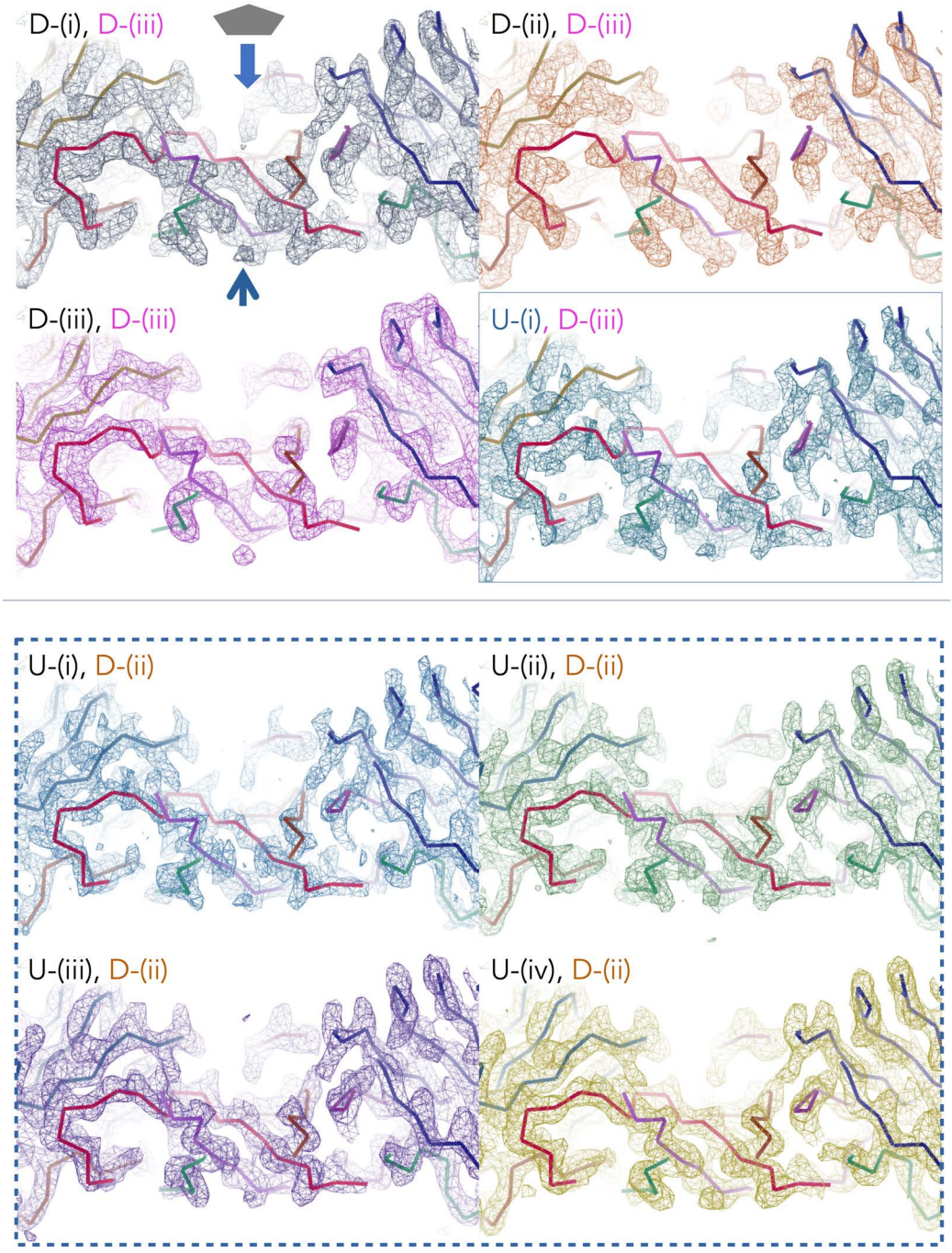


Figure 5.14: Cross-Section of the 5-fold axis for maps D-(i)-(iii) and U-(i)-(iv). The 5-fold axis is indicated by a pentamer symbol on the first image (D-(i), D-(iii)). For each image, the first label, in black text, corresponds to the map shown, and the second label, the model shown. The first four images show the more expanded D-(iii) model as a point of comparison. The blue arrow shows a spot of density that was observed plugging the base of the five-fold axis, something that is commonly observed for most FMDV VLP structures. The blue boxes correspond to the DTT untreated classes, with the bold dashed box using the model of D-(ii) as a point of comparison to demonstrate that there is little expansion between U-(i-iv) and D-(i-ii).

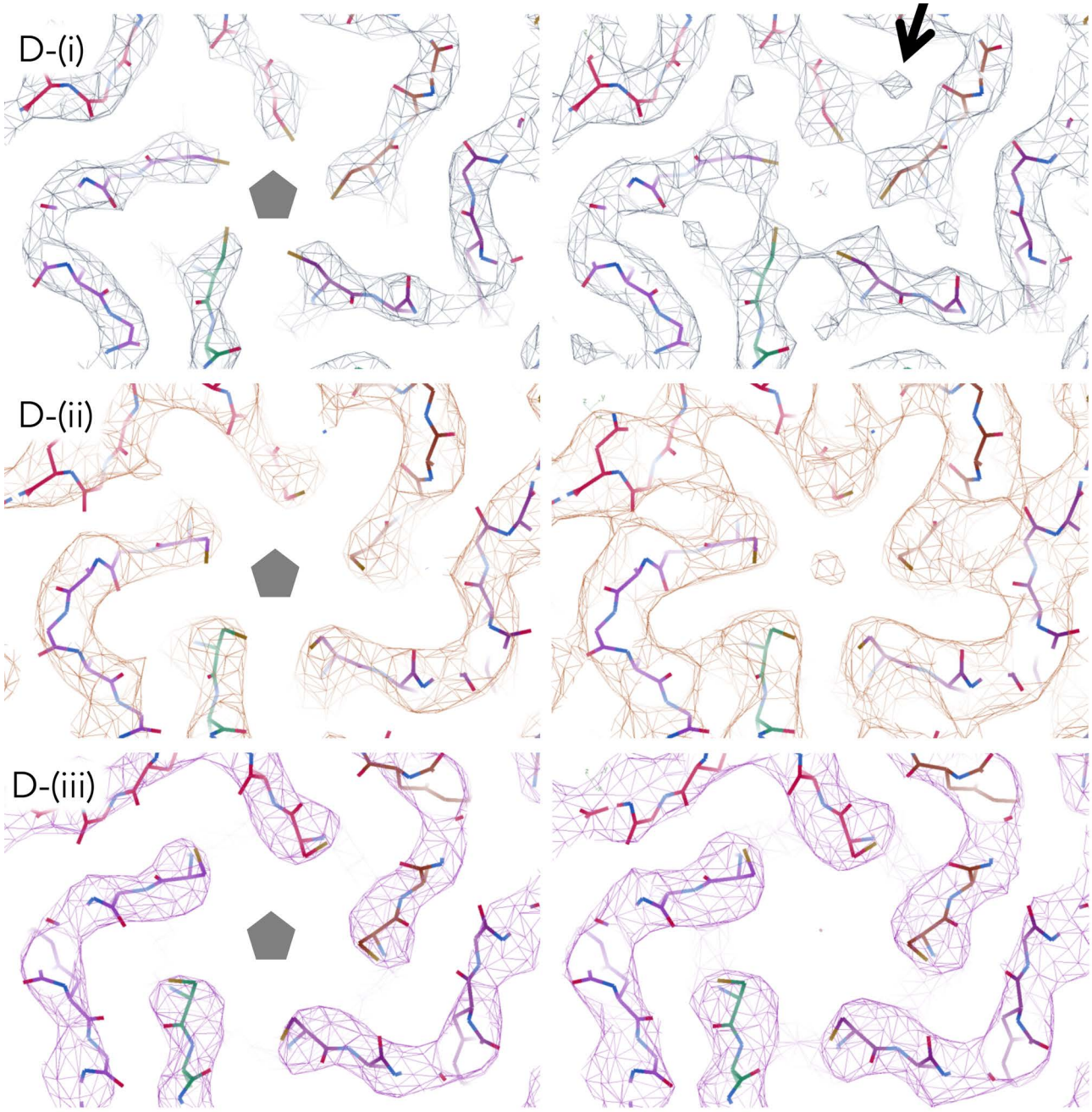


Figure 5.15: View of the top of the 5-fold axis, at the N-terminal portion of VP3, with C3007 from each protomer contributing to a 5-fold vertex shown at high (left) and low (right) contour levels. Labels D-(i)-(iii) correspond to each row of images, and the five-fold axis centre is marked with a pentamer for the high contour images. The arrow shows a small extra blob of density that may correspond to an alternative position for residue C3007, as also seen in the untreated dataset, Figure 5.11 above.

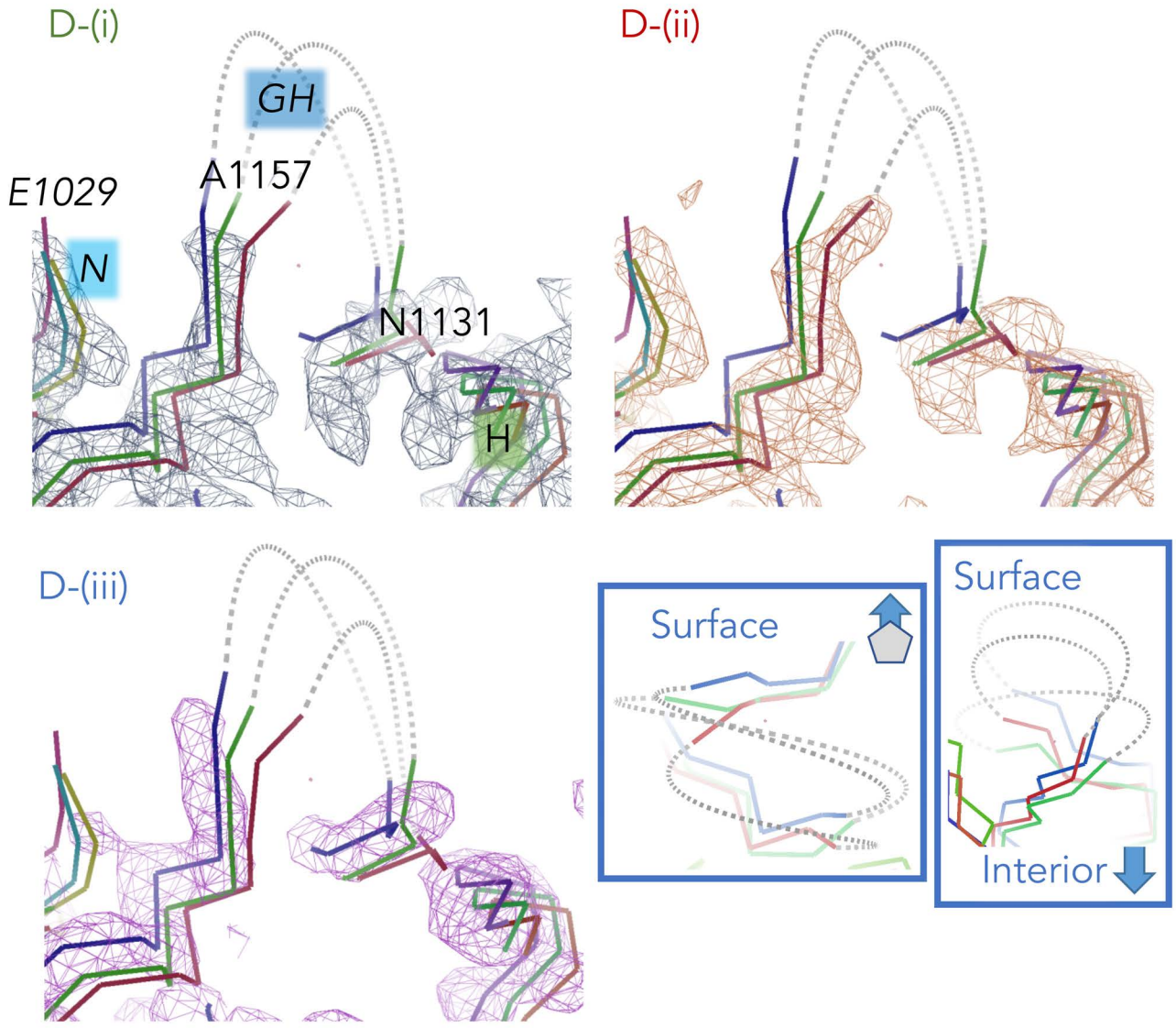


Figure 5.16: Green ribbons correspond to D-(i), red D-(ii) and blue D-(iii). For orientation purposes, key residues are shown on the image for D-(i). The H-strand and N-terminal region of VP1 from an adjacent protomer are also marked, with box colours corresponding to their associated structural unit (i.e. blue VP1, green VP2). Note that the missing portion of the loop is shown as a dashed line, and does not reflect the length of the 'real' GH-loop.

Table 5.4: RMS distances (Å) determined by SHP between protomers of Asiahcc U and D, as labelled in the table (Stuart et al., 1979). The number of equivalences for each VP was the same for each VLP, as follows: VP1, 162 ; VP2, 176 and VP3, 219. Cells are coloured by RMS distance (Å) from short distances (pale red) to far distances (pale indigo).

VP1	D-(i)	D-(ii)	D-(iii)	U-(i)	U-(ii)	U-(iii)	U-(iv)
D-(i)		0.24	0.48	0.45	0.44	0.44	0.50
D-(ii)	0.24		0.43	0.44	0.42	0.43	0.49
D-(iii)	0.48	0.43		0.47	0.47	0.47	0.52
U-(i)	0.45	0.43	0.47		0.21	0.31	0.35
U-(ii)	0.43	0.41	0.46	0.21		0.28	0.37
U-(iii)	0.43	0.42	0.47	0.31	0.28		0.35
U-(iv)	0.50	0.48	0.52	0.35	0.37	0.34	

VP2	D-(i)	D-(ii)	D-(iii)	U-(i)	U-(ii)	U-(iii)	U-(iv)
D-(i)		0.20	0.44	0.39	0.41	0.38	0.44
D-(ii)	0.20		0.45	0.43	0.42	0.40	0.50
D-(iii)	0.43	0.45		0.53	0.47	0.55	0.50
U-(i)	0.39	0.43	0.52		0.31	0.36	0.34
U-(ii)	0.41	0.42	0.50	0.31		0.31	0.38
U-(iii)	0.38	0.40	0.47	0.35	0.31		0.38
U-(iv)	0.44	0.49	0.54	0.34	0.38	0.38	

VP3	D-(i)	D-(ii)	D-(iii)	U-(i)	U-(ii)	U-(iii)	U-(iv)
D-(i)		0.28	0.46	0.48	0.48	0.47	0.54
D-(ii)	0.28		0.47	0.5	0.51	0.51	0.53
D-(iii)	0.46	0.47		0.57	0.56	0.56	0.59
U-(i)	0.48	0.5	0.57		0.28	0.36	0.44
U-(ii)	0.47	0.51	0.55	0.28		0.33	0.43
U-(iii)	0.46	0.51	0.56	0.37	0.34		0.42
U-(iv)	0.53	0.54	0.59	0.45	0.44	0.42	

State	θ / °
U-(i)	67.8
U-(ii)	71.5
U-(iii)	64.8
U-(iv)	66.0
D-(i)	71.9
D-(ii)	72.8
D-(iii)	83.8

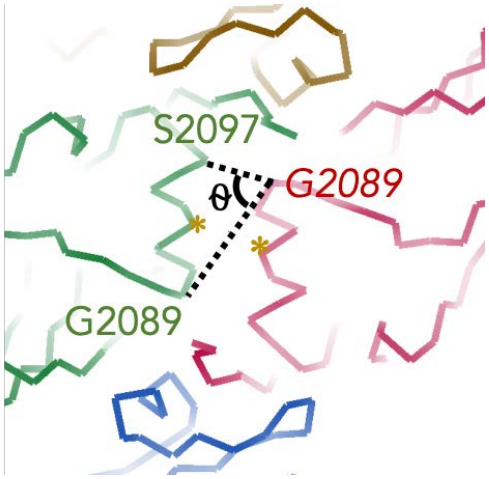


Figure 5.17: (Right) Model of 2-fold showing the angle, denoted θ, measured for each state described above and residues whose C_αs were used to define this angle. The C2093 S-S bridging residues are shown as gold asterisks for reference. VP2 for protomer 1 is coloured green, and, for protomer 2 in red. (Left) list of angles measured for each expansion state.

instance, corroborate the observed sliding of the VP2 helix for D-(iii), and demonstrates that the majority of classes have VP2 2-fold helices positioned in very similar places. Given the crude nature of this metric, slight variations (for instance, between classes U-(i)-(iv) and D-(i)-(ii)) are probably not significant. It is interesting, however, that classes where the S-S bond is most clearly formed, U-(ii) and D-(i), show comparable θ angles (Figure 5.17).

For all classes, we see some conformational switching at the highly conserved residue H3140 (see the additional blob of density in Figures 5.9,5.13). The two positions are shown in Figure 5.18. In the primary position, H3140 contacts S2216, distance approximately 3 Å. In the second position, modelled with thin lines, H3140 contacts either R2060 from the adjacent protomer, or L2214-O from the same protomer, depending on the position of the indole ring. Both residues are conserved across A, Asia1 and O serotypes (see Appendix). For A1061, the corresponding residue, H2142, appears to make no contacts with either conserved residue.

Provisional Characterisation of the 2-fold Interface

To identify key interface residues at the 2-fold axis to aid with sequence comparison, Pisa analysis (Krissinel and Henrick, 2007) was conducted using VP2/VP3 of two opposing protomers for each of the expansion classes, U-(i-ii), with disordered VP2 residues 2039-2050 omitted (see Figure 5.19). Results are shown in Figure 5.20, but should be treated with extreme caution, since, even for

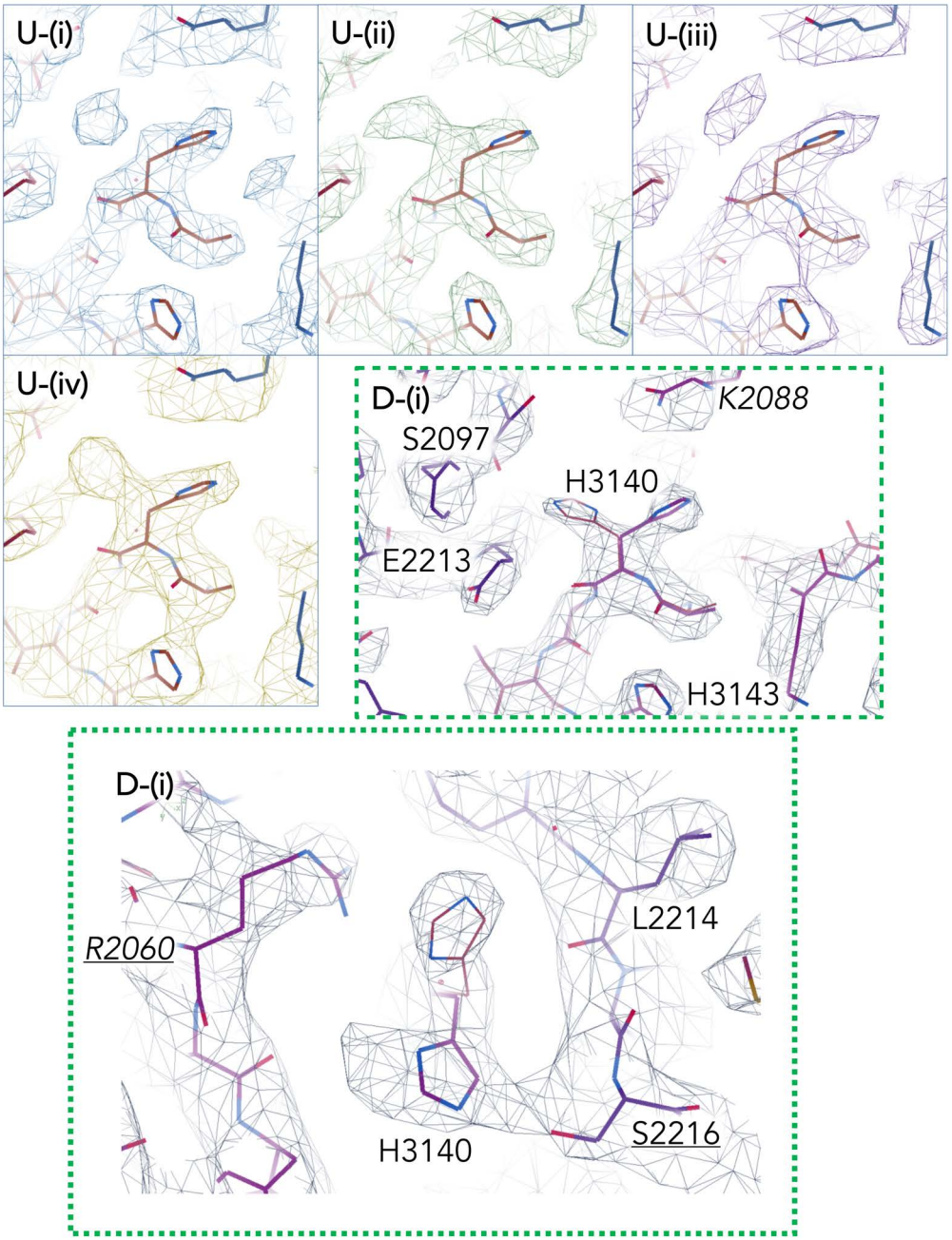


Figure 5.18: Blue boxes: Multiple conformation of H140 shown from the side view for maps U-(i-iv). Green dashed boxes show the two positions of H3140 mapped for D-(i), which is almost identical to U-(i), as shown by the use of the model for D-(i) in the preceding blue boxes. The secondary position for H3140 is shown as narrow sticks. Residues belonging to the adjacent protomer to the H3140 are shown in italics. The bottom green box shows H3140 from the top view, and contacting residues for each position are underlined.

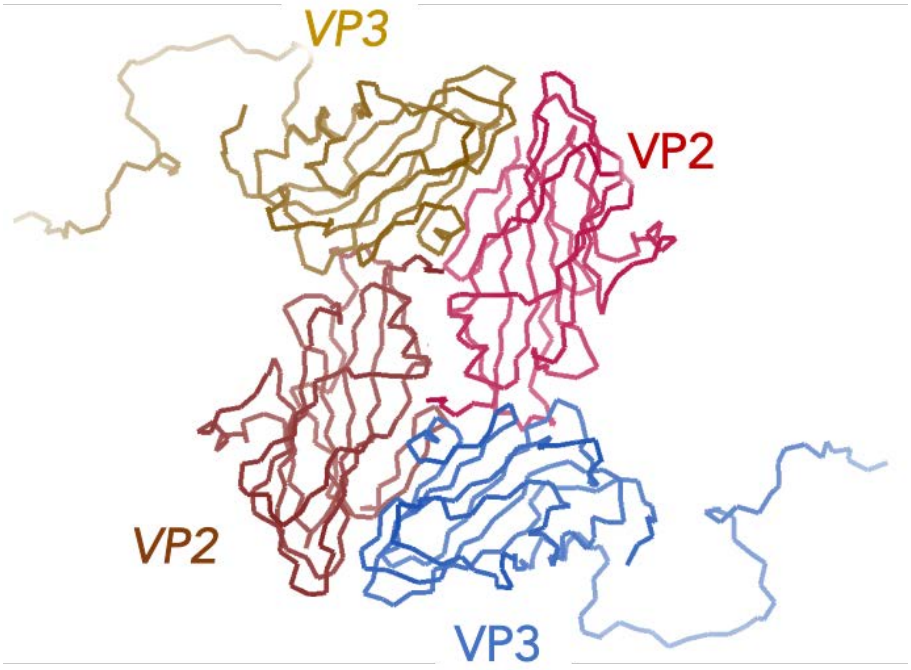


Figure 5.19: Example of capsid proteins used for input into pisa interfaces analysis (this is from model U-(i)). Units labelled in italics are from one protomer, those in normal type from the other.

these two high resolution classes, side chain density is weak at the 2-fold interface and side chain positioning cannot be always trusted. A more robust approach would have been to ensure that all low-confidence residues occupy a similar pose, or are removed entirely. Nevertheless, the results give some indication of the key residues that can form stabilising interactions close to the 2-fold axis.

In parallel, to try to pinpoint unconserved residues at the 2-fold interface between FMDV serotypes, sequences were aligned with ClustalW for FMDV sero/sub-types AIrn, OTur, A22cc, AsiaShcc, A1061 and O1K (see Appendix) (Sievers et al., 2011). Variable interface residues shown in bold in Figure 5.21.

As may be seen in Figure 5.21(a), AIrncc and AsiaShcc differ at residue 2086, with a longer Glu residue in the latter (purple asterisk). AsiaShcc also contains a bulkier Met at position 2095 (blue box), but this seems insignificant; the residue faces away from the inter-protomer contact helix, and is well accommodated for by a K2172A switch in the adjacent VP2 chain. We see an additional difference at the C-terminal VP2 residue of OTurcc, which is a bulkier Phe, rather than Leu for all non-O serotypes (green), but its significance is unclear.

Residue 3137 at the C-terminal end of the VP3 EF-loop, located at the 2-fold, in close proximity to the C-terminal portion of VP2 is also variable (Figure 5.21(a)). For A22cc and AIrncc, a lysine occurs in this position, while, OTurcc possesses a histidine and thus is more pH sensitive to contacts with the adjacent VP2 C-terminus.

HB	VP2-B	Dist / Å	VP2-A
1	b:GLN 57[NE2]	3.66	2:ARG 54[O]
2	b:THR 53[OG1]	2.99	2:GLN 57[OE1]
3	b:SER 216[OG]	2.54	2:LYS 88[O]
4	b:CYS 93[SG]	3.71	2:GLY 89[O]
5	b:ARG 60[NH1]	3.3	2:GLU 213[OE2]
6	b:ARG 54[O]	3.64	2:GLN 57[NE2]
7	b:GLN 57[OE1]	2.99	2:THR 53[OG1]
8	**b:LYS 88[O]**	2.56	2:SER 216[OG]
9	b:GLY 89[O]	3.74	2:CYS 93[SG]
10	b:GLU 213[OE2]	3.2	2:ARG 60[NH1]
SB			
1	b:ARG 60[NH1]	3.3	2:GLU 213[OE2]
2	b:GLU 213[OE2]	3.2	2:ARG 60[NH1]

K2088 all except for A106I, where H.

HB	VP2-B	Dist / Å	VP2-A
1	2:GLN 57[NE2]	3.32	b:ARG 54[O]
2	2:CYS 93[SG]	3.08	b:CYS 93[SG]
3	2:ARG 54[O]	3.21	b:GLN 57[NE2]
SB			
1	2:HIS 21[ND1]	2.64	b:GLU 213[OE2]
2	2:HIS 21[NE2]	3.86	b:GLU 213[OE2]
3	2:GLU 213[OE2]	3.13	b:HIS 21[ND1]
4	2:GLU 213[OE2]	3.94	b:ARG 60[NH2]

HB	VP3-A	Dist / Å	VP2-A
1	**3:HIS 137[ND1]**	3.15	2:SER 216[O]
2	3:SER 162[N]	3.64	2:TYR 168[OH]
3	3:PRO 127[O]	3.53	2:GLN 170[NE2]
4	**3:HIS 137[O]**	3.32	2:SER 216[N]
5	3:TYR 160[O]	2.97	2:ARG 167[NH2]
6	3:SER 162[O]	3	2:ARG 167[NE]
7	3:SER 162[O]	3.09	2:ARG 167[NH2]

L316I. Less bulky V for A/m.
H3137 K for A22cc and A/m; A for OTur

HB	VP3-A	Dist / Å	VP2-A
1	3:ASP 165[OD1]	2.31	2:ASN 47[ND2]
2	3:TYR 160[O]	3.68	2:SER 49[N]
3	**3:LEU 161[O]**	3.19	2:TYR 100[OH]
4	3:ALA 163[O]	2.99	2:ARG 167[N]
5	3:SER 162[O]	3.76	2:TYR 168[OH]
6	**3:HIS 137[O]**	3.12	2:SER 216[N]
7	3:SER 162[OG]	3.67	2:ASN 47[O]
8	3:SER 162[N]	3.63	2:TYR 168[OH]
SB			
1	**3:HIS 137[ND1]**	3.76	2:GLU 218[OE1]
2	**3:HIS 137[NE2]**	3.97	2:GLU 218[OE1]

HB	VP2-B	Dist / Å	VP3-A
1	b:THR 25[N]	2.78	3:GLU 145[O]
2	b:THR 25[OG1]	2.57	3:GLU 145[O]
3	b:SER 24[OG]	3.78	3:GLU 145[OE1]
4	b:SER 28[OG]	3.16	3:ASP 147[OD2]
5	**b:SER 194[OG]**	3.45	3:HIS 190[O]
6	b:ASN 114[ND2]	3.49	3:GLY 191[O]
7	b:TYR 200[OH]	3.74	3:ARG 71[NE]
8	b:SER 28[OG]	3.21	3:LYS 117[NZ]
9	b:THR 26[OG1]	3	3:ARG 119[NH1]
10	b:THR 25[O]	3.77	3:ARG 119[NH2]
11	b:SER 24[OG]	3.35	3:HIS 143[NE2]
12	**b:SER 194[OG]**	2.56	3:LYS 192[N]
SB			
1	b:LYS 63[NZ]	2.84	3:GLU 145[OE1]

HB	VP2-B	Dist / Å	VP3-A
1	b:MET 154[SD]	3.38	3:LYS 117[NZ]
2	**b:SER 194[OG]**	3.39	3:LYS 192[N]
3	b:TYR 200[OH]	2.86	3:ARG 71[NE]
4	b:ARG 60[NH2]	2.37	3:HIS 140[O]
5	b:ASN 202[ND2]	2.82	3:GLU 145[OE2]

S2194 unique to AsiaSh

U(i)

1	Atoms	Residues	Solvent SA / Å²	2	Atoms	Residues	Solvent SA / Å²	Interface area / Å²	ΔiG kcal/mol	ΔiG P-value	Hbonds	Salt bridges	S-S bridges
VP3-A	91	22	14565	VP2-B	87	27	10582	878.0	-2.7	0.798	5	0	0
VP3-A	73	20	14565	VP2-A	73	19	10581	715.9	-9.9	0.248	7	0	0
VP2-B	55	17	10582	VP2-A	53	16	10581	564.3	-5.1	0.369	7	3	0
										TOTAL	19	3	

U(ii)

1	Atoms	Residues	Solvent SA / Å²	2	Atoms	Residues	Solvent Surface	Interface area / Å²	ΔiG kcal/mol	ΔiG P-value	Hbonds	Salt bridges	S-S bridges
VP2-B	122	36	12238	VP3-A	114	28	14421	1090	-4.7	0.731	12	1	0
VP2-A	87	21	11231	VP3-A	82	25	14421	833	-12	0.202	8	2	0
VP2-A	70	20	11231	VP2-B	70	20	11238	692.6	-9.3	0.156	3	4	0
										TOTAL	23	7	

Figure 5.20: Summary of output from Pisa for U-(i) and U-(ii) (Krissinel and Henrick, 2007). The top tables show the overview of the three interfaces (inter-protomer VP2/VP3; interprotomer VP2/VP2 and intra-protomer VP2/VP3). Note that the S-S bridge for U-(ii) was not detected, as highlighted in the orange box. Residues that are not conserved are shown in bold. VP2-A/VP2-B and VP3-A/VP3-B should be identical, and this was ensured by applying the 'copy NCS chain' function in Coot prior to upload to Pisa (Emsley and Cowtan, 2004; Krissinel and Henrick, 2007; Emsley et al., 2010).

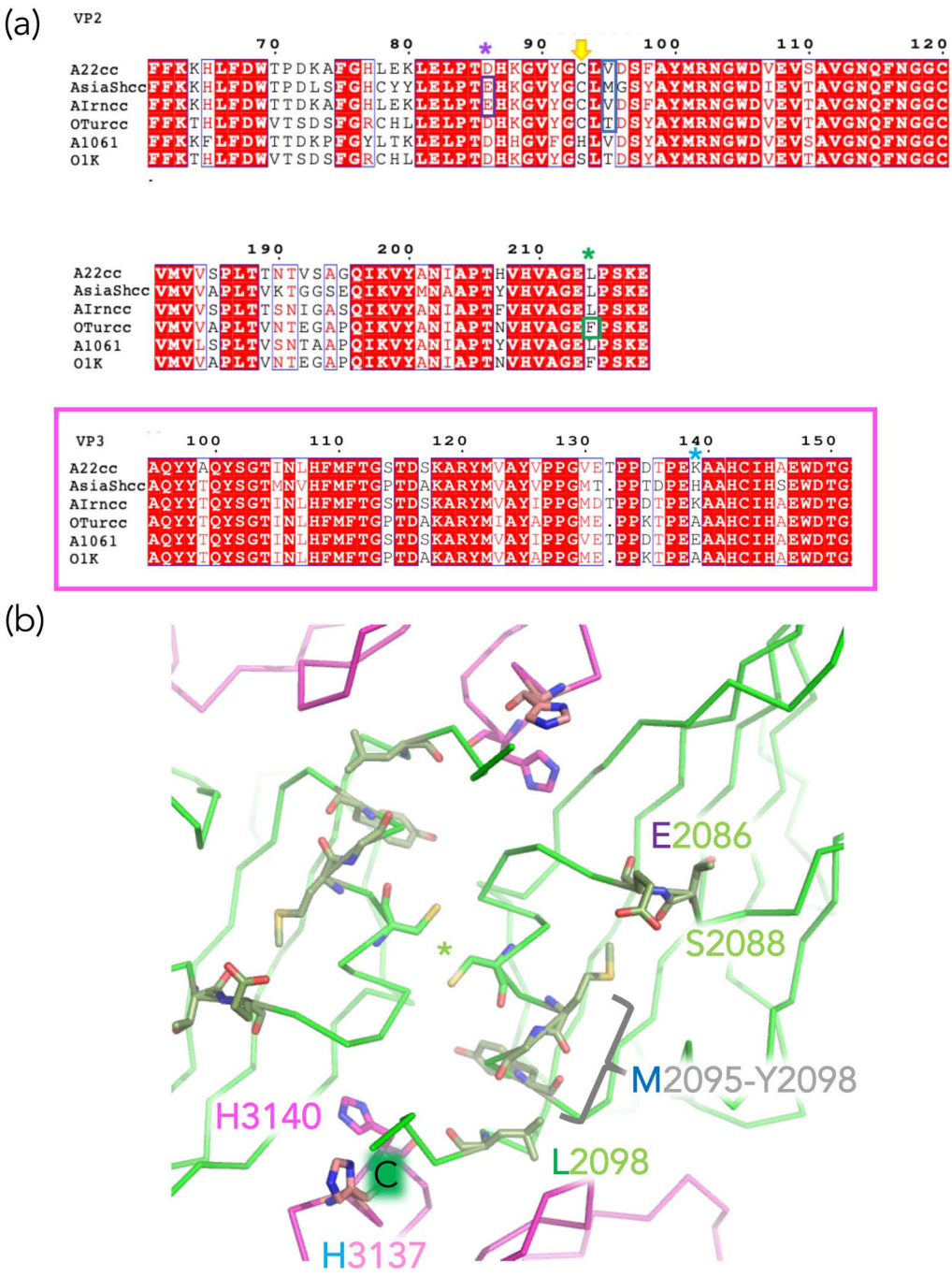


Figure 5.21: (a) Portion of sequence alignments for selected serotypes of VP2 and VP3. The VP3 except is shown in a pink box for clarity. A full sequence alignment for VP1-3 may be found in the Appendix. The position of the engineered S-S bridge for the 'cc' VLP sequences is shown as a yellow arrow. Numbering is based on A22cc, but is very similar for all types shown. (b) Position of some mutated residues in the context of the 2-fold axis for our U-(i) model. Residues from one protomer are labelled, with residue name colours corresponding to markers in (a) where appropriate. The 2-fold bridge site is shown as an asterisk, and C2093 is shown.

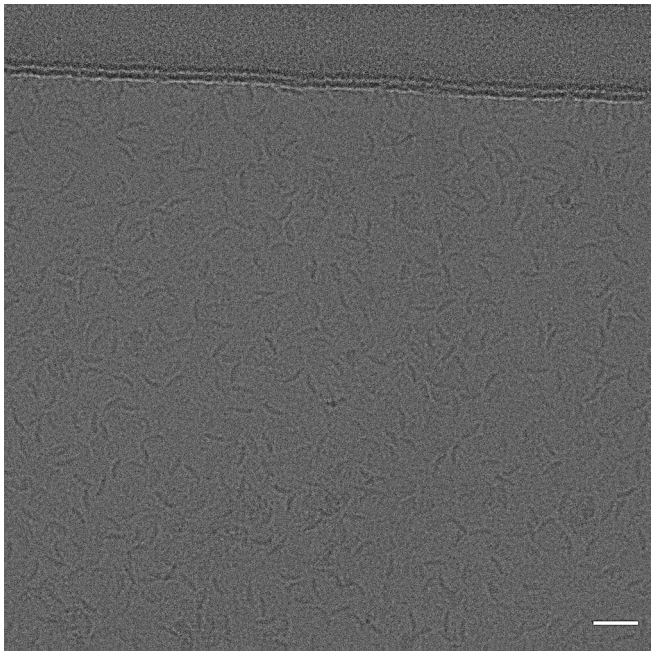


Figure 5.22: Micrograph of FMDV A22cc post 3 hr treatment with 10 mM DTT. The finger-nail clipping-like structures are all pentamer side views. Scale bar = 300 Å.

5.2.2 10 mM DTT treated A22cc

As may be seen in Figure 5.22, treatment of A22cc VLPs with 10 mM DTT concentrations caused particles to degrade into 'finger-nail-clipping-like' particles, consistent with side-view pentameric units (Malik et al., 2017). It is unclear whether reduction of disulphide bridges at the 2- and 5-fold axes by DTT directly caused this observed degradation, or simply made particles more susceptible to degradation in weakly acidic solutions (the pH of the final 10 mM DTT solution was measured, using indicator paper, as close to neutral pH i.e. 6.5-7.5, and FMDV can degrade at pHs as high as 6.5 (van Vlijmen et al., 1998)).

A recent publication describing a 8 Å cryo-EM structure of acid-treated wild type A FMDV-A1061 pentamer reported the presence of exclusively side view orientations in vitreous ice, and a tendency of the pentamers to aggregate, collectively limiting the attainable resolution (Malik et al., 2017). The lacey continuous carbon grids used for our analyses helped to reduce particle clumping (Figure 5.22), but, consistent with these previous observations, few top views were observed. It was therefore decided to collect data at a 20 ° tilt despite tilting itself potentially compromising attainable resolution, since the microscope stage can be less stable at higher tilts, the specimen thickness is greater and there is a steeper defocus gradient across the micrographs

(Drulyte et al., 2018). As may be seen in Figures 5.23-5.25, this limited tilt enabled improved sampling of orientational space.

A total of 1323 micrographs were collected (894 tilted; 429 untilted), and particles were picked following two rounds of training using Warp (Tegunov and Cramer, 2018). Selected particles were then processed with Relion, with C5 symmetry imposed during refinement (see Figure 5.23-5.25) (Scheres, 2012). Each dataset was initially processed independently (classification, refinement, Figures 5.23, 5.24), yielding structures of respective resolutions of 4.03 and 4.09 Å (gold standard FSC = 0.143 criterion), before filtered particles (from classification) were combined (total of 34,743) and refined, yielding a final structure of 3.91 Å (gold standard FSC = 0.143 criterion), Figure 5.26 (Scheres, 2012; Zivanov et al., 2018).

As observed for our AsiaShcc VLP structure presented above, the N-terminal portions of VP1 and VP2 for the A22cc pentamer are totally disordered, and residues 1001-1025 and 2001-2031 were omitted from the model. This is unsurprising, since both regions are essential for interactions at the pentamer edge, and thus become disordered when such contacts cannot be made e.g. upon dissociation or expansion.

A pentamer of A22cc VLP (PDB: 4IV3), lacking VP4, in addition to the N-termini of VP1 and VP2, was initially rigid body fitted in Coot (Emsley and Cowtan, 2004; Emsley et al., 2010). However, clear deviation of the model from density was observed at the pentamer edges in particular.

To achieve a more robust fit, MD fitting was performed using Namdinator (Kidmose et al., 2019). A summary of the model-map improvement may be seen in Table 5.5. MD fitting gave rise to a substantially better map-to-model fit (CC increasing from 0.52 to 0.80) and the RMS distance between the input and output models was quite significant; 1.20 Å (Stuart et al., 1979), Figure 5.27.

As seen in Figures 5.28-Figure 5.29(a), the resulting model fitted the map well, albeit deteriorating towards the pentamer edges and at the surface loops where map quality was poor (Figure 5.29(b)). The harsh DTT treatment used was expected to disrupt the five-fold ring of disulphide bridges between neighbouring C3007 residue, in addition to disulphide bridges between neighbouring C2093 residues at each 2-fold axis. However, there remains, at low contour levels, appreciable density at the 5-fold axis between adjacent C3007 residues, see Figure 5.29(a,b), which suggests that this disulphide ring may not have been totally reduced.

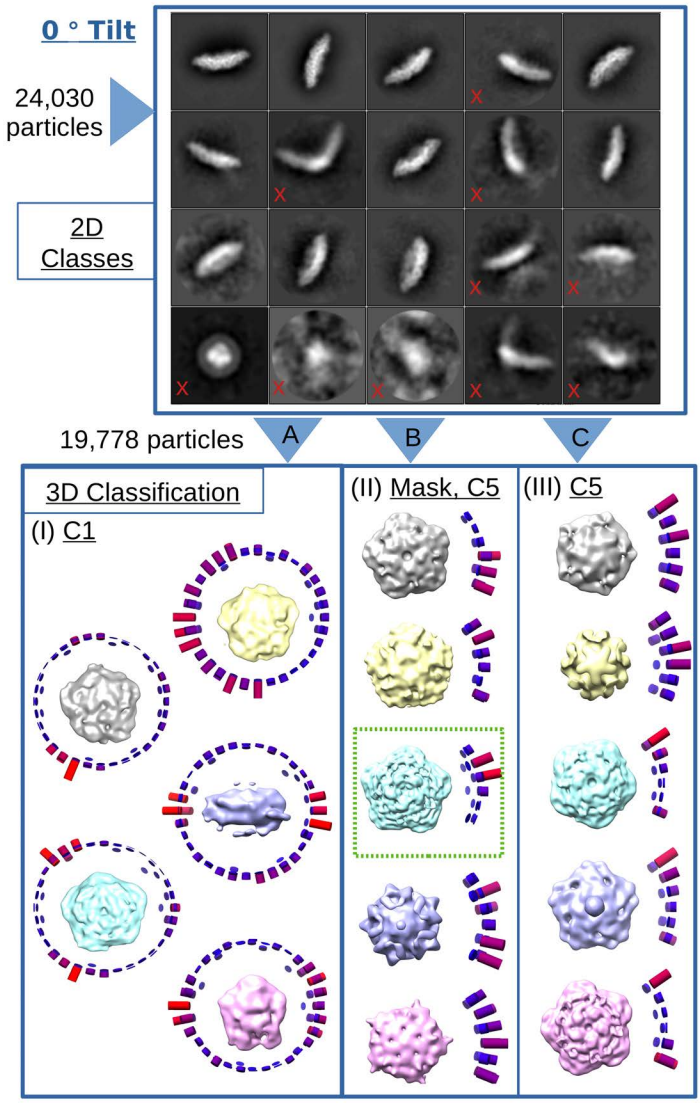


Figure 5.23: Summary of workflow used in data processing following particle picking, motion correction and CTF correction for the untilted (i.e. 0 °) dataset. First, particles from each dataset were 2D classified into 100 different classes, and, the top classes were selected on the basis of secondary structure and translational and rotational accuracies (below 2). The top 20 classes, sorted by rotational accuracy, are shown for each tilt, and discarded particles marked with a red 'x' (top blue box). Particles were then further classified using three different strategies, the resulting volumes and angular distributions from which are shown in the bottom blue box: (I) with C1 symmetry (leftmost panel), where orientational sampling seemed poor, and pentamers were less well defined as with the 20 ° tilted dataset (Figure 5.24). (II) with C5 symmetry, with a tight solvent mask around the pentamer, (III) with C5 symmetry and no solvent mask (rightmost panel). Volumes were reasonably similar for (II) and (III) and the green boxed class in run (II) was used for further refinement (see Figure 5.25 below). With hindsight, classes from all runs should have been refined and compared.

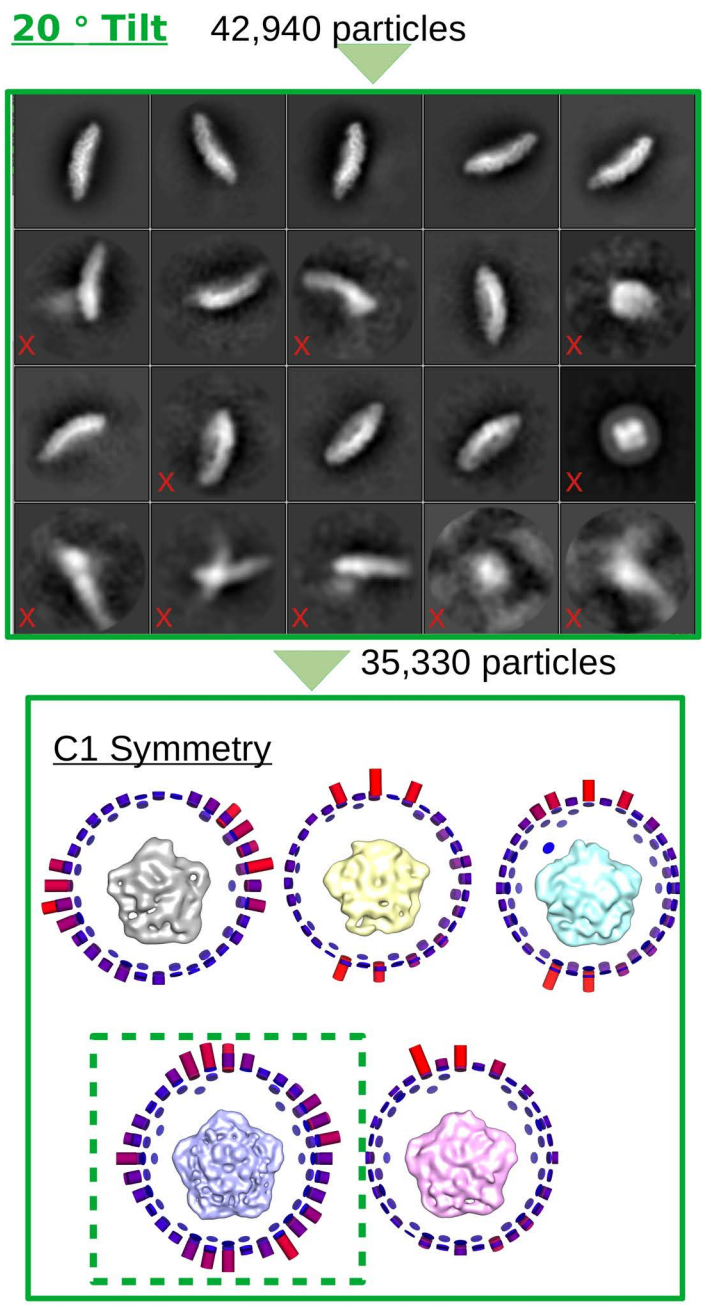


Figure 5.24: Summary of workflow used in data processing following particle picking, motion correction and CTF correction for the 20 ° dataset. First, particles from each dataset were 2D classified into 100 different classes, as in Figure 5.23. Filtered particles were then classified into 5 classes, using a 50 Å low pass filtered pentamer reference, and no symmetry (C1). One of the classes seemed to show much better angular sampling (purple and red tubes surrounding the particle), and appeared to closely resemble a symmetric pentamer that we would expect.

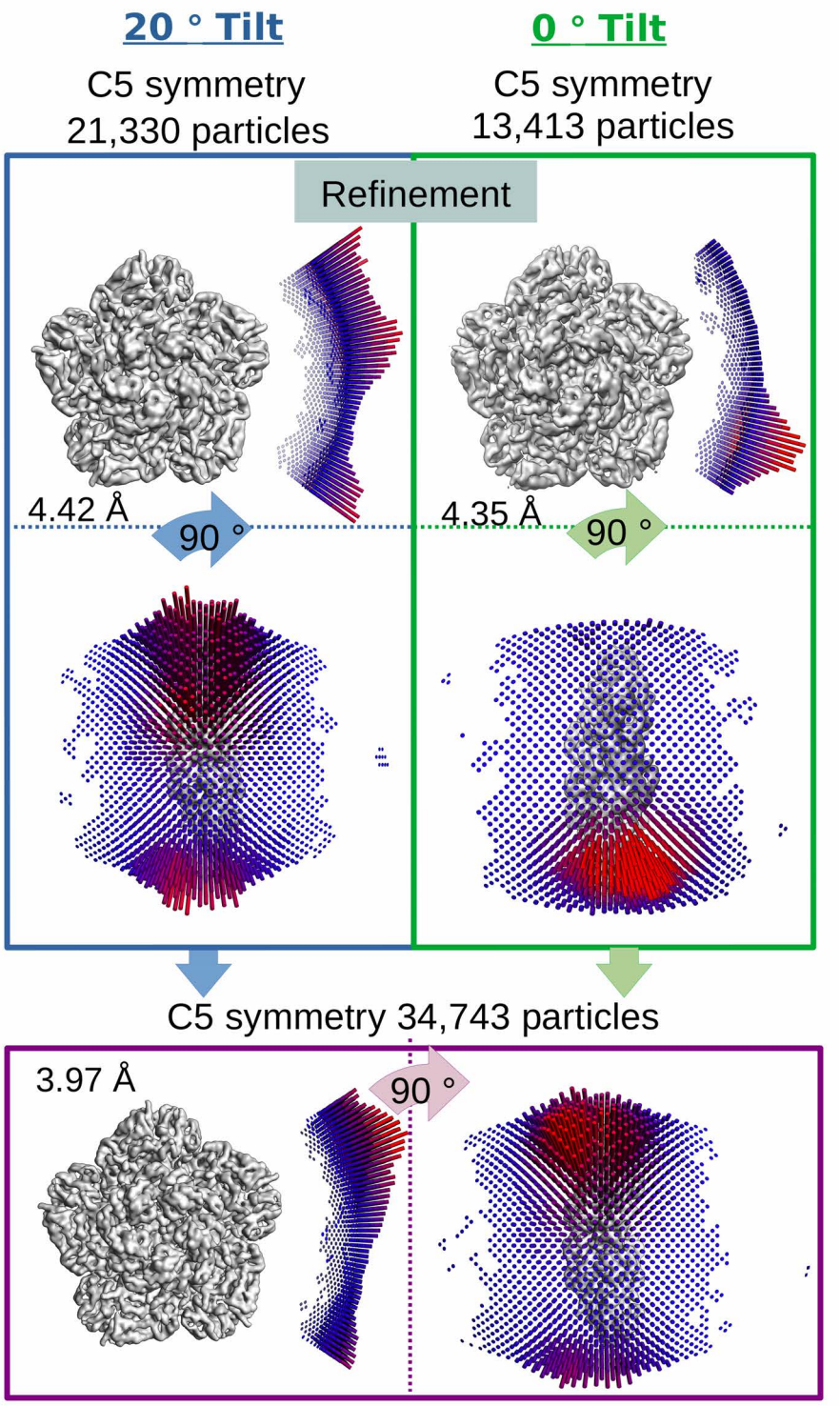


Figure 5.25: Results from refinement of selected pentamers from untilted and 20 ° tilted datasets. the resulting refined reconstruction showed major orientational bias, relative to that of the tilted structure. Given the appreciably better sampling for the tilted dataset, it is possible that the CTF fit was poorly defined with the tilt, and, perhaps in future, this should be investigated thoroughly to see if a higher resolution can be attained. The resulting reconstructions were then combined and refined together, yielding a structure of 3.97 Å resolution, which was then post-processed, yielding a final resolution of 3.91 Å.

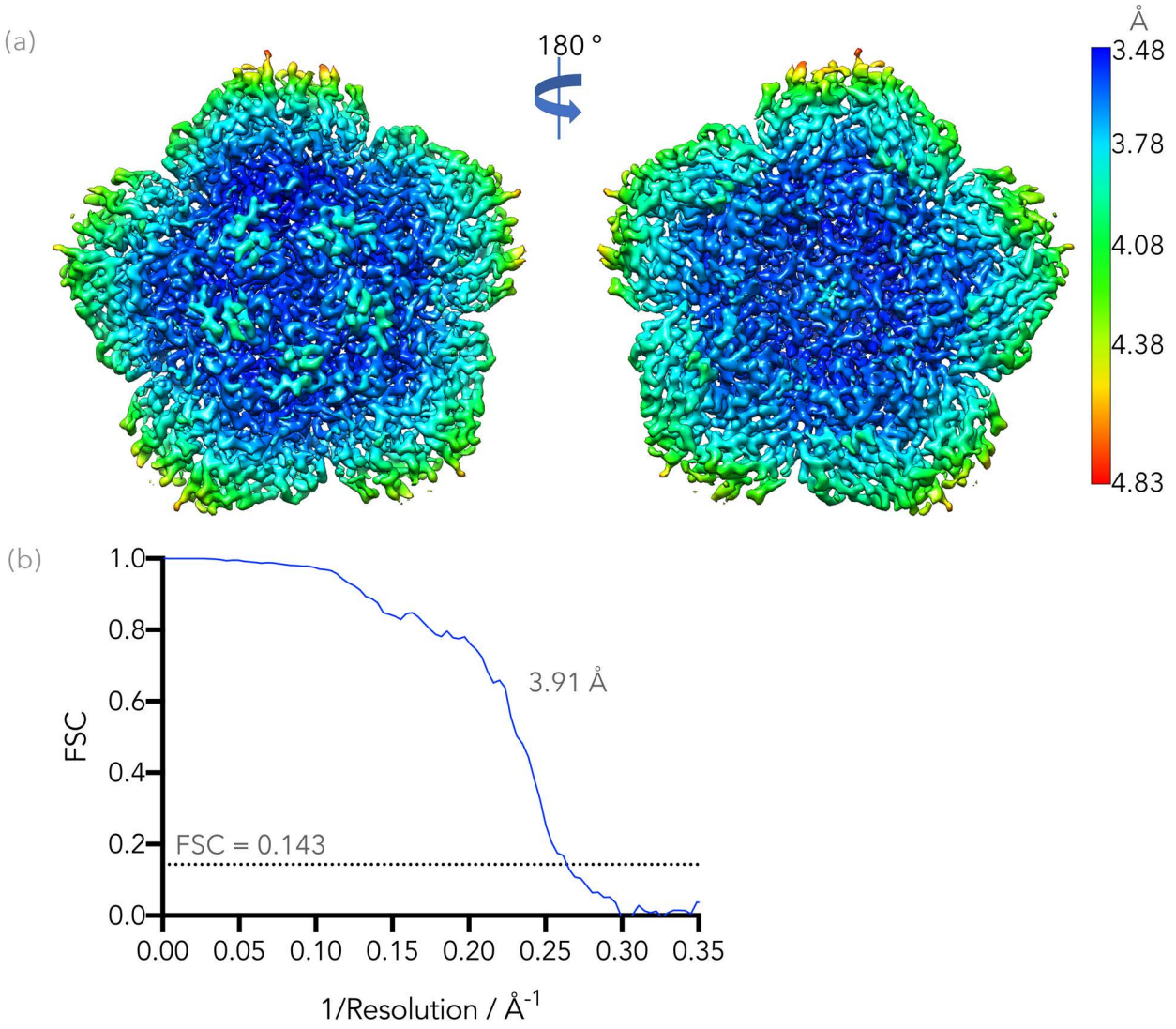


Figure 5.26: (a) Local resolution estimation determined by Relion 3.0's 'Local Resolution' tool (Zivanov et al., 2018), with colours corresponding to the key, in Å on the right. (b) FSC plot for 10 mM DTT treated A22cc, with reported resolution, evaluated using the gold standard FSC, threshold = 0.143 criterion (marked as a dashed line).

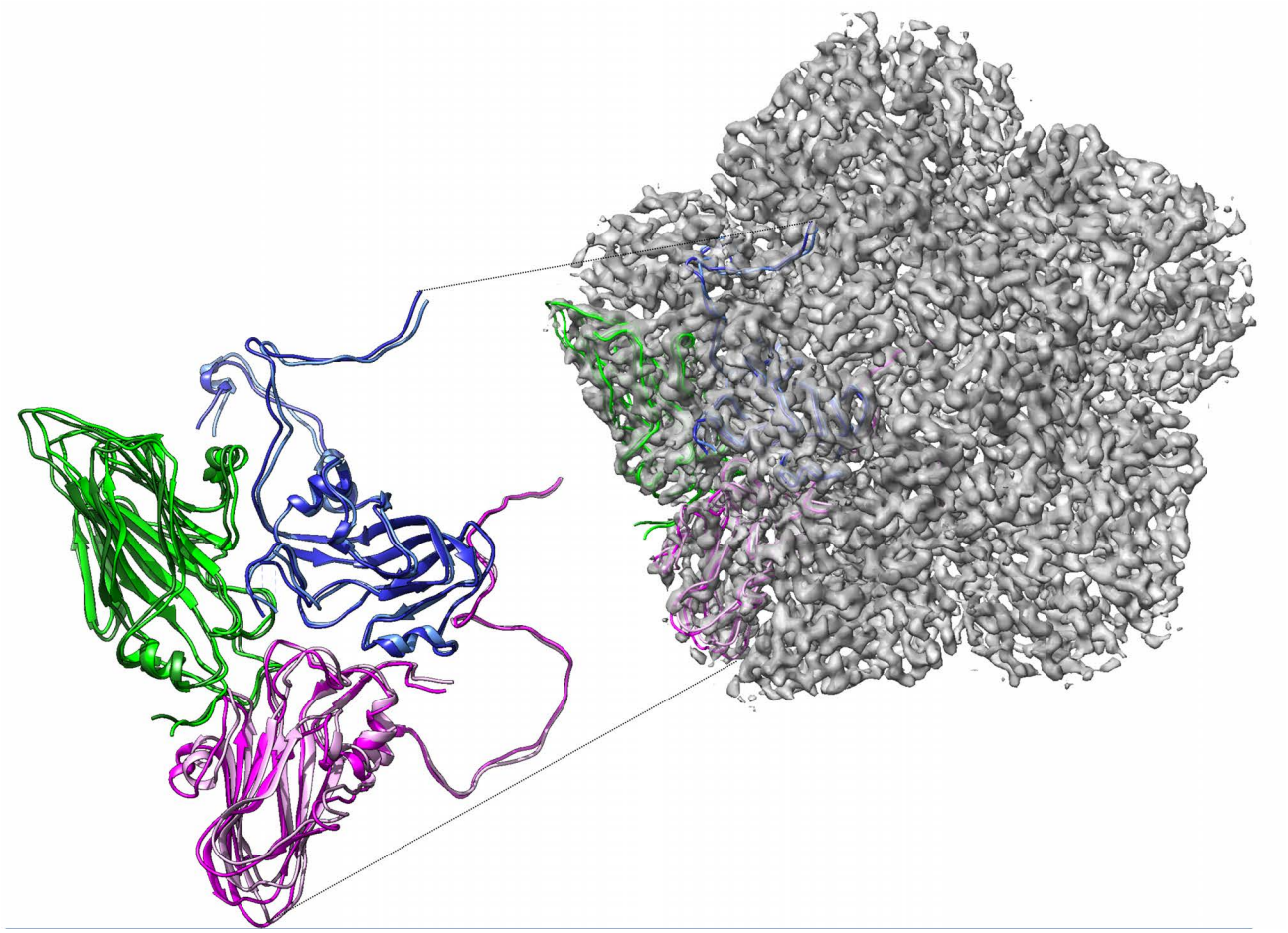


Figure 5.27: Comparison of input (pale colours) and output (bold colours) structures from refinement in Namdinator, and corresponding fit to our pentamer density.

Table 5.5: Processing statistics from Namdinator; 'Input PDB' refers to the protomer that had been rigid body fitted in Chimera and then Coot prior to upload to Namdinator. The remaining columns correspond to the output from Namdinator following MD simulation and 5 rounds of real space refinement with Phenix (Pettersen et al., 2004; Emsley and Cowtan, 2004; Emsley et al., 2010; Kidmose et al., 2019). CCC refers to the cross correlation coefficient between the input map and each of the respective PDB files. Phenix CC refers to the correlation coefficient between map and model, as determined by Phenix (Adams et al., 2010, 2011; Kidmose et al., 2019).

	Input PDB	MD Simulation	Phenix RSR
Clashscore	4.85	0.63	1.90
Protein atoms			4502
Ramachandran / %			
Favoured	93.1	92.0	94.2
Allowed	6.64	5.48	5.81
Outliers	0.17	2.16	0.00
Rotamer outliers / %	2.12	4.04	0.38
VMD CCC	0.27	0.36	0.38
Phenix CC	0.52	0.75	0.80

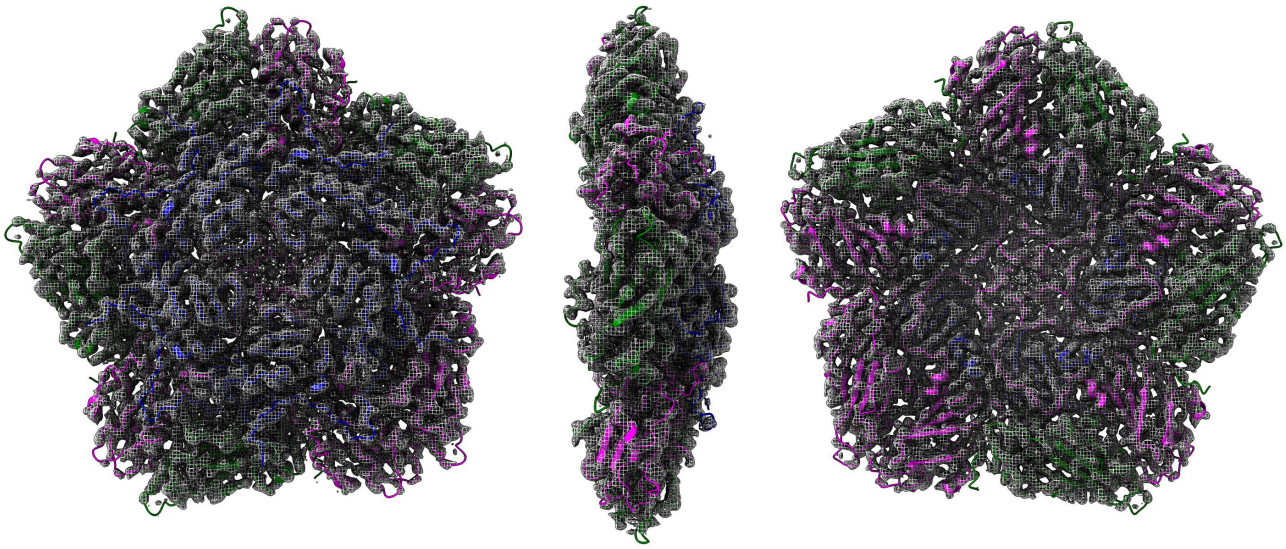


Figure 5.28: Final model fitted into EM map from front (left), side (middle) and back (right) views. Demonstration of goodness of fit of our model with the EM map. Each structural protein is coloured by convention, density is shown as a grey mesh.

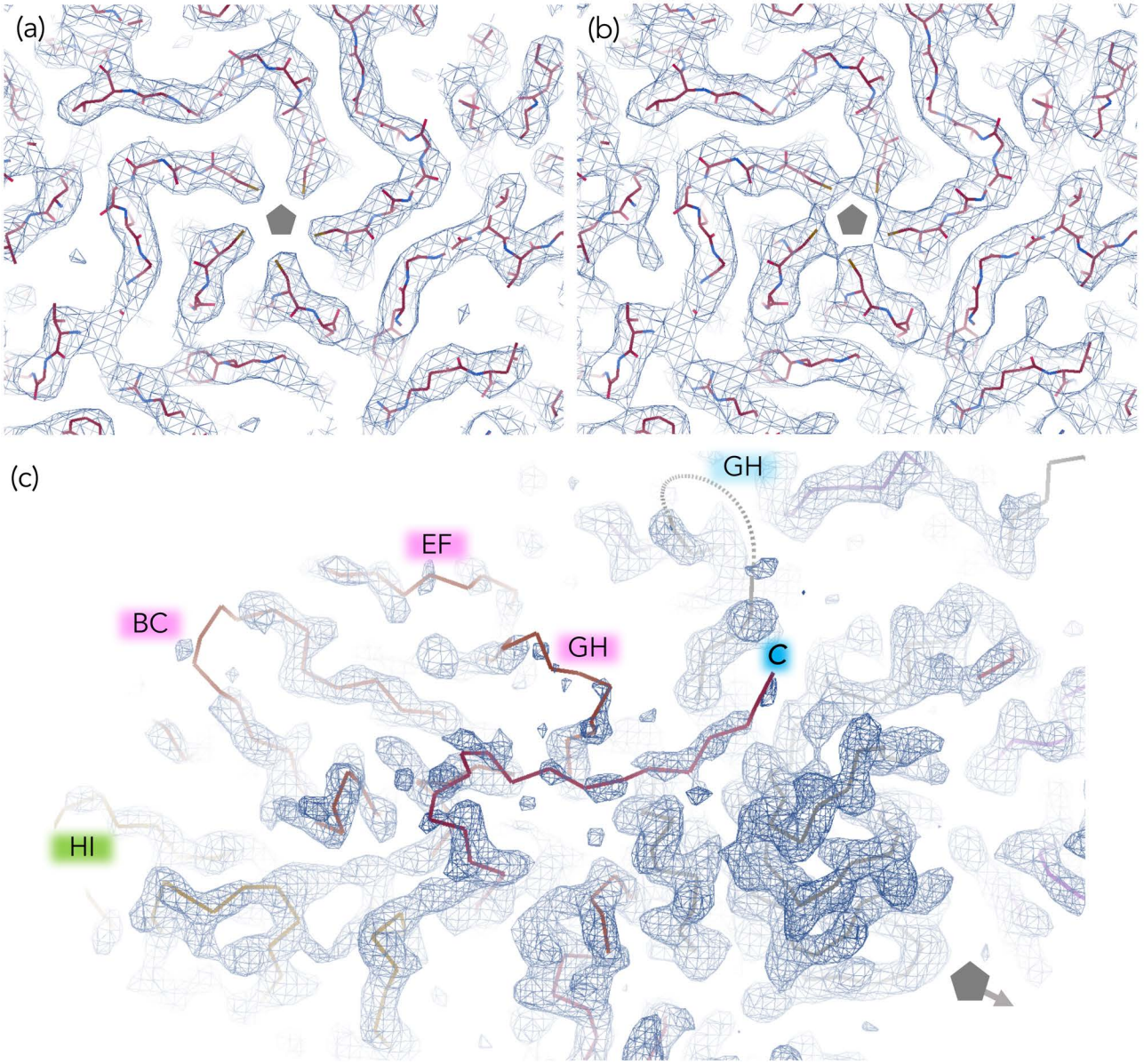


Figure 5.29: Zoomed in view at the pentamer centre, demonstrating the model fit in map at higher (a) and lower (b) contour levels. (c) An overview of fit for some surface loops. The model is shown as ribbons for clarity and major loops are marked, with coloured boxes reflecting the structural protein to which they belong (blue VP1, green VP2, pink VP3. The C-terminus of VP1 is marked in bold italics.

Comparison with existing FMDV structures

To compare our pentamer model (A22cc-P) with that of published FMDV native and VLP models, RMS distances were computed with SHP (Stuart et al., 1979) (Table 5.6, corresponding equivalences in Table 5.7).

In addition, for a better visual comparison, per-residue C_α RMS distances between a protomer from our pentamer and that of FMDV A22cc (PDB: 4IV3, Porta et al. (2013a)), FMDV A22 (PDB: 4GH4, Curry et al. (1996)), and FMDV A1061 (PDB: 1ZBE, Fry et al. (2005)) were determined in Pymol (DeLano et al., 2002), and mapped onto the A22cc protomer structure (Figures 5.30-5.32).

Generally, RMS distances between our A22cc-P model and the A22cc VLP structure are closer than those between A22cc-P and A22 native structures for all VPs (Tables 5.6, 5.7). There also seems to be a closer similarity between our pentamer model and each of the full capsid A22 native structures, relative to that of the published A1061 pentamer model and native and A1061 (for a similar number of equivalences for each VP, 5.7). Both observations may be due to A22cc VLPs already displaying characteristics associated with uncoating, such as disordering of N-terminal VP1/2 loops.

VP1

Analogous to A22cc VLP and A1061 pentamer structures, the VP1 N-terminal portion of A22cc-P, up to residue 1026, was totally disordered, and thus omitted (Malik et al., 2017; Porta et al., 2013a).

For VP1, the RMS distances are relatively low between our pentamer structure and all of native VLP and virus structures (ca. 0.6 Å for 160 equivalences, see Tables 5.6-5.7). This is by contrast to the A1061 pentamer structure, which, for a similar number of equivalences, has RMS distances of ca. 1 Å when compared against native VLP and virus structures (Tables 5.6-5.7).

Significant structural differences, relative to A22cc, A22 and A1061, occur at the antigenic GH-loop (see below), and along a patch towards the C-terminus, Figures 5.30-5.32. In this divergent region towards the C-terminus (residues N1188-V1193), the strand is shifted closer to the adjacent VP2 EF-loop residues R1135-E1136, compared to A22cc VLPs. This is consistent with the A1061 pentamer and A1061 miss-assembled particles, versus native A1061 particles (Fry et al., 2005; Malik et al., 2017).

Table 5.6: RMS distances (Å) between protomers of: our A22cc-P, native FMDV A1061 (A1061-F, PDB: 1ZBE), native FMDV A22 (A22-F, PDB: 4GH4), FMDV A22cc VLP (A22cc-F, PDB: 4IV3), FMDV A1061 miss-assembled particle (A1061-M, PDB: 5OWX), (A1061-P, PDB: 5OYI), determined by SHP (Stuart et al., 1979; Curry et al., 1996; Fry et al., 2005; Porta et al., 2013a; Malik et al., 2017). Values are coloured by distance (Å), with closer distances being a warmer pale red colour; further distances pale indigo. The A22cc-P distances are shown in bold for clarity. Associated equivalences are shown in the matrix below, Table 5.7.

VP1

1 ↓ | 2 →	**A22cc-P**	A1061-F	A22-F	A22cc-F	A1061-M	A1061-P
A22cc-P		**0.62**	**0.67**	**0.62**	**1.12**	**0.95**
A1061-F	**0.62**		0.26	0.22	1.24	1.02
A22-F	**0.66**	0.23		0.18	1.24	1.06
A22cc-F	**0.61**	0.23	0.18		1.22	1.04
A1061-M	**1.09**	1.30	1.30	1.27		0.60
A1061-P	**0.90**	1.05	1.08	1.07	0.59	

VP2

1 ↓ | 2 →	**A22cc-P**	A1061-F	A22-F	A22cc-F	A1061-M	A1061-P
A22cc-P		**0.91**	**0.92**	**0.94**	**1.06**	**0.95**
A1061-F	**0.94**		0.21	0.27	1.38	1.22
A22-F	**0.98**	0.21		0.19	1.36	1.21
A22cc-F	**0.92**	0.27	0.19		1.33	1.18
A1061-M	**1.04**	1.43	1.41	1.38		0.51
A1061-P	**0.96**	1.32	1.31	1.21	0.5	

VP3

1 ↓ | 2 →	**A22cc-P**	A1061-F	A22-F	A22cc-F	A1061-M	A1061-P
A22cc-P		**1.27**	**1.27**	**1.17**	**1.29**	**1.13**
A1061-F	**1.30**		0.32	0.32	1.80	1.67
A22-F	**1.24**	0.33		0.23	1.84	1.65
A22cc-F	**1.17**	0.32	0.23		1.80	1.59
A1061-M	**1.25**	1.88	1.82	1.79		0.72
A1061-P	**1.16**	1.66	1.62	1.62	0.72	

Table 5.7: Equivalences for RMS distances (Å) between protomers of: our A22cc-P model (A22cc-P), native FMDV A1061 (A1061-F, PDB: 1ZBE), native FMDV A22 (A22-F, PDB: 4GH4), FMDV A1061 miss-assembled particle (A1061-M, PDB: 5OWX), (A1061-P, PDB: 5OYI), shown in Table 5.6, as computed by SHP (Stuart et al., 1979; Curry et al., 1996; Fry et al., 2005; Porta et al., 2013a; Malik et al., 2017). Cells are coloured by equivalence and the A22cc-P equivalences are shown in bold for clarity.

VP1

1 ↓ | 2 →	**A22cc-P**	A1061-F	A22-F	A22cc-F	A1061-M	A1061-P
A22cc-P		**161**	**162**	**162**	**160**	**160**
A1061-F	**161**		190	176	158	159
A22-F	**162**	189		178	156	159
A22cc-F	**162**	177	178		157	160
A1061-M	**159**	160	159	159		161
A1061-P	**159**	161	160	161	161	

VP2

1 ↓ | 2 →	**A22cc-P**	A1061-F	A22-F	A22cc-F	A1061-M	A1061-P
A22cc-P		**177**	**178**	**179**	**177**	**178**
A1061-F	**178**		207	207	174	175
A22-F	179	207		207	176	176
A22cc-F	179	207	207		175	175
A1061-M	177	176	176	176		182
A1061-P	177	176	176	175	182	

VP3

1 ↓ | 2 →	**A22cc-P**	A1061-F	A22-F	A22cc-F	A1061-M	A1061-P
A22cc-P		**217**	**215**	**220**	**220**	**219**
A1061-F	**218**		215	219	214	218
A22-F	**216**	215		217	213	213
A22cc-F	**220**	219	217		218	217
A1061-M	**217**	217	211	215		221
A1061-P	**219**	217	212	217	221	

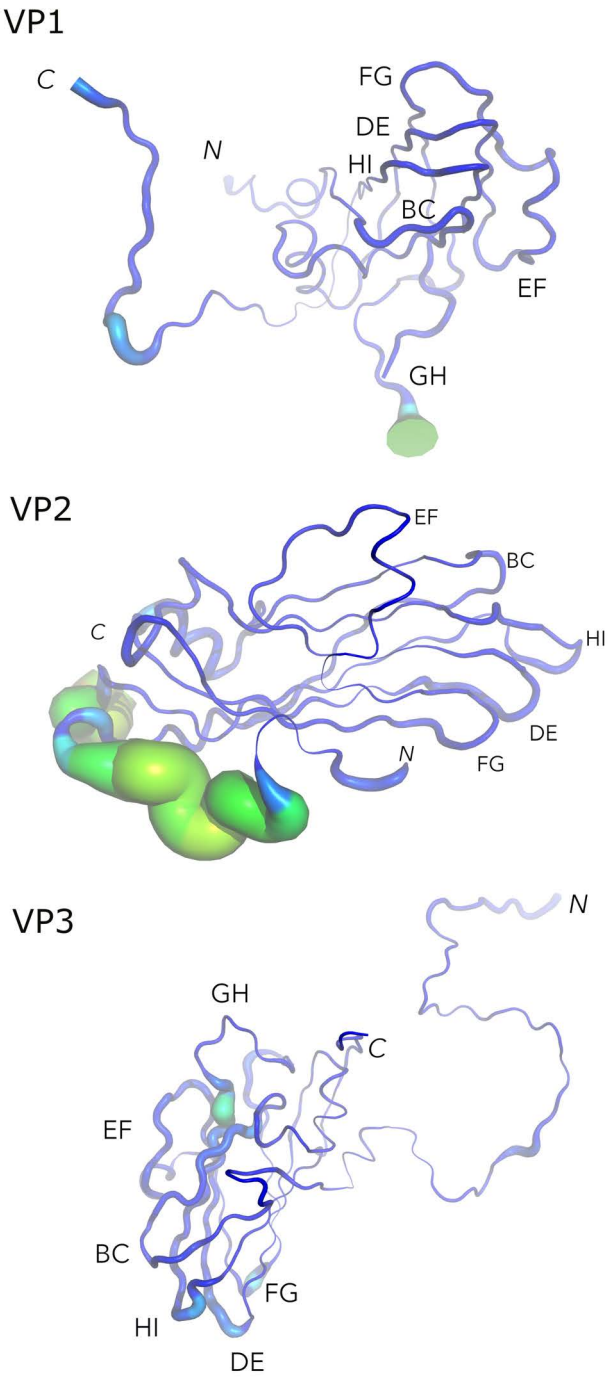


Figure 5.30: Putty representation of (from top to bottom) A22cc-P model VP1, VP2 and VP3. The colour and width of the lines reflects the RMSD between our A22cc-P and the A22cc VLP model (PDB: 4IV3) (Porta et al., 2013a), as determined by Pymol (DeLano et al., 2002). Note that the colours and thicknesses are relative within each individual model. N and C termini are labelled in italics, and loops are labelled in normal type.

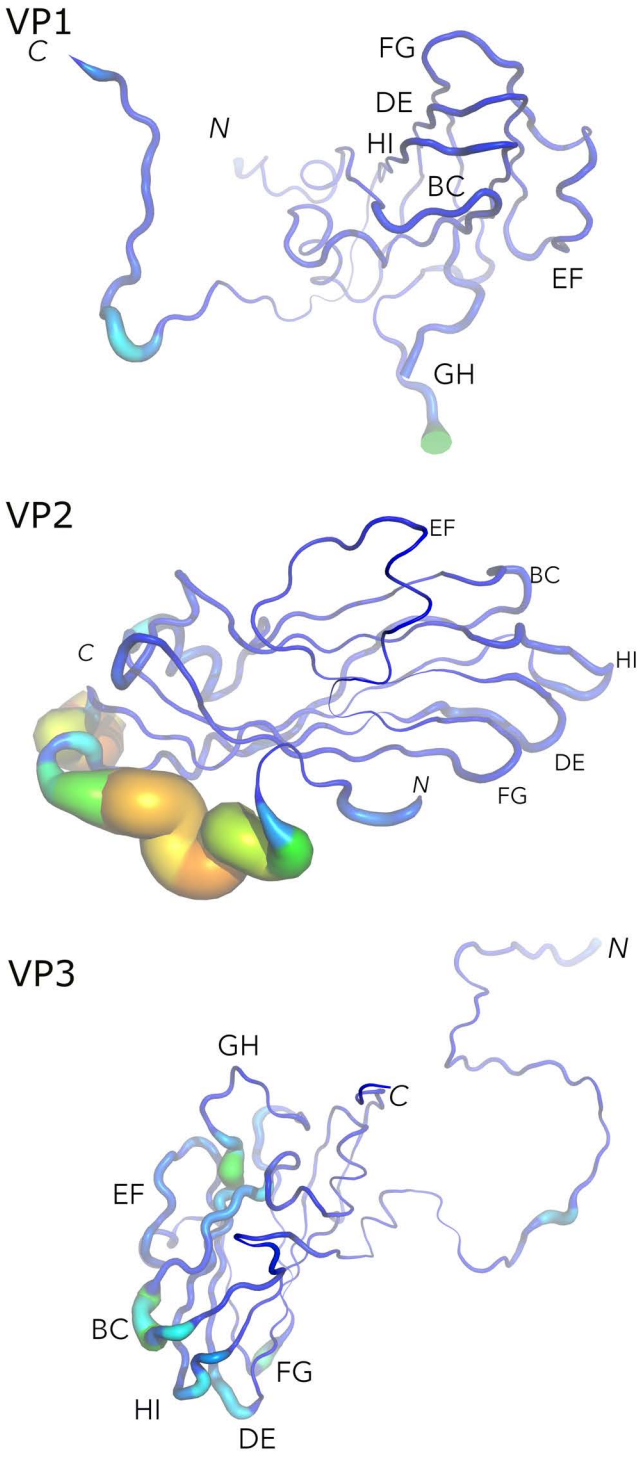


Figure 5.31: Putty representation as with Figure 5.30, except showing the RMSDs between the A22cc-P model and A1061 native virion. PDB:1ZBE (Fry et al., 2005).

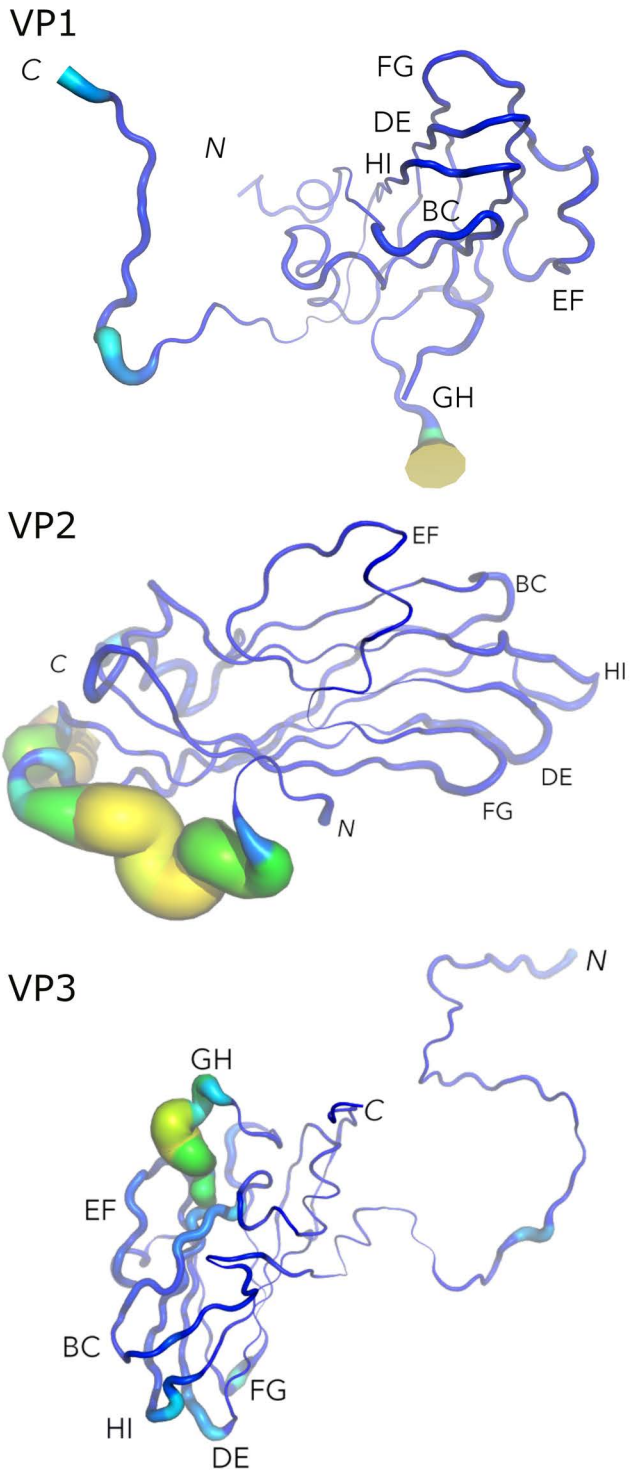


Figure 5.32: Putty representation as with Figure 5.30, except showing the RMSDs between our A22cc-P model and a model for the A22 native virion, PDB:4GH4 (Curry et al., 1996).

VP2

The most striking difference shown in Figures 5.30-5.32 occurs towards the N-terminal portion of VP2. Here, a significant movement in the loop comprising residues 2037-2049 of VP2 could be modelled for A22cc-P (see Figure 5.33(a)). Contrasting the A22cc VLP model, this loop moves away from the bulk pentamer interior (Figure 5.33(b-d)), towards the site occupied by the ordered N-terminal VP2 hairpin that sits below the 2-fold axis (Figure 5.33(c,d)). This seemingly compensates for the significant shift in VP3 of the neighbouring protomer (pink arrow, Figure 5.33(d)), a movement reported previously in the A1061 pentamer structure (Malik et al., 2017).

Another significant difference can be seen at the C-terminus of VP2, which, contrasting native and VLP FMDV particles, extends towards the outer edge of the pentamer, adjacent to VP3, and towards where, in an assembled particle, H2021 (just above the 2-fold helices) would sit, Figure 5.34. This is reminiscent of observations of an antibody-induced uncoating intermediate for HRV-B14, where these residues were also internalised to interact with VP3 of the 5-fold related protomer and VP2 of a 2-fold related protomer, an interaction not observed for other EV A-particles (Dong et al., 2017). Perhaps the disorder of the VP2 N-terminal hairpin loops allow the VP2 C-terminus of the 2-fold related protomer to relax into the structure. At very low contour levels, a slither of density can be seen in an analogous position to the A1212-L1214 portion of the modelled A22cc VLP C-terminus, perhaps indicative of a less favourable C-terminal conformation that was not distinguished during classification.

VP3

RMS distances between VP3 in the A22cc-P and the A1061 native particle and pentamer are much greater than for VP1 and 2 (greater than 1 Å, Table 5.6). Note that the number of equivalences is relatively higher for VP3 (over 200, under 180 for VP1, 2).

The closest VP3 structure to our pentamer, based on RMS distances, was that of the A1061 pentamer (overall RMSD 1.13 Å, 219 equivalences), by contrast to the A22cc full particle (1.17 Å, 220 equivalences) and A22 native structure (1.27 Å, 215 equivalences). A much greater difference can be seen between the A1061 pentamer and its related full native particle (1.62 Å, 218 equivalences).

Although a number of the VP3 surface loops differ to native structures (for instance, see Figure

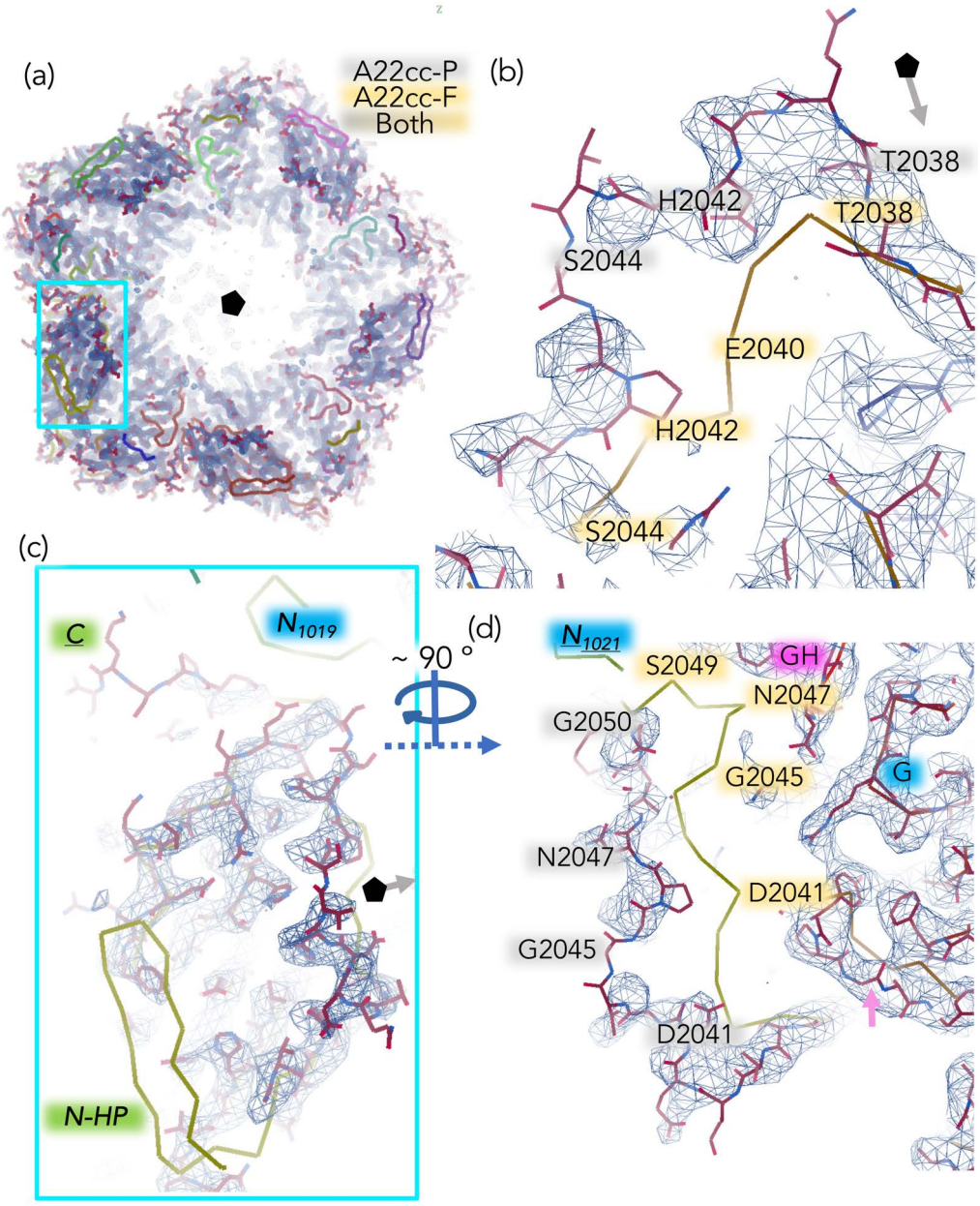


Figure 5.33: (a) Overview of A22cc-P model and map 'interior' i.e. the side that sits inside a capsid in native full-capsids. The section where we see major variation towards the N-terminus of VP2 is shown by a blue box, and a zoomed in portion of this showing our model (sticks) overlaid with that of A22cc VLP (ribbon) can be seen in (b-d). (b) A close-up of the loop region, with the approximate orientation relative to the pentamer centre shown by the arrow and pentamer symbol. (c) A directly zoomed-in view of the area boxed out in blue in (a), and (d) the same region rotated by ca. 90 °. 'N-HP' refers to the N-terminal hair pin of VP2. Label colours corresponding to specific residues correspond to the key shown to the left of (a). The N-terminal region VP1 that are not resolved in our pentamer structure are marked as a blue box. The VP1 G-strand, C-terminus of VP2 from an adjacent protomer and VP3 GH-loop are also labelled, with box colours corresponding to convention.

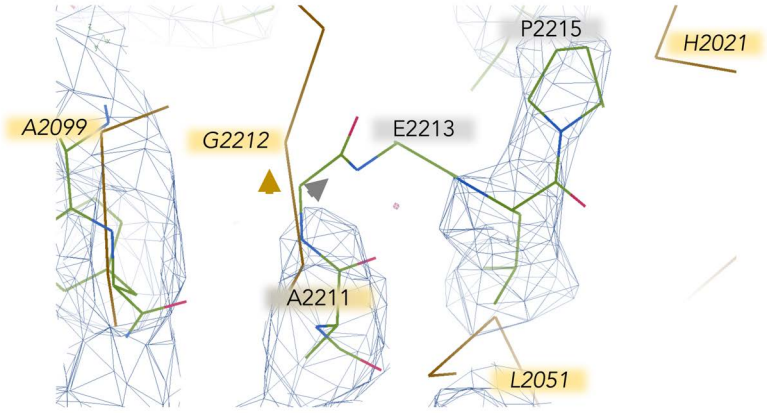


Figure 5.34: View of VP2 N-terminus for our A22cc-P model (green sticks) and that of A22cc VLP (gold ribbon, PDB: 4IV3 (Porta et al., 2013a)) sliced through the capsid (exterior on the left, capsid interior towards the right). Residue labels corresponding to A22cc VLP are shown in gold, and those for A22cc-P in grey. The direction of the VP2 C-termini for both models are shown by arrows (gold for A22cc VLP, grey for A22cc-P).

5.35 and lowermost panel, Figure 5.30-5.32), stabilising contacts are still maintained in a number of cases, such as those between opposing Arg residues in the VP3 BC loop, and the $\pi - \pi$-stacking interaction between Phe and Trp residues in the VP3 FG-loop (Figure 5.36).

Antigenic Sites

Major antigenic sites identified for FMDV A serotype based on A10 and A12 subtypes are listed in Chapter 1, Section 1.2.3, and regions are summarised in Figure 5.37(a).

Site 1 comprises the VP1 GH- and HI-loops and C-terminus. The VP1 HI-loops encircle the 5-fold axis channel; relatively close to the pentamer centre, and we see little change between our pentamer structure and that of A22cc VLP, by contrast to the C-terminus (Figure 5.30, 5.37(b)).

A large proportion of the GH-loop (residues 1132-1160, with residues 1135-1155 unmodelled for A22cc) is disordered for all X-ray capsid structures, excepting O serotype under reducing conditions, limiting structural comparison (Logan et al., 1993) (for instance, see: Lea et al. (1995, 1994); Curry et al. (1996); Fry et al. (2005); Porta et al. (2013a)). For A22cc-P, the initial, relatively ordered, stretch of residues (N1131-T1133), located at the start of the GH-loop, extend away from the bulk pentamer surface (Figure 5.37(b), Figure 5.38(a)), an orientation consistent with that of A1061, but not A22cc (Figure 5.31, versus Figure 5.30). For the A22cc, this initial part of the VLP GH-loop is similar to that of the ordered GH-loop of FMDV O1M (DTT treated, PDB: 1FOD); sitting flat, across

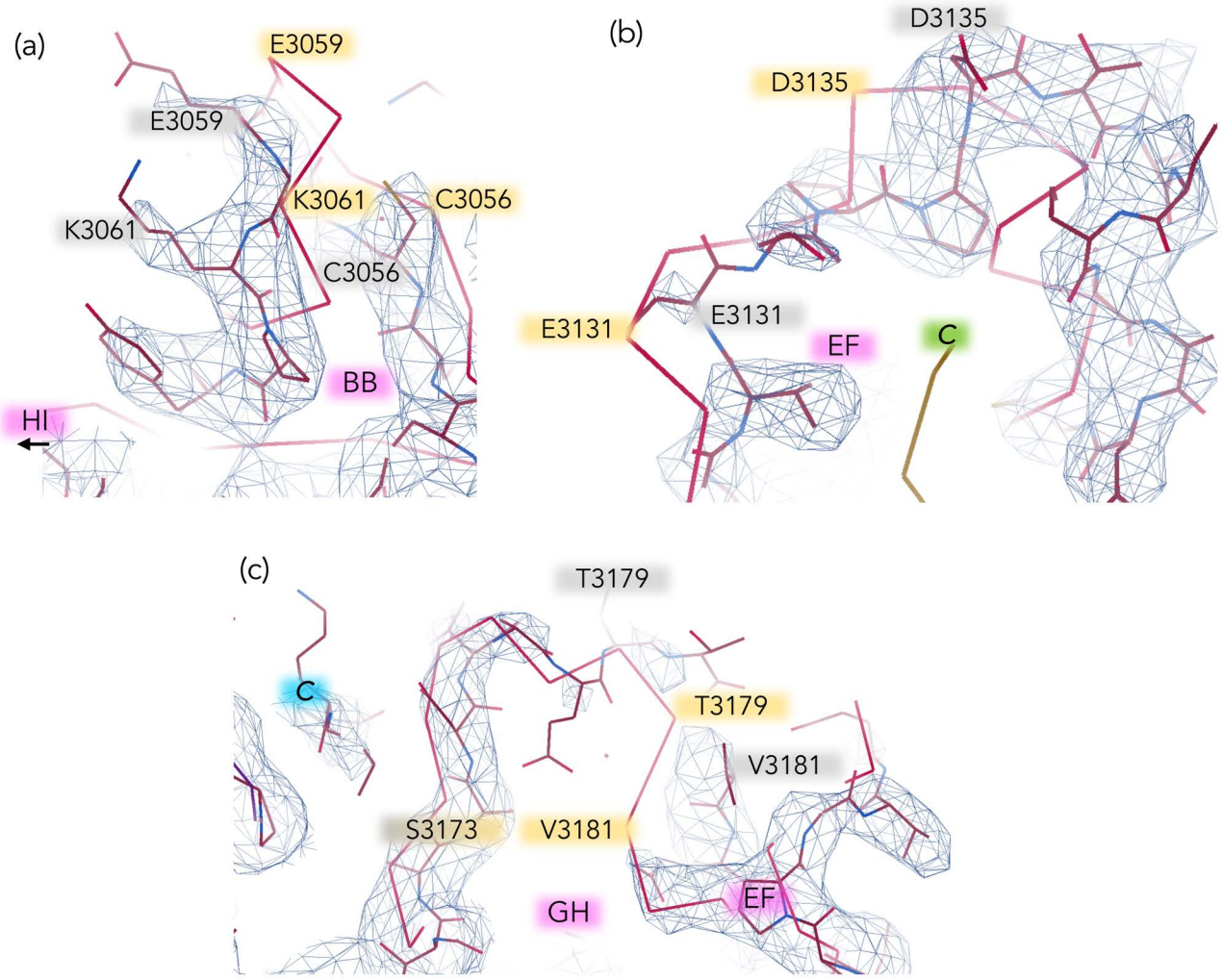


Figure 5.35: A22cc-P map and model, with the A22cc VLP model overlaid (PDB: 4IV3) at (a) VP3 BB knob, (b) VP3 EF-loop and (c) the VP3 GH-loop. A22cc VLP VP3 is shown as a red chain trace and that of VP2 as a gold trace. Residue labels corresponding to A22cc-P are shown in grey, those for A22cc VLP in gold. Our A22cc-P model is shown in stick representation. Key loops and termini are marked and are coloured by convention (blue VP1, green VP2, magenta VP3).

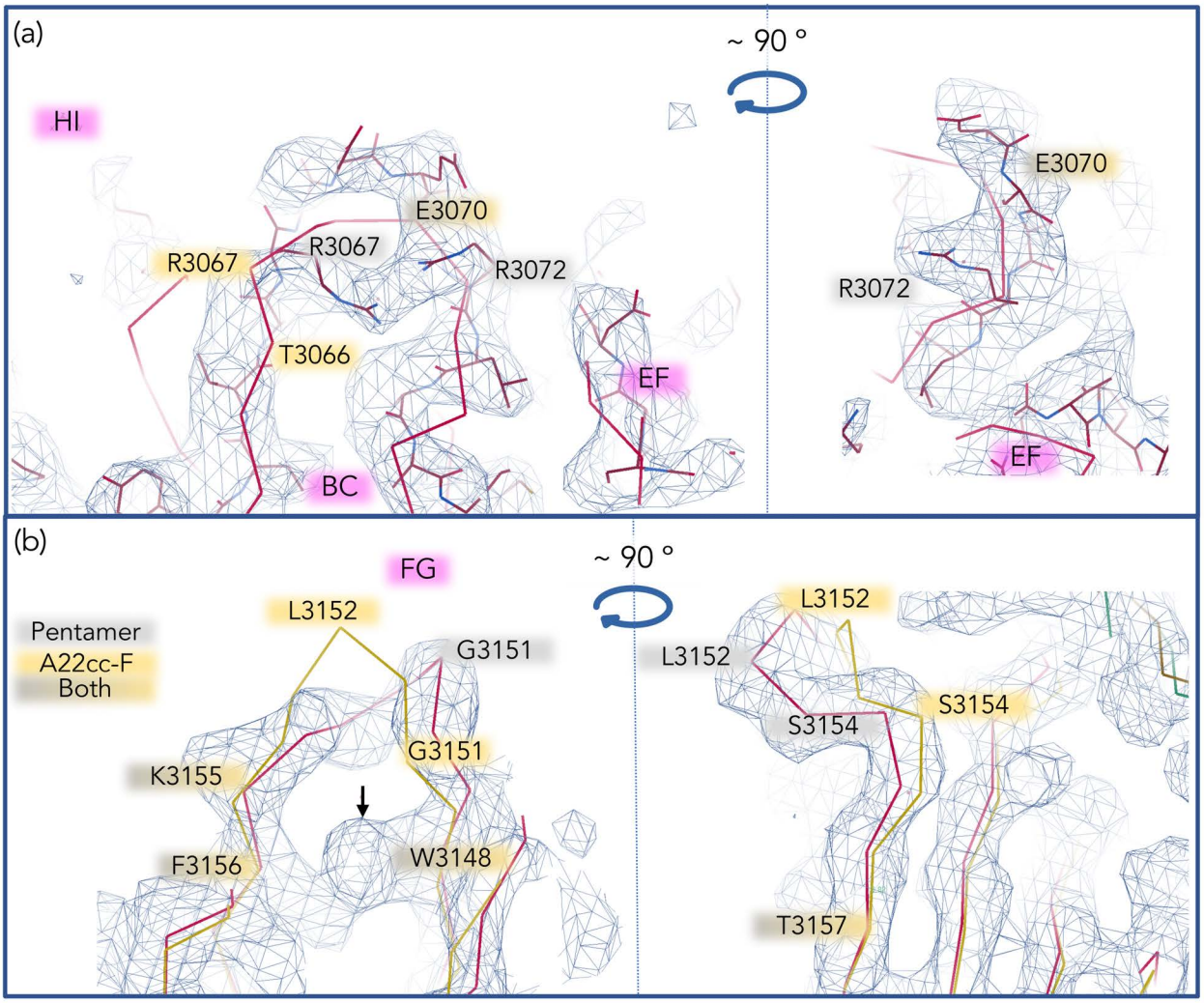


Figure 5.36: (a) Comparison of the BC-loop for A22cc-P (red sticks) and A22cc VLP (pink strands). The left image faces the BC-loop, and, for orientation purposes, nearby antigenic loops are labelled, coloured by structural protein (magenta represents VP3). (b) VP3 FG-loop from front (left) and side (right). The density for the bridge between F3156 and W3148 is indicated by an arrow. A22cc VLP is shown as gold line for clarity; our A22cc-P model is a red ribbon.

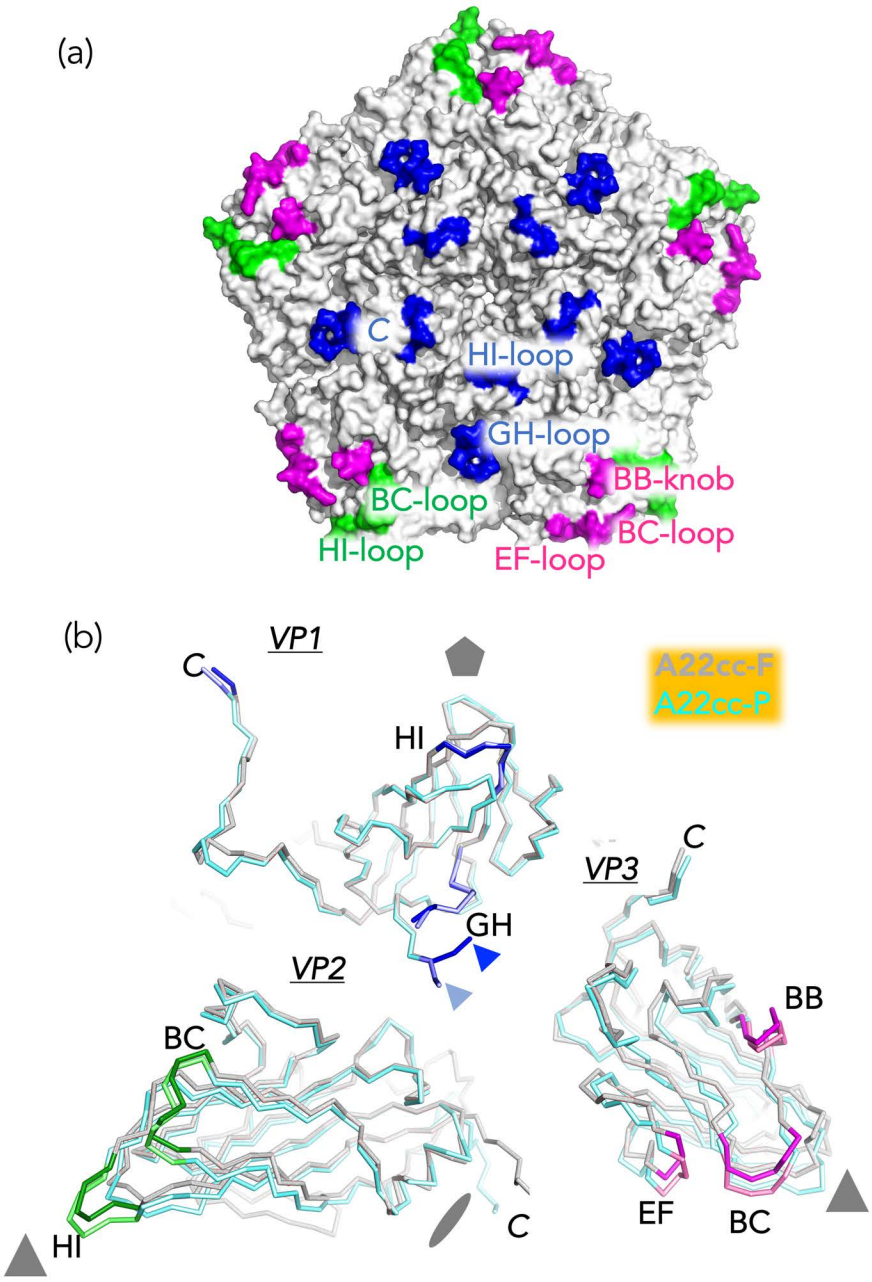


Figure 5.37: (a) Pentamer of A22cc (VLP) (PDB: 4IV3) in surface representation (white), with loops that have been identified as antigenic sites for FMDV A10 and A12 subtypes, coloured by VP; site 1, blue; site 2, green; site 3, magenta (Thomas et al., 1988; Baxt et al., 1989; Grazioli et al., 2013; Porta et al., 2013a). (b) Ribbon representation of protomers from A22cc (PDB: 4IV3, grey) and A22cc-P (cyan), separated by VP. The loops containing antigenic sites are shown in bold colours for A22cc VLP, and lighter shades for the A22cc-P (site 1, blues; site 2, greens; site 3, pinks) (Porta et al., 2013a). The two blue arrow heads next to the VP1 models point to the early portion of the VP1 GH-loop for A22cc-pentamer (light blue) and A22cc (bold blue).

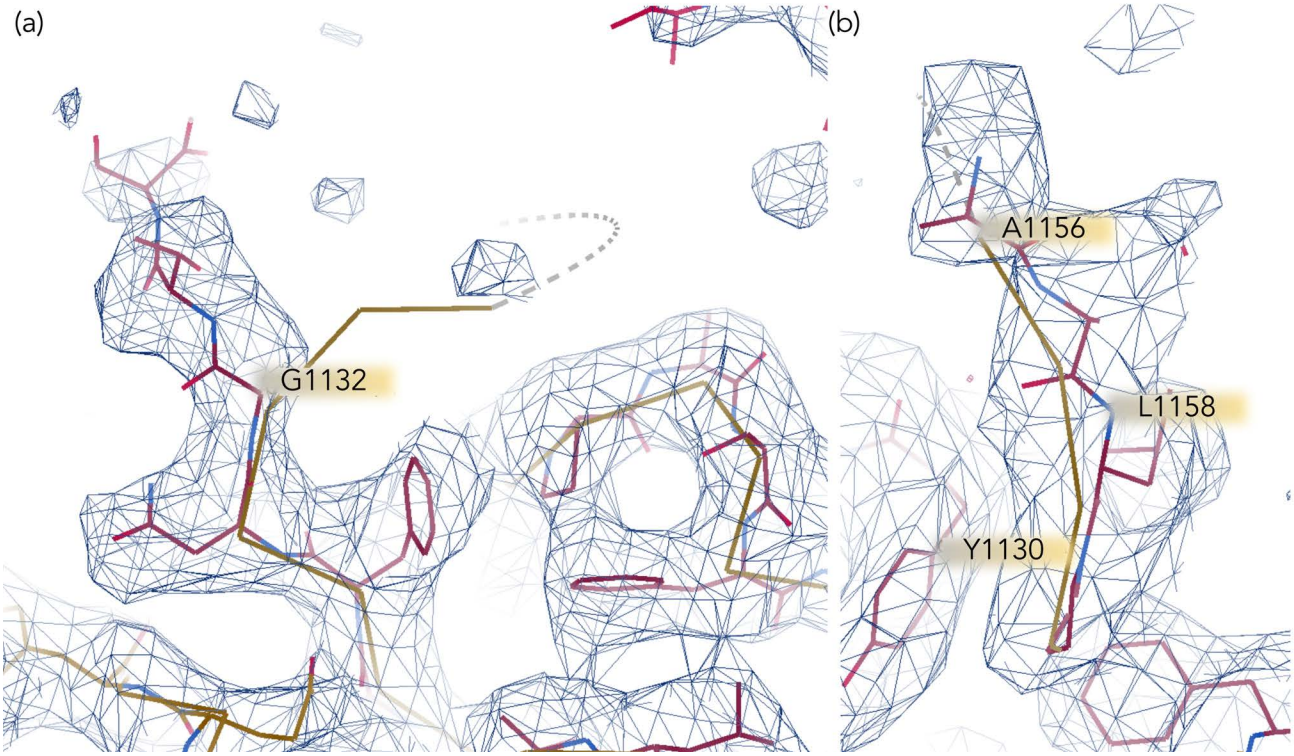


Figure 5.38: View of the VP1 GH-loop for our pentamer model (sticks) with that of the A22cc-VLP model (PDB: 4IV3, gold ribbon) (Porta et al., 2013a). (a) Side view of the initial part of the VP1 GH-loop. (b) View of the latter portion of the VP1 GH-loop.

the capsid surface from residue G1132 (Logan et al., 1993), Figure 5.38(a). Pentamer residues immediately following the GH-loop (beyond A1156) are, by contrast, similar to A22cc, Figures 5.37, 5.38(b). Unfortunately, density was too poor to confidently model an alternative loop conformation (Figures 5.29(c), 5.38), but our observations certainly account for the altered antigenicity observed in disassembled vaccines (Cartwright et al., 1980).

For site 2, the VP2 BC- and HI-loops of A22cc-pentamer are slightly flatter than that of the A22cc VLP; closer to the bulk surface. Similarly, the VP3 BB-knob, BC-loop and EF-loop also appear relatively flattened (Figure 5.37(b)). For the VP3 BC-loop, a H-bond interaction between N3067 and N3072 that occurs in A22cc VLP is maintained, despite a substantial twisting of the loop (Figure 5.36(a), above). These loops are located at the pentamer edges, where, unlike the full VLP capsid, there are fewer protein-protein contacts to constrain loop positions (Figure 5.37).

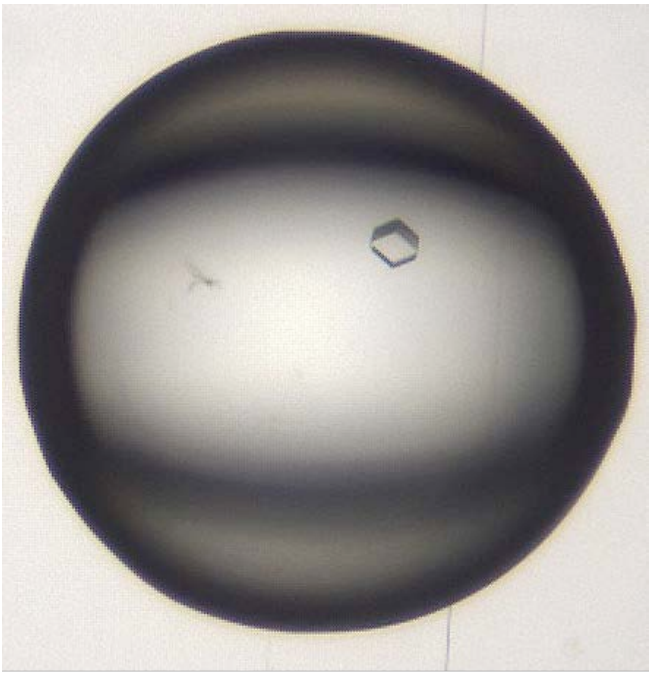


Figure 5.39: Example crystal of A22cc, crystallised at pH 8, 3.45 M ammonium acetate. The drop measures approximately 1 mm across.

5.2.3 Investigating the Metal Ion composition of FMDV VLPs

As described in Chapter 1 (Section 1.1.5), there is a body of evidence for the importance of metal cations on FMDV capsid stability. Ca^{2+} ions have been shown to stabilise ERAV empty capsids, indicating that their stabilising role may not be genome-specific, and thus may prove useful for the stabilisation of FMDV vaccine VLPs (Tuthill et al., 2009). FMDV A22cc was chosen as a vehicle to investigate (i) Calcium ions were naturally present in vaccinia-produced FMDV A22cc's capsid, or (ii) Calcium ions can bind to specific sites in the capsid following soaking.

A22cc VLP particles have been crystallised previously (Porta et al., 2013a), and can be reliably reproduced under the same crystallisation conditions (see Chapter 2, Section 1.4.1). Indeed, for our crystallisation trays of with A22cc vaccinia VLP sample, prepared as described previously, typically one crystal per well, of 50 μm diameter, grew within a month, and an exemplar crystal may be seen in Figure 5.39.

Subsequently, some crystals were soaked with 5 mM $CaCl_2$ and these, together with some unsoaked crystals, were gradient glycerol cryoprotected (Chapter 2, Section 2.4.2) for ca. 20 min prior to cryo-cooling.

Table 5.8: Theoretical f′ and f″ values for Ca and S at the probe wavelengths used for the FMDV A22cc data collection. Note that these values fail to take into account local environmental effects, and are merely a guide. Taken from Brennan and Cowan (1992).

Wavelength		Ca		S	
Å	keV	f′	f″	f′	f″
3.179	3.90	-3.33	0.43	0.15	1.95
3.024	4.10	-3.94	3.98	0.24	1.79

Table 5.9: Data processing statistics for I23 anomalous data collection.

Soaked with Ca^{2+}?	No	Yes	Yes
Data Collection			
Wavelength / keV	4.1	3.9	4.1
Wavelength / Å	3.024	3.179	3.024
Space group	*I*222	*I*222	*I*222
Cell dimensions			
a, b, c / Å	324.15,337.74,359.64	323.19, 336.81, 358.56	323.55, 337.18, 358.94
α, β, γ / °	90,90,90	90,90,90	90,90,90
Resolution / Å	138.3-2.5 (2.5-2.56)	116.7-2.4 (2.4-2.46)	122-2.4 (2.4-2.46)
Rmeas	0.249 (–)	0.206 (–)	0.247 (–)
I / σI	5.23 (0.46)	5.97 (0.89)	4.58 (0.55)
CC0.5 / %	96.8 (12.6)	98.7 (46.5)	98.2 (39.2)
Completeness / %	94.8 (82.2)	88.1 (76.8)	89.6 (77.4)
Refinement			
Resolution / Å	138.3-2.5	116.7-2.4	122-2.4
No. reflections	1244976 / 3813	1303966 / 3863	1320723 / 3883
Rwork / Rfree	20.7/20.9	19.4/19.6	20.7/20.4
Protein Atoms	4792	4839	4839
Solvent Atoms	220	232	233
Ramachandran / %			
Favoured	95.6	95.8	95.7
Allowed	3.86	2.29	4.04
Outliers	0.51	0.50	0.50
R.m.s. deviations			
Bond lengths / Å	0.010	0.009	0.009
Bond angles / °	1.1	1.1	1.1

Diffraction data were collected from soaked and unsoaked crystals, at 100 K at Diamond Light Source, Didcot, UK, beamline I23, with a focussed beam (Chapter 4, Section 4.1.2). Only one of the unsoaked crystals diffracted well, and it was only possible to collect a dataset at 3.024 Å with this crystal. For one of the soaked crystals, it was possible to collect data above (3.9 keV) and below (4.1 keV) the calcium K-edge, see Table 5.8 for theoretical f″ values.

Data to ca. 2.5 Å resolution from each of these three collections were processed using XDS, and merged and scaled with XSCALE (Kabsch, 2010*b*), before initial rigid body refinement followed by positional and B-factor refinement with Phenix (Adams et al., 2010, 2011) using the published A22cc structure (PDB: 4IV3, Porta et al. (2013a)) as an initial model, with strict 15-fold NCS constraints.

Resulting models were then manually inspected in Coot (Emsley and Cowtan, 2004; Emsley et al., 2010) before further positional and B-factor refinement with Phenix and NCS constraints as before (Adams et al., 2010, 2011). Data collection and refinement statistics may be found in Table 5.9.

For data collected above the calcium edge, higher background from the Ca^{2+} fluorescence would be anticipated, in addition to increased absorption in the solvent around and the channels inside the crystal, which may account for the slightly worse data quality (see Table 5.9).

Comparison with Room-Temperature FMDV A22cc

In contrast to the structures presented in Section 5.2.1, our crystallographic models showed no expansion at the 2-fold axes (Figure 5.40), and are consistent with the PDB deposited room temperature structure of A22cc (4IV3) that was used as an initial model (Porta et al., 2013a). This could simply be due to a greater propensity of non-expanded particles to crystallise under the conditions used.

Unit cell shrinkage associated with cryo-cooling has been known to cause shearing of disulphide bridges in sample preparations (Ginn et al., 2015), but we see no such effects in our structure, as evidenced by the disulphide bridge that may be seen in Figure 5.40(a).

Could Calcium be bound?

To establish whether calcium could be bound, anomalous differences and refined models collected above and below the calcium edge were each inputted into ANODE (Chapter 2 Section 2.4.5. Note that Phenix automatically computes anomalous difference maps (ANOM, PHANOM columns in mtz, too) (Adams et al., 2010, 2011; Thorn and Sheldrick, 2011). A summary of the top 12 residue peak heights from ANODE may be found in Table 5.10 (Thorn and Sheldrick, 2011).

The resulting substructures were then 15-fold NCS averaged and inspected in Coot (Emsley and Cowtan, 2004; Emsley et al., 2010).

Significant anomalous signal was observed seemingly exclusively for sulphur atoms from the large number of Met and Cys residues present in the model (Table 5.10). Clear signal could be seen at anticipated disulphide bridges 5.40(c); bridging adjacent C3007 residues at the five-fold axis and adjacent C2093s at the 2-fold axis (see Figures 5.40(c; d-f), 5.41). The presence of disulphide bridges at the five-fold axis also accounts for the positive peaks seen in the Fo-Fc maps for each dataset, Figure 5.40(d).

Based on theoretical f"s listed in Table 5.8, for a single calcium ion, at 4.1 keV wavelength, we anticipate 4 anomalous electrons, and 1.8 anomalous electrons for a single sulphur atom.Therefore, for calcium ion at full occupancy, we would expect a calcium peak height to be ca. 2 times greater than that of sulphur (i.e. 4 electrons / 1.8 electrons). The highest, unaveraged peak height for the calcium soaked dataset collected at 4.1 keV is 11 σ (residue C2120 Table 5.10), and so, with 1.8 anomalous electrons, we see a peak height of 11 σ for a given B-factor. Therefore, in theory, a fully occupied calcium site would have a peak height of (4 electrons / 1.8 electrons) * 11 = 24 σ. Thus, to achieve a peak height of, say, 4.5 σ (the sigma level just above the noise in an anomalous map), the minimum occupancy for calcium, with 4 anomalous electrons per atom, would be ca. 19 % (4.5 σ/(24 σ)). This is assuming that any bound calcium ions have an identical B-factor to residue C2120, and that anomalous electrons are similarly distributed for both calcium and sulphur atoms, coming from each respective atom's innermost shell. With 15-fold NCS averaging, we would anticipate the minimal detectable occupancy, at 4.5 σ, to be as little as 5 % (4.5/(2 * 11 * $\sqrt{}$15)). Thus, our results suggest that no calcium is bound at an occupancy of above ca. 5 %.

For the Ca^{2+}-soaked dataset collected at 3.9 keV only, for which we anticipate 0.4 electrons per calcium ion, a weak anomalous signal (3.7 σ, averaged) can be seen at the 5-fold axis (Figure

Table 5.10: Top 12 peak heights (without NCS averaging) and associated atoms, as determined by ANODE for each dataset (Thorn and Sheldrick, 2011). The residue numbers do not include the VP number i.e. the VP1 residue 125 is listed as Met125, not Met1125.

Wavelength / keV	Height / σ	Distance / Å	Atom:residue	VP
3.9	13.06	0.29	SD:MET125	VP2
	12.97	0.18	SD:MET182	VP2
	12.78	0.16	SD:MET86	VP3
	12.64	0.37	SG:CYS186	VP1
	12.62	0.14	SD:MET122	VP3
	12.62	0.15	SG:CYS120	VP2
	12.62	0.22	SD:MET36	VP1
	12.44	0.16	SD:MET179	VP1
	12.14	0.41	SD:MET182	VP2
	12.06	0.11	SD:MET125	VP2
	11.54	0.30	SG:CYS56	VP3
	11.52	0.03	SG:CYS143	VP3
4.1 (soaked)	11.15	0.38	SG:CYS120	VP2
	10.47	0.21	SD:MET182	VP2
	10.40	0.34	SD:MET125	VP2
	10.31	0.11	SG:CYS186	VP1
	9.90	0.14	SG:CYS143	VP3
	9.80	0.25	SD:MET101	VP2
	9.63	0.16	SD:MET122	VP3
	9.49	0.41	SD:MET86	VP3
	9.39	0.46	SD:MET36	VP1
	9.35	0.26	SD:MET179	VP1
	9.05	0.32	SD:MET54	VP1
	8.76	0.33	SG:CYS51	VP3
4.1 (unsoaked)	11.35	0.43	SD:MET125	VP2
	11.28	0.14	SD:MET86	VP3
	10.96	0.11	SDMET122	VP3
	10.89	0.17	SD:MET182	VP2
	10.87	0.32	SG:CYS186	VP1
	10.68	0.37	SD:MET101	VP2
	10.63	0.37	SD:MET179	VP1
	10.47	0.39	SG:CYS143	VP3
	10.33	0.26	SG:CYS120	VP2
	10.31	0.44	SD:MET36	VP1
	9.94	0.29	SD:MET110	VP3
	9.93	0.40	SG:CYS56	VP3

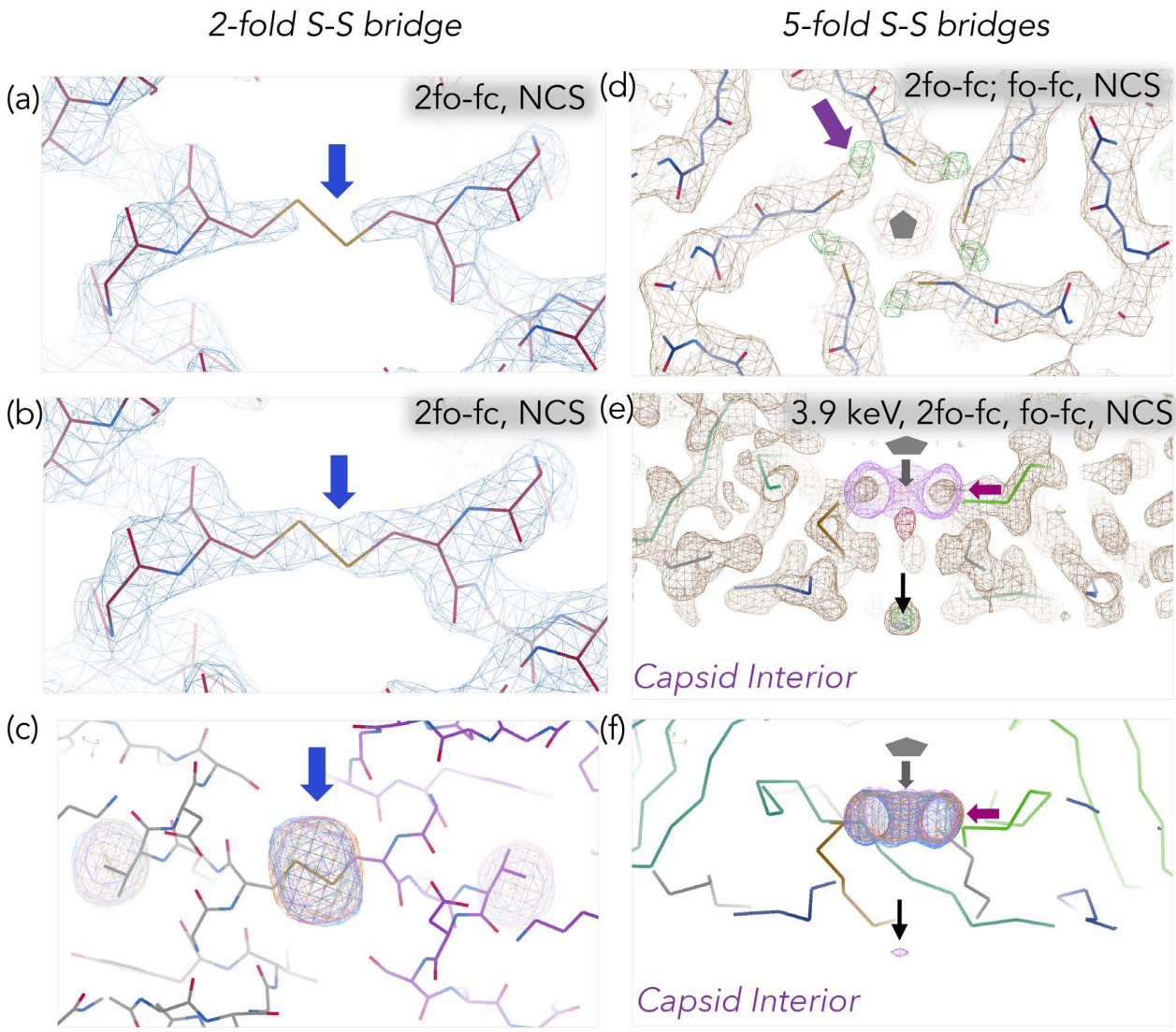


Figure 5.40: (a,b) 15-fold 2mFo-dFc map from the 3.9 keV calcium soaked dataset facing the 2-fold axis, at the junction between 2093C residues of adjacent protomers, the single mutation site that was engineered for our VLP constructs, at high (a) and lower (b) contour levels. The blue arrows indicate the position of the S-S bridge at this site. (c) Anomalous densities for all three datasets, overlaid at the 2-fold axis, indicated by the blue arrow. Densities for a M2101 residue from each protomer may be seen on the left and right of the 2-fold axis, in the background. Pink density: calcium soaked at 3.9 keV, orange density: no calcium soak at 3.9 keV, and blue density: calcium soaked at 4.1 keV. (d) 2mFo-dFc (brown) and Fo-Fc maps (green) at the 5-fold axis, using the 3.9 keV dataset as an example. Fo-Fc densities support the presence of a bridging interaction between adjacent S residues of C3007. (e) Side view of the 5-fold axis base. Protomers are coloured separately and are shown as ribbons for clarity. The 2mFo-dFc map (brown), Fo-Fc map (green and red) and anomalous density from the 3.9 keV soaked condition are overlaid. The black arrow points to a small plug of density, also observed in panel (f) that is unique to the 3.9 keV dataset. (f) As with (e), except with all anomalous densities overlaid and scaled with respect to one another based on the 2-fold density cut-off described in the main text.

5.40(f)). The significance of this is unclear.

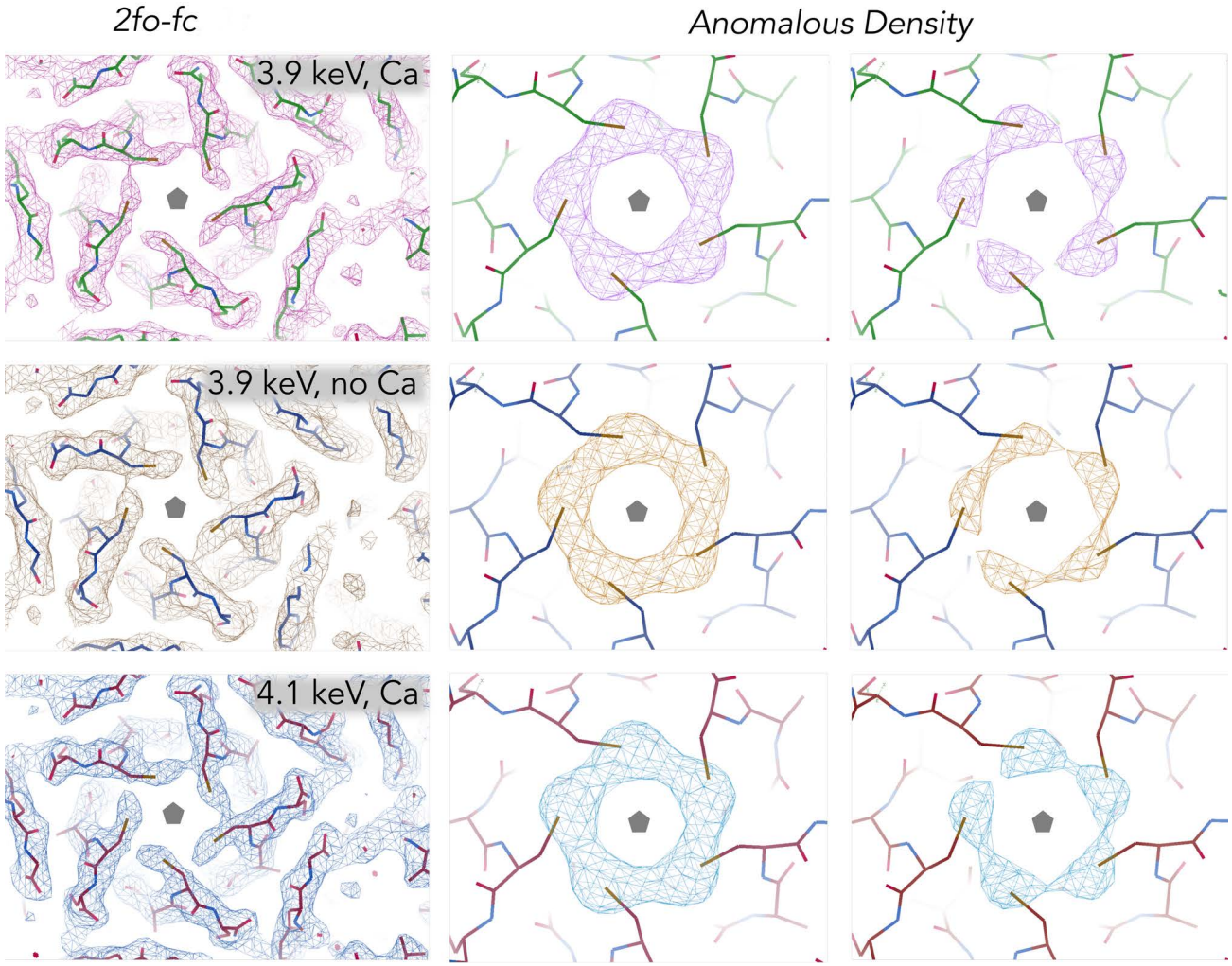


Figure 5.41: Possible disulphide bridges at the 5-fold axis. Left panel shows the 2mFo-dFc maps for each condition tested, contoured to the same level. Middle panel and rightmost panels show the anomalous density for each condition mapped onto the 5-fold axis at two contour levels. These should be approximately equivalent, and were derived using the minimum contour level at which the 2-fold disulphide bridge could no longer be discerned for each condition. Densities are 15-fold NCS averaged.

5.2.4 Thermofluor analysis of VLPs

FMDV VLP capsids containing an engineered disulphide bridge at the 2-fold axis can be produced for the majority of subtypes across all FMDV serotypes, and have been shown to have superior thermo- and acid-stability as well as retaining antigenicity after prolonged storage, relative to native virions. However, as illustrated in the preceding Sections, the reliable production of FMDV VLP particles in a native antigenic state, sufficiently stable to maintain that state for a prolonged period of time, is not completely understood.

An assay that provides a rapid means of checking VLP thermostability, as well as screening potential stabilising additives and mutations, would be highly desirable. Currently, such feedback is only achieved via running heat- or acid- treated samples on gradients, western blots and/or ELISAs (personal communication, Claudine Porta, Porta et al. (2013a)). As described in Chapter 2 Section 2.7, with a so-called PaSTRy assay, dyes sensitive to genomic and hydrophobic residue exposure may be employed to probe viral capsid stability over a range of temperatures in a 96-well format that permits the screening of a range of conditions in a single experiment, using minimal equipment and sample (Walter et al., 2012). The absence of genomic material in VLPs means that assays are restricted to the significantly less sensitive, and often more difficult to interpret, dyes that specifically recognise hydrophobic residues (Walter et al., 2012).

An existing hydrophobic dye: SYPRO Orange

To check the utility of one such hydrophobic dye, SYPRO Orange (Invitrogen) as a marker for VLP stability, a thermofluor assay was set up using a panel of vaccinia-derived VLPs. As Figure 5.42 demonstrates, consistent with previous observations (Kotecha et al., 2016), a small bump in the SYPRO Orange signal, around 50 °C, can be seen (Figure 5.42), consistent with melting curves for genomic release dyes, such as SYTO9 (Invitrogen) for native FMDV (Chapter 7, Section, 1.2.4, Kotecha et al. (2016)). However, the signal from this peak is too weak for reliable interpretation.

More sensitive hydrophobic dyes were therefore sought to try to amplify signal from potential capsid modulations. Suitable candidate dyes needed to also be, at the very least, thermostable and ideally pH stable. Searches suggested GloMelt^{TM} (Biotium) as a potential candidate.

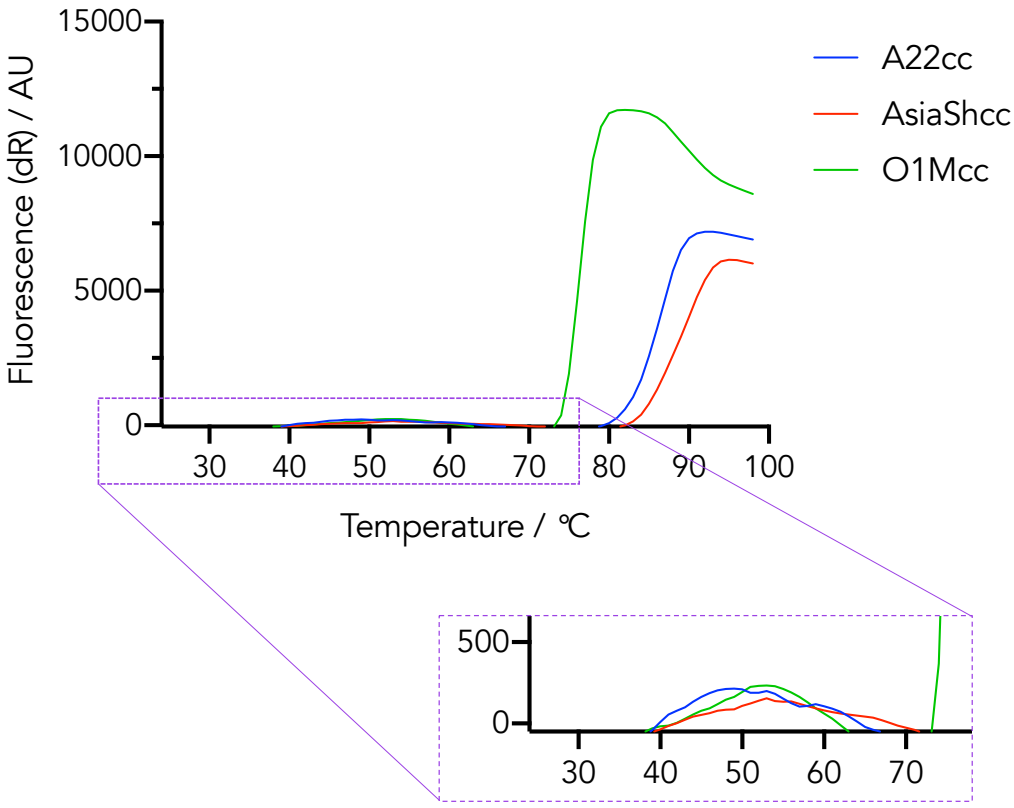


Figure 5.42: First derivative curve from thermofluor assay of vaccinia-produced A22cc, AsiaShcc and O1Mcc VLPs. The purple box is 'zoomed in' to the 'small bump' referred to in the main text.

Comparison of SYPRO Orange with GloMelt^{TM}

GloMelt^{TM} dye is purported to be more robust against additives such as DTT, imidazole, glycerol and sorbitol and under high detergent conditions (0.5-1 %), by contrast to SYPRO Orange and Red dyes (Invitrogen), and is allegedly more sensitive (biotium.com). Sensitivity is, however, highly dependent on the ratio of dye to protein, and is often system dependent, thus it was necessary to ascertain the optimal dye/protein ratio to assess GloMelt^{TM} utility.

For provisional analysis of GloMelt^{TM} sensitivity relative to SYPRO Orange (Invitrogen), 3C protease was used as a test system, since this is cheap and easy to obtain. In addition to protein:dye ratios, the concentration of both sample and dye has a massive effect on assay efficacy and, thus, a logarithmic range of concentrations of protein (3-0 μg) and dye (3-0.1 x) need to be assayed to critically assess dye performance. Establishing maximal dye concentration is particularly crucial, since too high a concentration results in collisional quenching.

Results are presented in Figure 5.43. Fluorescence was read with emission wavelengths of 510 and 585 nm and excitation wavelengths of 440 nm and 490 nm for GloMelt^{TM} and SYPRO Orange respectively. The emission and excitation wavelengths for GloMelt^{TM} were selected based on literature and corroborated with nanodrop recordings. Concentrations are described in terms of 'x', with '1 x' corresponding to the advised dye concentration for GloMelt^{TM} (sold as a 200 x concentrated solution) and SYPRO Orange (sold as 5000 x concentrated solution).

Results: New versus old Dyes

As may be seen in Figures 5.43 and 5.44, there is increase in signal from approximately 60 °C for both dyes. Consistent with associated advertisements, GloMelt^{TM} melting profiles declined the least at higher temperatures. However, SYPRO Orange still gives approximately three-fold more signal for 3 x dye concentration and comparable protein concentrations, Figure 5.44(a) versus Figure 5.43(a).

Decreasing SYPRO Orange dye concentration for a fixed protein concentration greatly reduces SYPRO Orange signal, with a nearly 10-times reduction in signal upon 3-fold decrease in dye concentration for a given protein concentration (Figure 5.44). We see a much less dramatic reduction in signal for GloMelt^{TM} dye concentration, where only a slight reduction in signal was observed between switching from 3 x and 1 x concentrations, suggesting that protein, and not dye concen-

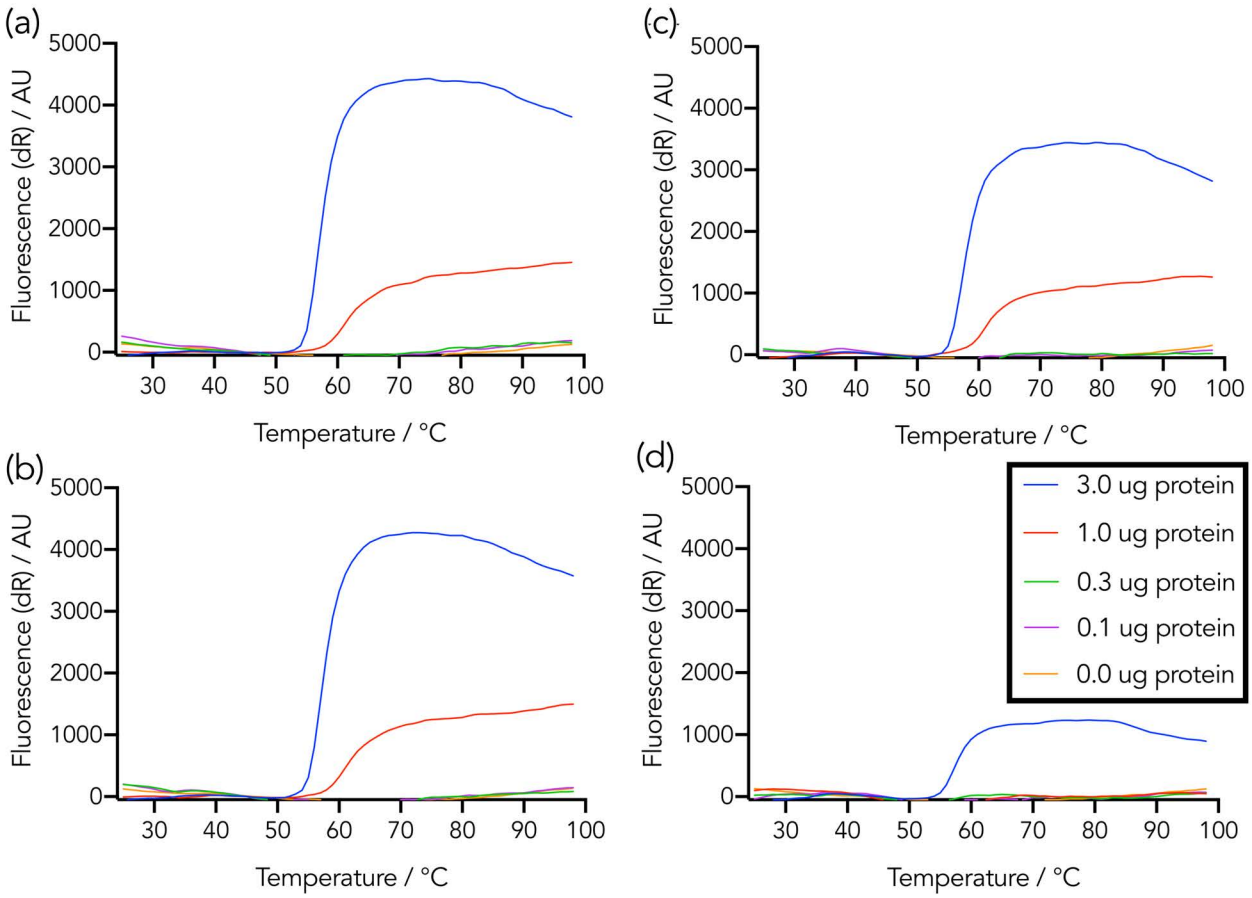


Figure 5.43: First derivative trace of GloMelt^{TM} dye. Lines are coloured according to 3C concentration shown in the key at the bottom left. Graphs (a) were collected with 3 x dye; (b) 1 x dye; (c) 0.3 x dye and (d) 0.1x dye concentrations.

tration, may have been the limiting factor in signal.

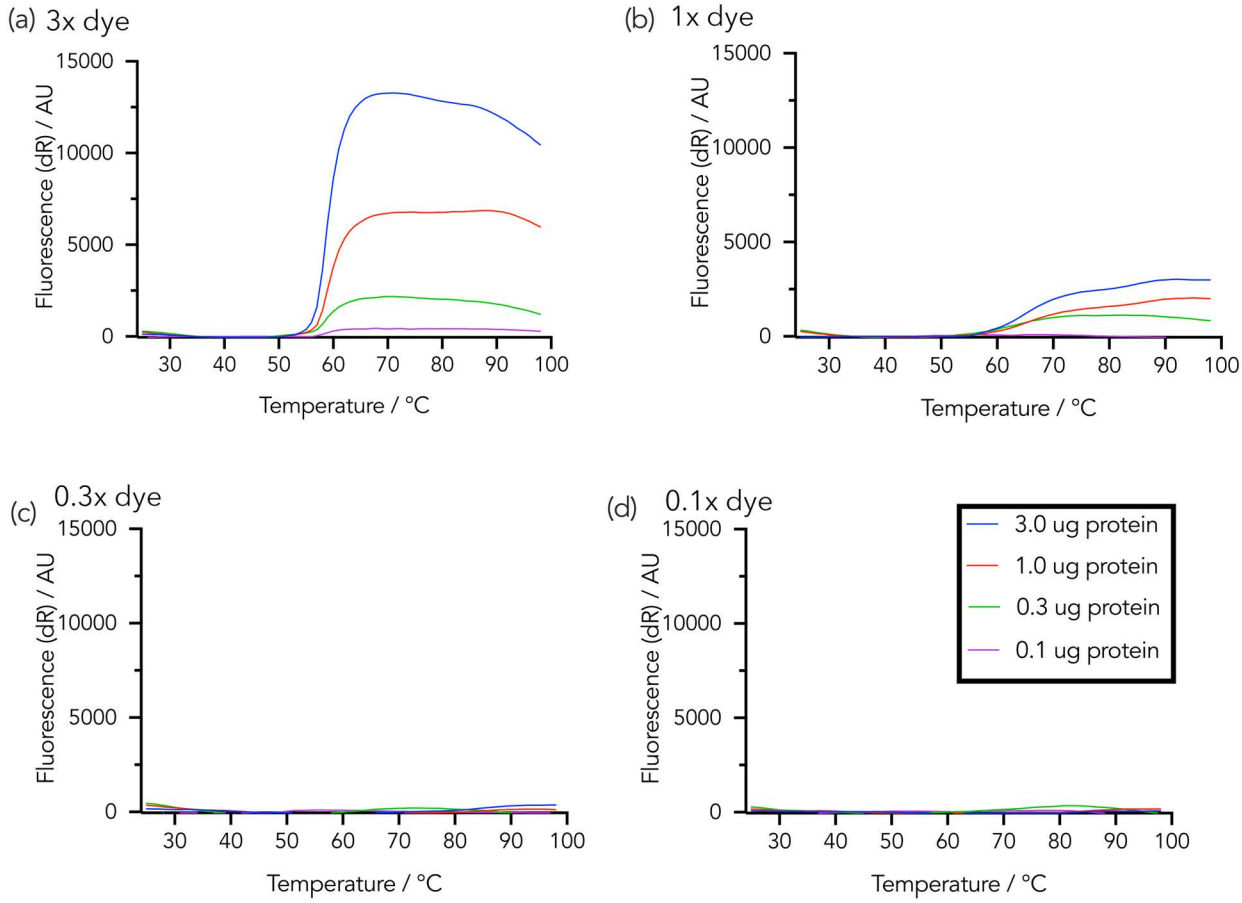


Figure 5.44: First derivative trace of control SYPRO Orange experiment with varied dye concentrations (defined as 'x' stock), as labelled above each graph. Colours for all graphs are shown in the key on the bottom right graph.

5.3 Discussion

This chapter has examined the basis of stability of FMDV VLPs through understanding the formation of S-S bonds and the role metal ions in VLP integrity.

Provisional cryo-EM analysis of AsiaShcc vaccinia-produced VLPs show that, in medium-term cold storage, VLP samples were expanded, with the majority devoid of an intact S-S bridge and almost identical to the subpopulation that did. This suggests that the S-S bridge itself can be maintained despite slight expansion, and that capsids can remain intact after reduction of the 2-fold S-S bridge. It would be interesting to see whether, for a freshly prepared VLP sample, we see a similar-such expansion. Storage itself, assuming that this was not under reducing conditions, should not degrade the bridging interaction, but the slight expansion that we see for all particles may be due to a slight instability over long term storage. The formation of an S-S bridge must somehow be thermodynamically and/or kinetically unfavourable. It would be useful, in future, to attempt crystallisation trials covering a broad range of crystallisation conditions to see if as for A22cc, crystals of a non-expanded subpopulation of AsiaShcc, where the S-S bridge is intact, can be grown, and whether expanded particles can be crystallised.

Analyses of AsiaShcc VLP cryo-EM data presented here are only cursory, and additional, in-depth interrogation of the structures is required to fully understand the driving forces for capsid distortion. Pisa analysis of inter-protomer contact surfaces of well-refined structures of VLPs from different serotypes, for instance, would better highlight conserved inter-protomer interactions between particle classes. Furthermore, given the particle numbers for the AsiaShcc dataset, it would be interesting to re-classify with relaxed symmetry to see if icosahedral averaging masks any significant differences between particles.

Protonation of residue H3140, in addition to H3143, has been previously suggested as a crucial switch for FMDV uncoating, and has been a target for introduction of thermostabilising mutations (Curry et al., 1995; van Vlijmen et al., 1998; Ellard et al., 1999; Scott et al., 2017). It may be informative to perform mutational studies of the highly conserved residue H3140 in our VLPs to establish whether this could be a complementary site for the S-S bridging interaction, and maintain particles in a non-expanded form.

DTT treatment disrupts VLP particles, especially under harsh,10 mM conditions, where particles completely dissociate into pentamers. Cryo-EM analysis of these pentamers show many sim-

ilarities to a previously published pentamer structure derived from acid-treated wild-type A1061 (Malik et al., 2017). Collection of cryo-EM data with a 20 ° tilt greatly improved the angular sampling for the pentamers and additional data processing, perhaps with relaxed symmetry, might reveal sub-populations of particles.

Anomalous density analysis of A22cc crystals with and without a calcium soak suggest that calcium does not bind to the VLP capsids at occupancies above ca. 5 %. This may be due to the calcium soaking time being too short, or the necessity for the presence of genomic material to interact with metal ions. It would be informative to perform stability assays to establish whether calcium ions confer a similar stabilising effect on VLPs to that seen in native FMDV. We do see clear evidence of the disulphide bridges at the two-fold and five-fold axes, confirming that, for vaccinia systems, it is possible for such bridging interactions to form. Crystallisation conditions might also favour the non-expanded S-S bonded particle crystals.

Provisional testing of alternative dyes GloMelt^{TM} (Biotium) for thermofluor analysis show that protein concentration was the limiting factor for signal with a 3 x dye concentration, while signal was almost three-times weaker for GloMelt^{TM}, relative to SYPRO Orange dyes. Collectively, these results indicate that GloMelt^{TM} fails to show significant signal enhancement at a comparable protein concentration to SYPRO Orange (Invitrogen).

Further work may be done to optimise thermofluor analysis of VLP samples. The ramp rate can also influence protein denaturation and is another parameter that could be optimised to tease out more information on VLP dissociation, but at the cost of time efficiency. Increasing the number of readings demands longer assay times, which may exacerbate sample degradation, and eliminate a key advantage of the method, speed. Alternative dyes could be investigated. For instance, 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM) is often used for membrane proteins and reacts with thiol groups of cysteine residues (Alexandrov et al., 2008). This could be an interesting choice for FMDV VLPs in particular, given the high proportion of cysteine residues present, including residues involved in disulphide bridges. Diminished selectivity at low pH and pHs above 8.0 and fluorescent quenching in certain salts may, however, limit its effectiveness (Alexandrov et al., 2008).

Chapter 6

Thermostabilised Poliovirus VLPs as Vaccine Candidates

6.1 Introduction

Poliovirus (PV) is the causative agent of the crippling, and sometimes deadly, infectious disease poliomyelitis. Global efforts towards eradication have made a substantial impact on the decline of global PV cases (Chapter 1, Section 1.1.6).

Moves towards eradication have been helped by the distribution of two types of licensed vaccines; a live-attenuated oral version (Sabin, OPV), generated by repeated passage of wild-type virus (Sabin, 1957); and a formaldehyde-inactivated virus (Salk, IPV) (Salk et al., 1954). Both vaccines are far from ideal; requiring the production of large quantities of infectious PV in containment. OPV has been favoured over IPV due to its easy administration and provision of gastrointestinal tract immunity (Hird and Grassly, 2012). Efficacy of IPV also come under question after inactivation procedures were found to alter antigenic sites for serotypes 2 and 3 in particular (Ferguson et al., 1993).

Vaccines traditionally contained a trivalent mix of all three PV serotypes, often Mahoney, MEF and Saukett strains, respectively. However, this has been replaced by a divalent vaccine, with IPV replacing PV-2 following PV-2's total eradication, and the observation that over 90 % of VDPV cases and 38 % of VAPP cases were due to the PV-2 component (http://polioeradication.org/who-we-are/polio-endgame-strategy-2019-2023/). This highlights a fundamental drawback of OPV; atten-

uated viruses have the ability to acquire virulent mutations, leading to vaccine-associated paralytic poliomyelitis (VAPP) in vaccinated individuals or their contacts. In rare cases, VDPVs can continue to circulate (cVDPV), thereby increasing the chances of virulence enhancing mutations. In addition, OPV's thermo-instability necessitates a cold-chain, making transport to more remote regions challenging. These factors have prompted the WHO to call for novel polio vaccines that circumvent issues associated with OPV and IPV (Zambon and Martin, 2018). As discussed in Chapter 1 Section 1.4.1, VLPs offer a tractable vehicle for vaccine development; mimicking native viral capsid structure and are thus, in principle, capable of generating a protective response from the host immune system, while remaining non-infectious and therefore safe.

The first PV VLPs were reported over 20 years ago, but particles were too unstable for vaccine preparations (Urakawa et al., 1989; Rombaut and Jore, 1997). As empty immature capsids, VLPs lack both stabilising RNA interactions, and the rearrangement of internal capsid proteins upon cleavage of VP0 into VP2 and VP4 (Basavappa et al., 1994), thereby making them more prone to D-to-C antigenic switching than their native counterparts (see Chapter 1). This has prompted efforts to maintain VLP capsids in the D antigenic form via a series of stabilising mutations, and/or by using capsid-stabilising additives (Adeyemi et al., 2017; Fox et al., 2017). In the former case, selection from temperature sensitive mutants at non-permissive temperatures, combined with deep sequencing, provided insight into potential stabilising mutations, Table 6.1 (Fox et al., 2017).

Conversion to the C form occurs concomitantly with pocket factor expulsion from within the VP1 core (see Chapter 1 Section 1.1.6) and it has been shown that incubation of native capsids with more potent fatty acid-like molecules increases the energetic barrier to capsid expansion to C-antigenic particles (Tsang et al., 2000; Rombaut et al., 1990).

A recently published plant-produced PV-3 Saukett strain, containing the panel of previously identified stabilising mutations (some of which are listed in Table 6.1) (Fox et al., 2017) is protective in mice and maintains native antigenicity (Marsian et al., 2017), but, intriguingly, little evidence of a bound endogenous pocket factor was seen in the 3.6 Å resolution cryo-EM reconstructions (Marsian et al., 2017; De Palma et al., 2008). Nevertheless, consistent with pocket factor stabilisation previously observed for yeast expressed PV VLPs (Rombaut and Jore, 1997), P-site binders do bind to plant-derived VLPs and enhance particle stability (Marsian et al., 2017).

Yeast expression systems are often attractive VLP production hosts, owing to their eukaryotic

Table 6.1: Summary of capsid stabilising mutations used for the two VLPs studied in this Chapter: Mahoney-SC6b and Saukett-SC8. The latter set of mutations are reported in Fox et al. (2017). A sequence alignment of both constructs, alongside wild-type PV-1 Mahoney and PV-3 Sabin may be found in the Appendix.

Construct	Mutation	Context
PV-1 SC6b		
	R4018G	Inner network near 3-fold.
	T2025A	Pentamer interface
	D2057E	Pentamer interface
	L3119M	VP2/VP3 interface
	Q3178L	Protomer interface
	G1006L	Capsid interior, close to 5-fold
	S1007E	Capsid interior, close to 5-fold
	S1009M	Capsid interior, close to 5-fold
	T1010I	Capsid interior, close to 5-fold
	H1248P	Protomer interface
PV-3 SC8		
	T4067A	Close to 3-fold, capsid interior.
	L2018I	β sheet at pentamer interface.
	L2215M	Protomer interface.
	D2241E	VP2/VP3 interface, buried.
	H3019Y	Internal network, below 5-fold.
	L3085F	β sheet at pentamer interface.
	T1105M	North wall of canyon.
	F1132L	Pocket.

post-translational modification (PTM) capabilities, robustness, cheapness (in terms of media) and scalability. Scalability comes with the caveat that stress responses induced by high-yield cultures abolishes quality checks by the yeast endoplasmic reticular (ER), resulting in a higher incidence of mis-folded proteins. Furthermore, fatty acid production is more akin to mammalian systems, than that of other eukaryotes such as plants, meaning that 'natural' pocket factors are more likely to be available for incorporation as a pocket-factor into nascent VLP capsids during assembly (see below) (Tuller et al., 1999; Viljoen et al., 1986). A number of other enterovirus VLPs have also been produced using yeast expression, including EV-A71 (Wang et al., 2016; Li et al., 2013; Lyu et al., 2015; Yang et al., 2019) and CV-A16 (Zhao et al., 2013; Zhang et al., 2016). Yeast expressed PV VLPs have been previously produced in yeast expression systems, but efforts were hampered by particle instability (Rombaut and Jore, 1997).

With previously identified thermostabilising mutations that stabilise particles in a D antigenic state (Fox et al., 2017), PV-1 and PV-3 VLP production was attempted using yeast expression by collaborators at the University of Leeds (Faculty of Biological Sciences). In this section, we report structures of PV-1 Mahoney and PV-3 Saukett VLPs from these experiments (details of VLP production and purification are outlined in Chapter 2, Section 2.2.3).

It should be noted that references to 'stability' in this chapter are to thermostability, which is assumed to be directly linked to D/C antigenic switching.

6.2 PV-1 SC6b

PV-1 SC6b grids were prepared on glow-discharged C-flat 2/1 grids (Protochips), plunge frozen in liquid ethane and screened on a Glacios microscope (Thermo Fisher, see Chapter 2, Section 2.5.2).

Grid holes were relatively sparsely populated with PV-1 SC6b particles, with the majority of sample immobilised on the inter-hole carbon. Collection set up was further hindered by the number of usable squares, since C-flat grids are susceptible to damage when mounting in autogrids. Given that a greater number of particles would give a more reliable ratio of antigenic states in the sample, rapid linear collection mode was set up using a Falcon III (FIII, Thermo Fisher) camera on a Glacios microscope at a 120Kx nominal magnification, corresponding to a pixel size of 1.22 Å/pix.

6.2.1 Data Processing

A total of 1204 particles were selected from 385 micrographs using Warp, before being filtered by 2D and 3D classification in Relion (Figure 6.1 and Chapter 2, Section 2.5.4) and one of the best micrographs can be seen in Figure 6.2(a) (Zivanov et al., 2018; Tegunov and Cramer, 2018).

The two predominant classes from 3D classification (58.2 % and 22.8 % of the original particles picked, see Figure 6.1) were refined separately to 5.7 Å and 6.8 Å resolution, respectively, Figure 6.2(b, c, d). The two predominant classes had a partially 'filled' capsid interior (see Figure 6.2(b) and (c)), with the interior of (b) relatively more condensed Figure 6.2(b). This internal density is unlikely to be due to the presence of genome, since the A260/280 readings for PV-1 SC6b, recorded as 1.32, corresponding to approximately 10 % RNA (Glasel, 1995).

From inspection of the reconstructions, the two classes were labelled C and D (C being the most populous and resembling the structure expected the C antigenic state, whilst the smaller population of particles appeared similar to the native (D) antigenic form). Indeed, as may be seen in radially coloured reconstructed volumes, relative to C, D has a less pronounced mesa, and no opening at the two-fold axis (Figure 6.3).

6.2.2 Structure Analysis

In order to confirm suspected antigenicity for particles C and D, protomers of both PV-1 Mahoney 135S (PDB: 3J48) (Butan et al., 2014) and PV-1 Mahoney native (PDB: 1HXS) (Miller et al., 2001) were rigid-body fitted into each map using Coot (Emsley and Cowtan, 2004; Emsley et al., 2010). For the native initial model used, the N-terminus of VP1 (1-24); N-terminus VP2 (1-30) and VP4 were removed since these were far extended from the bulk density, and were not resolved in the higher resolution PV-3 SC8 structures presented below. The resolution of both C and D reconstructions (5.7 Å and 6.8 Å) was deemed too low for real-space refinement.

For C, the map-to-model CC value of both native and 135S rigid body fitted maps was very low (42 % and 48 % for 135S and native structures, respectively) (Adams et al., 2010, 2011). For the lower resolution map D, CCs were marginally better with CC values of 55 % and 53 % for native and 135S models, respectively.

To attain a better model, a MD flexible fit was run using Namdinator (Kidmose et al. (2019),

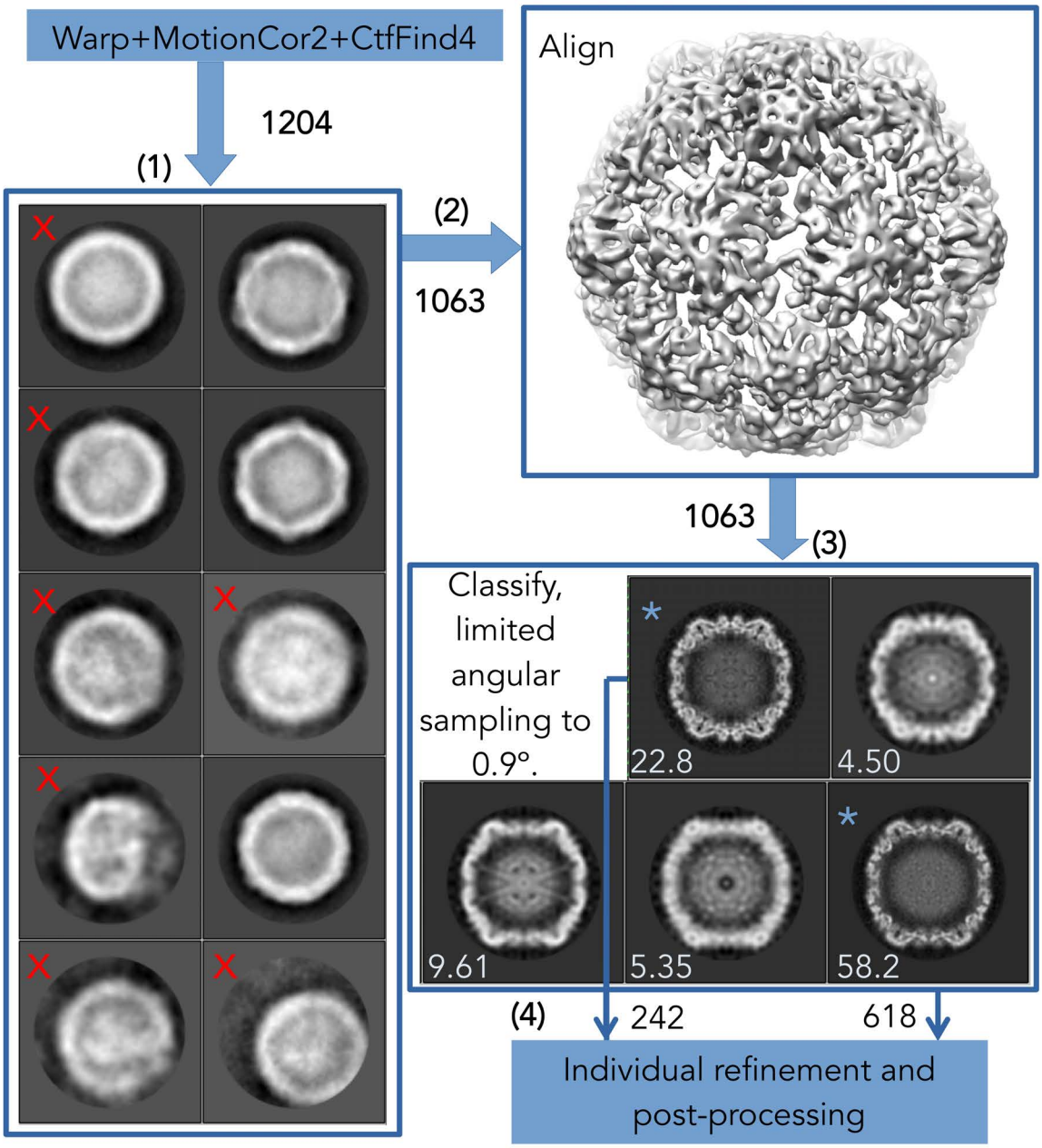


Figure 6.1: Processing workflow for PV-1 SC6b. Each stage of processing is indicated by a blue arrow next to which associated numbers of particles used is stated. Prior to processing with Relion, movies were initially aligned and particles picked by Warp (see Chapter 2 Section 2.5.4 (Tegunov and Cramer, 2018). The coordinates from Warp were then used to extract 1204 particles from micrographs processed with MotionCor2 and CtfFind4 (Zheng et al., 2017; Rohou and Grigorieff, 2015). (1) Extracted particles were classified into ten classes to filter out bad particles. The bad particles are marked with red crosses. (2) Particles were then aligned, and the resulting initial reconstruction is shown in grey. This appeared to have an opening at the 2-fold axis. (3) Aligned particles were classified with a limited angular sampling, and the two predominant classes were selected and refined separately before undergoing post-processing. The final projections are shown in Figure 6.2.

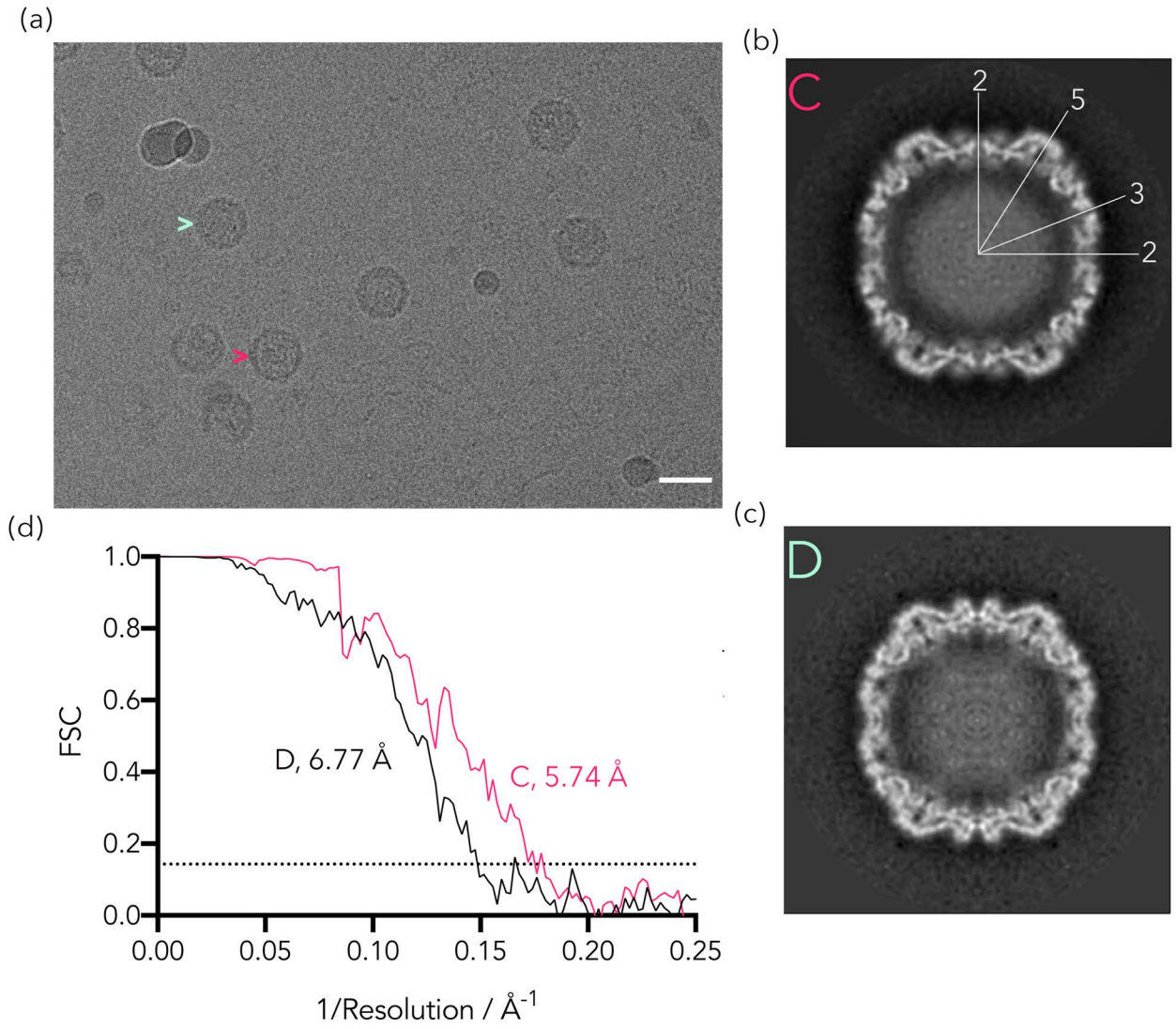


Figure 6.2: (a) Example cryo-EM micrograph of PV-1 SC6b. Particles seemed heterogeneous, with some centres appearing more granular in appearance, as indicated by '>' arrows. Scale bar = 300 Å. (b) and (c) Projection images from refinement of the two predominant 3D classes. (d) FSC curves for 'D' antigenic (dark grey) and 'C' antigenic (pink), see labels particles. The dashed line marks the FSC of 0.143.

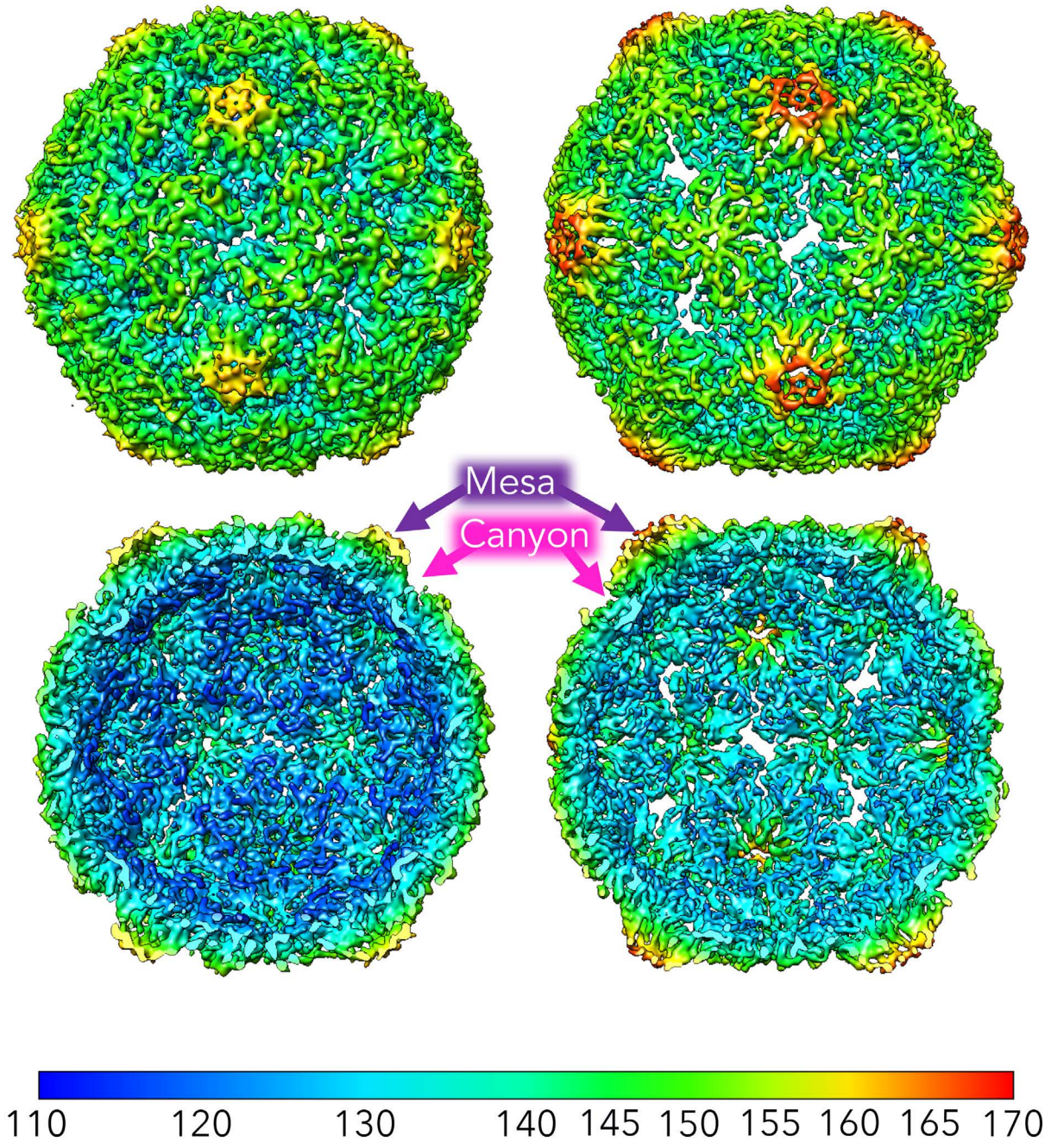


Figure 6.3: Top: Volumes of the two predominant PV-1 SC6b classes; D (left) and C (right) antigenic particles, viewed along the two-fold axis. Surfaces are radially coloured, each from the centre of each VLP. Colour bar represents the radius in Å from each capsid centre. Bottom: Central slices through each antigenic form showing the genome-free internal surface along the two-fold axis.

Chapter 2 Section 2.5.5) for both C and D maps and each respective rigid body fitted native model (PDB: 1HXS) (Miller et al., 2001). It was not possible to run the same analysis using the 135S model used because this only contains coordinates for C_α and the AutoPSF component in the Namdinator pipeline (www.ks.uiuc.edu/Research/vmd/plugins/autopsf/) requires full residues as input to generate parameter files. In future, a higher resolution model, such as that of 135S PV complexed with VHH PVSS21E (reported resolution 3.8 Å, PDB: 3JBG) should be used (Strauss et al., 2016). For our analyses, the RMS of the Namdinator output versus the original 135S model was computed using SHP (Stuart et al., 1979), and results are summarised in Table 6.2 (Kidmose et al., 2019). For both C and D, a better map-to-model agreement, based on CC values, was achieved following Namdinator (CC of 69 % versus 44 % for C and 72 % versus 55 % for D, Table 6.2(a)).

Assessment of the C/D Antigenic Divide

The resulting model for D seems to be more similar to that of the native model (RMS of 0.89 Å, 686 equivalences) than C (RMS 1.26 Å, 695 equivalences), Table 6.2(b), consistent with our assignment of particle D being closer to the native antigenic state. RMS distances for both C and D versus 135S PV-1 Mahoney were relatively higher (1.29 Å, 1.32 Å, respectively), but for fewer equivalences (591, 592).

Features in C/D maps, and comparison with the rigid-body fitted native and 135S models also support the C/D assignment. The dominant class, C, derived from 619 particles, has a significant opening at the two-fold axis of approximately 14.6 Å (determined by taking the average distance of C_αs of the two opposing VP2 helices; G2090-*Y2098*, Q2094-*Q2094*, Y2098-*G2090*, where residues from each protomer are distinguished by italic text), see Figure 6.4(a,c), 6.5. By contrast, the gap between opposing 2-fold helices was significantly shorter for the smaller subset of data, D, Figure 6.4, 6.5(b-d), with an average spacing between opposing 2-fold VP2 helices of ca. 9.3 Å (approximated as before).

Little to no density can be seen for the VP2 EF loop for C, by contrast to that of D (Figure 6.4(a,b)). The VP1 GH, EF and VP3 GH loops are also significantly shifted for C, relative to D, although density fails to account for the entire loop conformation (Figure 6.5(c,d)) and no density could be seen in the pocket factor site for either reconstruction, although the limited resolution probably precluded pocket factor resolution.

Table 6.2: (a) Summary of correlation coefficients prior to ('Initial', i.e. from rigid body fitting in Coot) and after ('MD') fitting with Namdinator for particles C and D (Emsley and Cowtan, 2004; Emsley et al., 2010; Kidmose et al., 2019). The CC mask and CC main chain scores were computed by Phenix (Adams et al., 2010, 2011). Note that the 'VMD CCC' score is the cross correlation coefficient between the input map and PDB file, computed by Namdinator. (b) RMS distances (Å) determined by SHP (Stuart et al., 1979) for particles C, D, native PV-1 Mahoney (PDB: 1HXS) and 135S PV-1 Mahoney (PDB: 3J48) (Miller et al., 2001; Butan et al., 2014). Cells are coloured according to RMS distance. (c) Matrix of equivalences for the RMS values listed in (b), also coloured by value. SHP RMS distances should be re-determined with a more permissive sigma threshold for more reliable comparisons (i.e. comparable equivalences) in future.

(a)

	C		D	
	Initial	MD	Initial	MD
Clash score	2.31	1.28	2.45	1.18
CC mask	0.44	0.69	0.55	0.72
VMD CCC	0.38	0.44	0.40	0.44
CC main chain	0.48	0.64	0.58	0.62

(b)

RMS	D	C	Native	135S
D		1.16	0.90	1.32
C	1.19		1.26	1.30
Native	0.89	1.26		1.28
135S	1.32	1.29	1.27	

(c)

Equivalences	D	C	Native	135S
D		662	686	591
C	670		695	593
Native	686	696		609
135S	591	592	606	

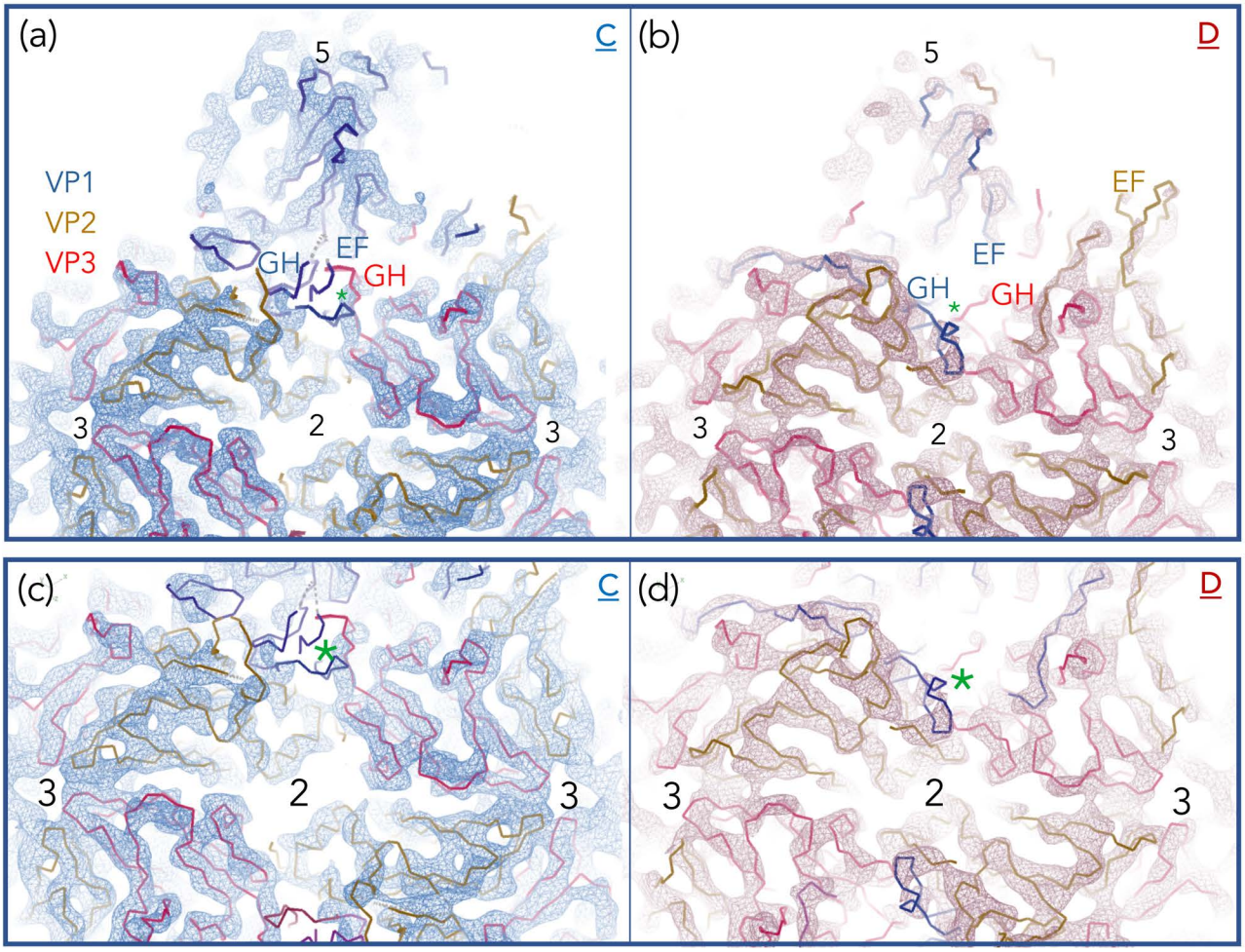


Figure 6.4: (a) Overview of the capsid surface topology for data sets C (blue) and (b) D (red), inspired by Figure 4 in Butan et al. (2014). Zoomed in versions of (a) and (b), centred at the 2-fold axis can be seen in (c) and (d). The 2- 3- and 5-fold axes are marked in black text. Pertinent loops are labelled in text associated with VP (blue VP1, green VP3 and magenta VP3). The green asterisk corresponds to an opening at the quasi-3-fold as identified previously for 135S particles Butan et al. (2014) that occurs following rearrangements of the labelled loops. As can be seen comparing the blue and red densities, there is a substantial opening at the 2-fold axis for the C reconstruction, and is shown more clearly in the view below, and in Figure 6.5.

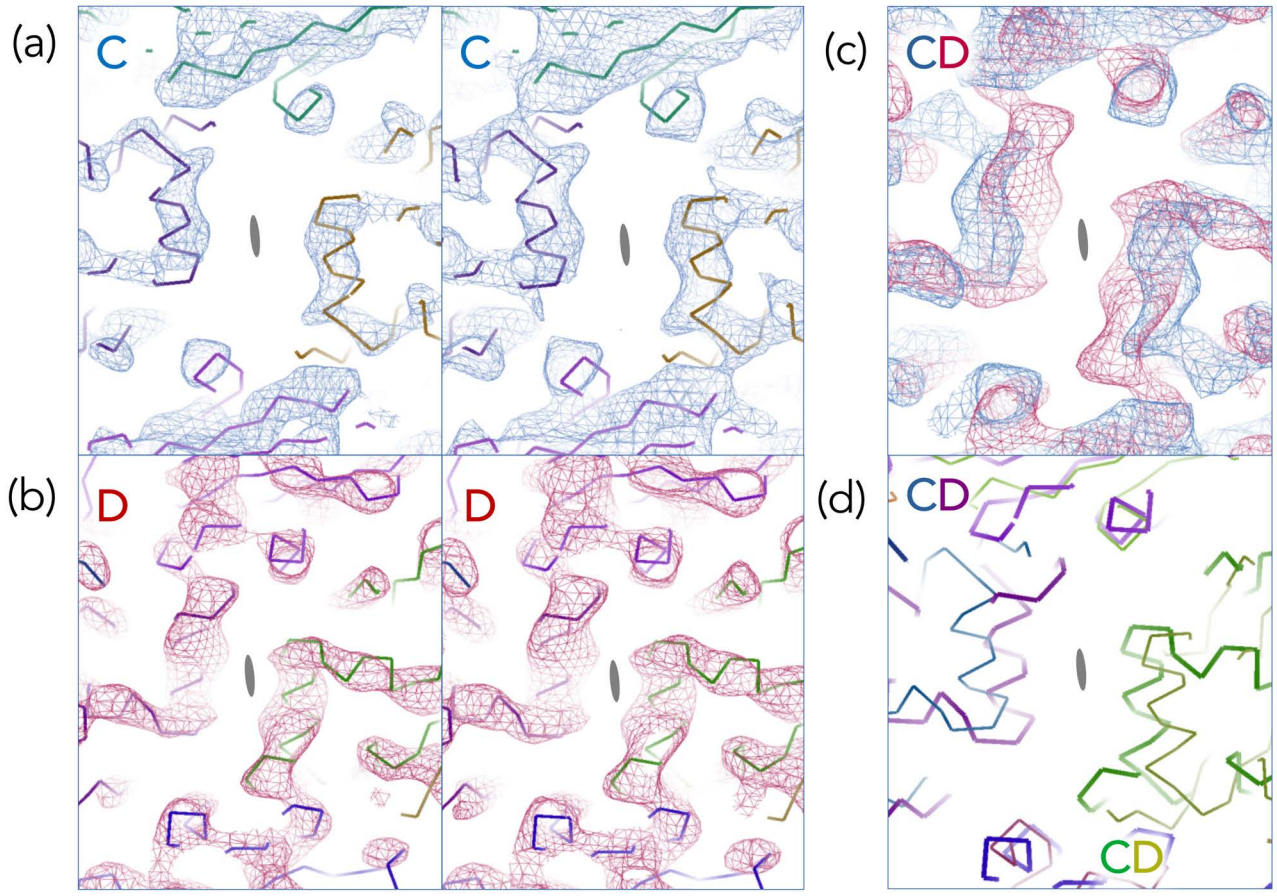


Figure 6.5: (a,b) Maps and models for C and D particles at (left) high and (right) low contour levels (15 % difference in sigma level between high and low). Models are shown in ribbon representation and each chain coloured separately. (c) Superposition of C and D maps to demonstrate the difference in 2-fold opening. (d) Overlay of models for C and D. C is shown with thin ribbons and its first protomer is in blue, second protomer in lime. D is shown in thicker ribbons and its first protomer is in purple, its second in green.

In addition, for C, the map clearly shows a more pronounced, slightly twisted, broad and open mesa (see Figures 6.2(b), 6.3 right, Figure 6.6(a)) with a wider channel down its centre, and, bigger holes surrounding its central flower (Figure 6.6(a-c)). For C, at the 5-fold mesa, there is no density for the DE loop residues 1148-1151, but it is conceivable that the tip of this loop, around residue 1149, sits closer into the 5-fold than modelled in PV-1 135S, serving to account for the central ring of density at the top of the mesa (arrow, Figure 6.6(c,d)).

Another typical feature of 135S particles is the flattening of the rim around the three-fold symmetry axis that is forged by the VP2 HI loop (residues 2240-2246). The fit of the 135S model in map C seems to be in good agreement in this region, with the exception of residues 2243-2244 (green arrow, Figure 6.7), which may instead occupy the region of unmodelled density that sits close to the adjacent protomer, indicated by a pink arrow in Figure 6.7. More broadly, the three-fold appears to be flattened and more obviously triangular in shape for C, relative to subset D (Figure 6.7).

Subtle differences between C and the rigid body 135S map can also be seen for the VP3 portion that encircles the 3-fold, where C's density supports a loop conformation slightly more retracted from the three-fold centre, and, in this case, the rigid fit for the mature PV-1 Mahoney model (PDB: 1HXS) is closer (Miller et al., 2001), see Figure 6.7(d). Similarly, for the BC loop residues 3073-3080 that rest above, but slightly further back from the VP3 three-fold HI loop, C's density supports a more upright loop conformation, which is closer to what we see for mature PV-1 (green and pink arrows, Figure 6.7).

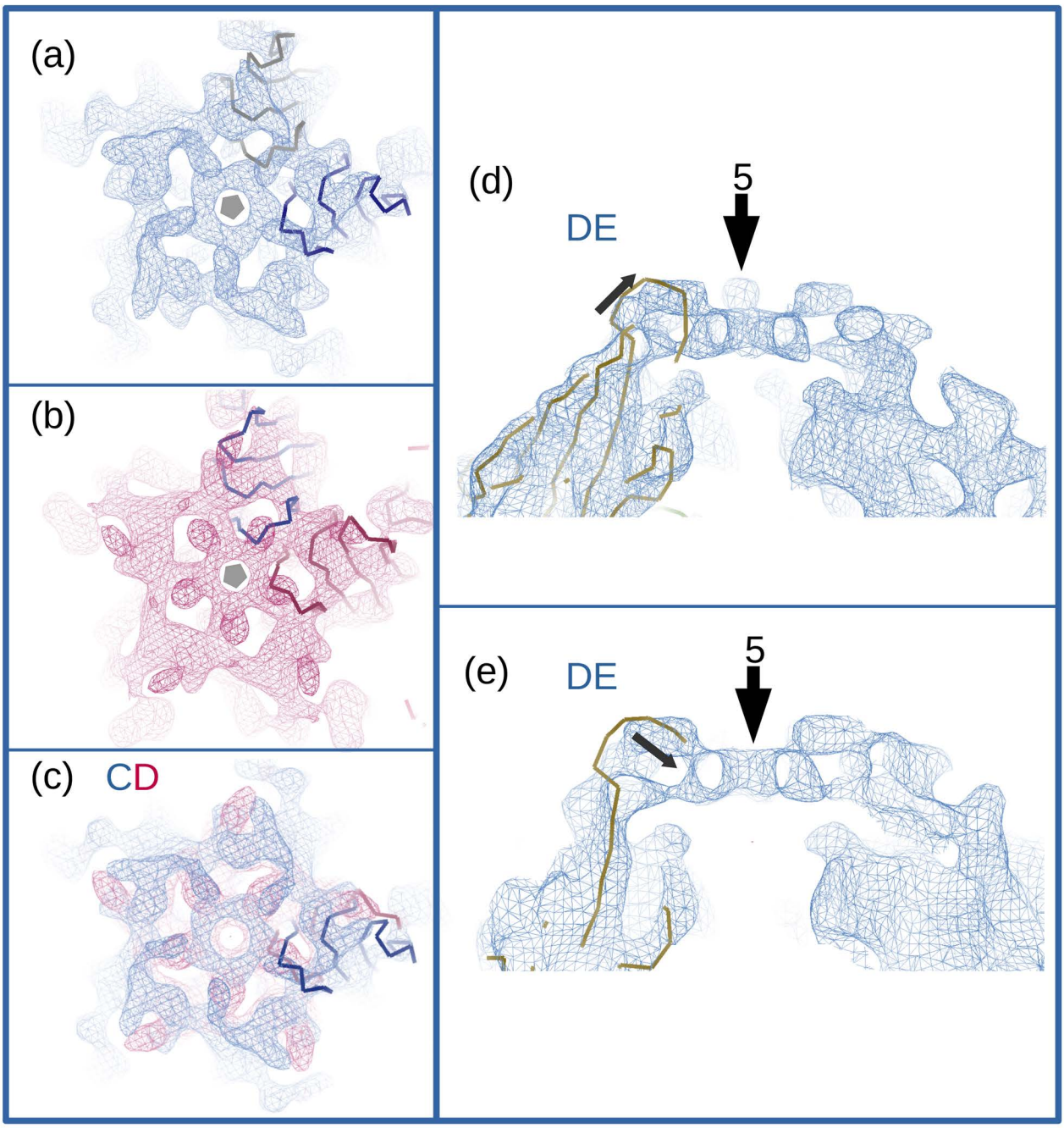


Figure 6.6: (a, b) Maps for C (blue) and D (red) particles. Models for two protomers are shown for two vertices to aid comparison. (c) C and D maps superimposed, with models from one protomer overlaid to demonstrate that the model for C (blue) sits much higher than for D (red), which is less clear to see. (d) Side view of protomer fitted into 5-fold. (e) Zoomed in view of (d).

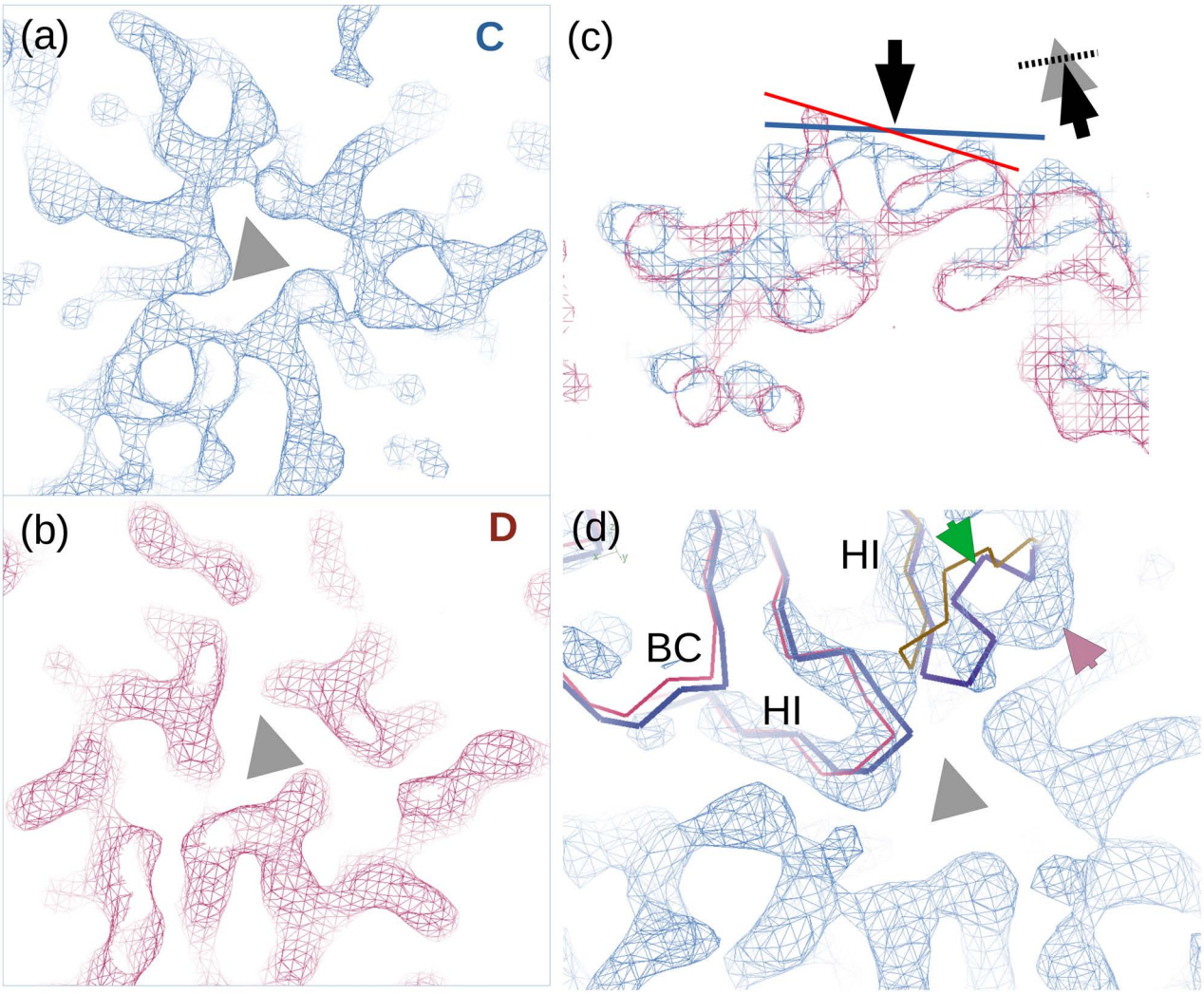


Figure 6.7: (a, b) Maps for C (blue) and D (red) particles. (c) Slice through the capsid maps for C (blue) and D (red), showing one wall of the three-fold channel, facing a vertex. The lines show the slope of the topmost part of the channel. (d) Model from one protomer and map for C, shown with bold ribbon. The green arrow points to the part of the model that fails to fit as well into the map, and the black arrow indicates the direction that this may shift to. Loop names are labelled. The thinner ribbons show a wild-type Mahoney PV structure (PDB: 1HXS), and are coloured by VP (red, VP3, gold VP2) (Miller et al., 2001).

6.3 PV-3 SC8

6.3.1 Data Collection

To circumvent particle accumulation outside of acquisition holes experienced with PV-1 SC6b, lacey grids (Agar Scientific) with an ultra-thin carbon layer were used for PV-3 SC8 preparations. Although the particle distribution seemed somewhat better (Figure 6.8), inconsistencies in the lace distribution (an inherent problem with lacey grids), and the less efficient acquisition set up for lacey grids in EPU led to the decision to collected in linear mode, this time at a somewhat lower nominal magnification of 92 kX, with the same microscope and detector as before, corresponding to a pixel size of 1.54 Å/pix. In hindsight, a higher magnification, and possibly even counting mode collection, would have been more suitable, especially given the more extended data collection time used for this sample.

Acquired micrographs contained a number of particles that appeared to be partially filled, possibly due to encapsulation of host material, and smaller particles were also observed (this latter observation applies to both PV-1 and PV-3 data collections, for instance, salmon circle, Figure 6.8(a)). These smaller particles appeared more rectangular relative to the 'flowers' observed in FMDV VLP micrographs (see Chapter 5, Section 5.2.1), appear to associate with PV capsids (arrow-marked particles in Figure 6.8(a)). These may be yeast derived, possibly chaperones produced by yeast cells during over-expression. Indeed, micrographs show some similarities to those recently resented for HSP104, see Figure 6.8(e) (Lee et al., 2019). It would be interesting to see if this can be picked up in an SDS-page gel, or, with more micrographs, picked out and processed. No viral pentamers were observed in either set of micrographs.

PV-3 SC8 data were processed as described in Chapter 2, Section 2.5.4. The PV-3 SC8 reconstructions were of higher resolution than that of PV-1-SC6b, due to the far greater number of particles used: 1204 PV1- SC6b particles, versus 18,716 PV-3 SC8 particles (see Table 6.3).

6.3.2 Data Processing

The processing workflow can be found in Figure 6.9. The final two reconstructions were at respective resolutions of 3.39 Å and 4.33 Å (FSC = 0.143 criterion, see Figure 6.8). For the sake of clarity, these two reconstructions will be termed states 'I' and 'II', respectively.

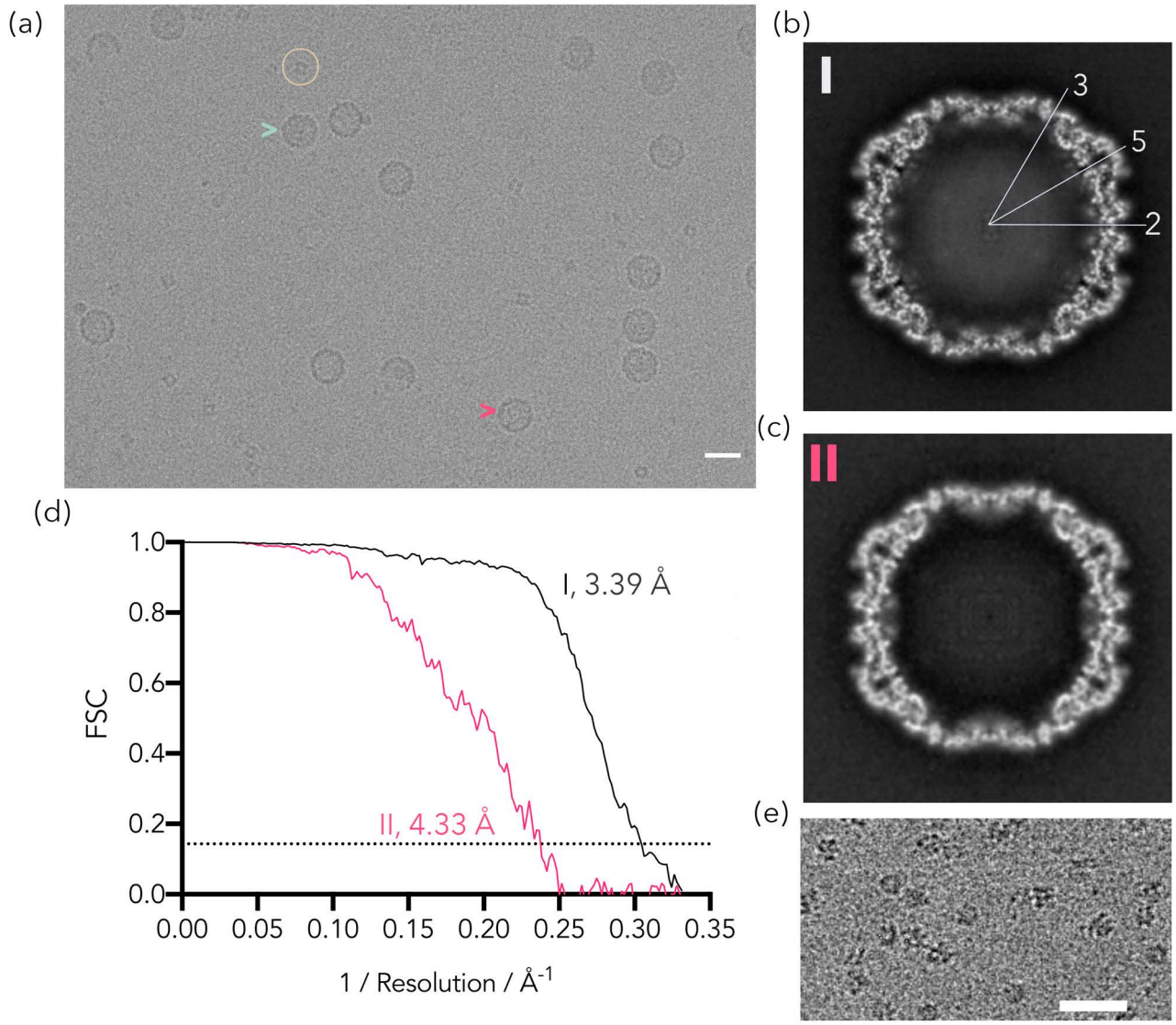


Figure 6.8: (a) Example micrograph, with particle I marked with a light green arrow, particle II with a pink arrow. The mysterious small particle is circled in salmon. Scale bar (bottom right) = 300 Å. (b) Central slice of final reconstruction for particle I. The 2, 3 and 5 fold axes are marked. (c) Central slice of the final reconstruction of particle II. Orientation should be analogous to that of panel (b). (d) Gold-standard FSC curves for particles I and II. The dotted line corresponds to FSC = 0.143. Final resolutions, using the FSC = 0.143 criterion are marked for each particle. (e) Cropped section of micrograph of HSP104 shown in Figure 1D from Lee et al. (2019). These HSP104 particles show some resemblance to the particle circled in salmon in (a). Note that this particular HSP104 dataset were expressed in *E. coli* BL21, but the final reconstruction was found to be very similar to yeast-derived HSP104 (Lee et al., 2019). The scale bar corresponds to 250 Å. Permissions have been granted via RightsLink from Elsevier under a Creative Commons license CC BY NC ND.

Table 6.3: Processing statistics for PV-3 SC8 classes. CCmask was determined by Phenix (Adams et al., 2010, 2011). The resolution reported corresponds to the resolution reported based on the gold-standard FSC=0.143 criterion. 'Particles' indicates the number of particles used in the reconstruction.

PV-3 SC8 class	I	II
Resolution Å	3.29	4.20
Particles	6749	3539
RMS		
Bonds / Å	0.007	0.005
Angles / °	0.70	0.70
Protein Atoms	5740	5597
Rotamer Outliers / %	4.62	6.31
Ramachandran / %		
Favoured	93.8	93.8
Allowed	6.22	6.09
Outliers	0.00	0.14
Model-map fit		
CCmask	0.79	0.78
CCvolume	0.79	0.78

A model of plant derived PV-3 SC8 (PDB: 5O5B) (Marsian et al., 2017) was rigid body fitted into each map using Coot (Emsley and Cowtan, 2004; Emsley et al., 2010), before real-space refinement in Phenix using protomers arranged into an 8mer (Adams et al., 2010, 2011). Data processing statistics may be found in Table 6.3.

Particle I versus Particle II

To compensate for the resolution difference between reconstructions I and II, a 1.2 σ gaussian filter was applied to I's map I, and, as shown in Figure 6.10, filtered I and II appear practically identical. Consistent with this observation, the models for I and II, refined independently from mature PV-3 (as described above), appear almost identical for VP1-3 Figure 6.11, excepting less well ordered regions, where map density was weaker or absent (see Table 6.4).

As may be seen in Figure 6.12, expansion at the two-fold axis is less pronounced than that of the 'C' particle described above (Figure 6.3), suggesting that both reconstructions may be of D

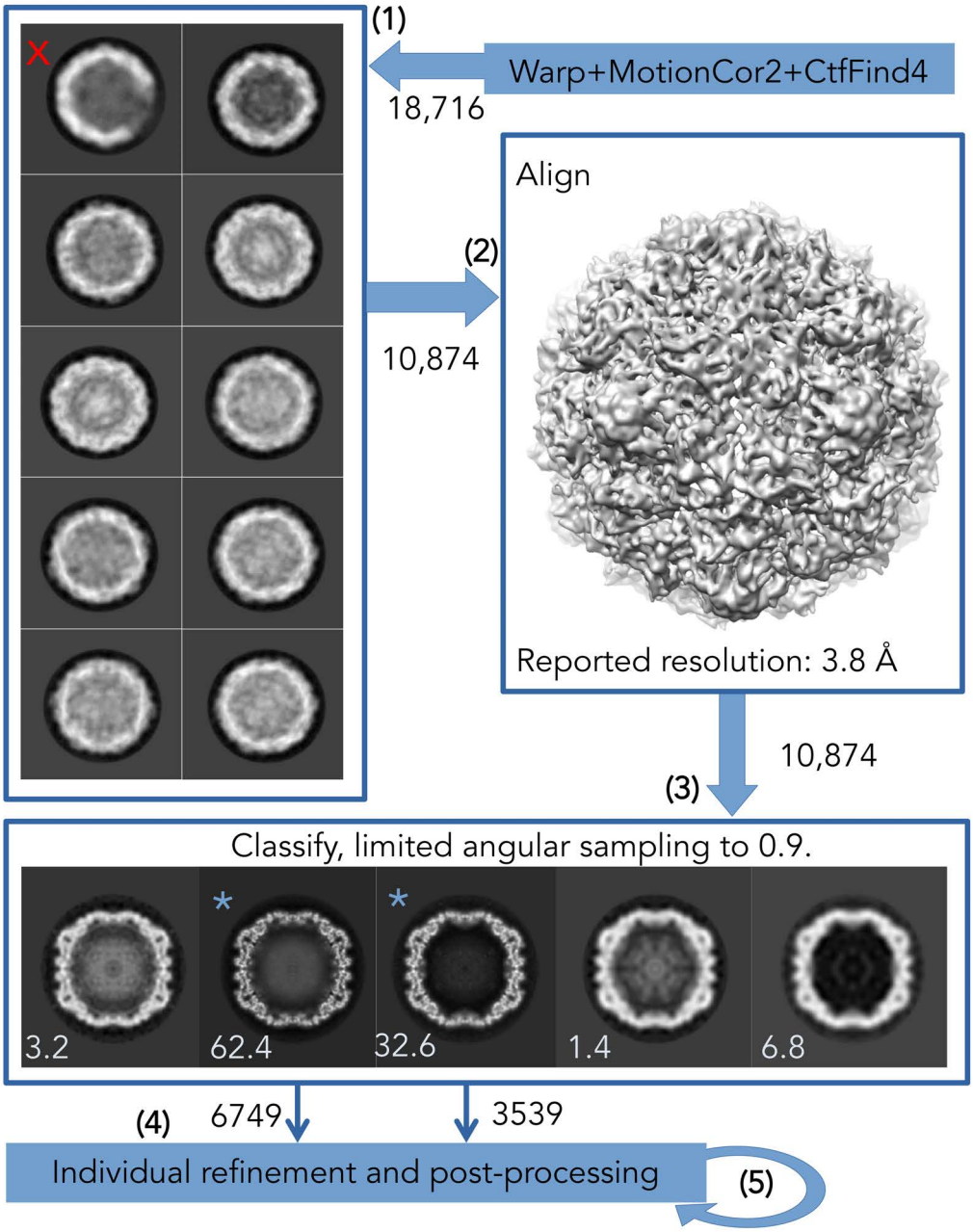


Figure 6.9: Processing workflow for PV-3 SC8. Numbers of particles at each stage are shown next to the blue arrow. (1) A total of 10,874 particles were filtered from 18,716 selected particles (selected from 2496 micrographs) during 2D classification. (2) Particles were then aligned, with icosahedral symmetry, without 3D classification, yielding a map of 3.8 Å resolution and appeared to be of D antigenicity (no gap at the two-fold, and no mesa expansion at the 5-fold). (3) Subsequent classification with limited angular searches unveiled two distinct classes of 6749 and 3539 particles, which were then (4) refined separately before (5) a single cycle of post-processing and per-particle CTF refinement. Electron density slices of the final reconstructions are shown in Figure 6.8.

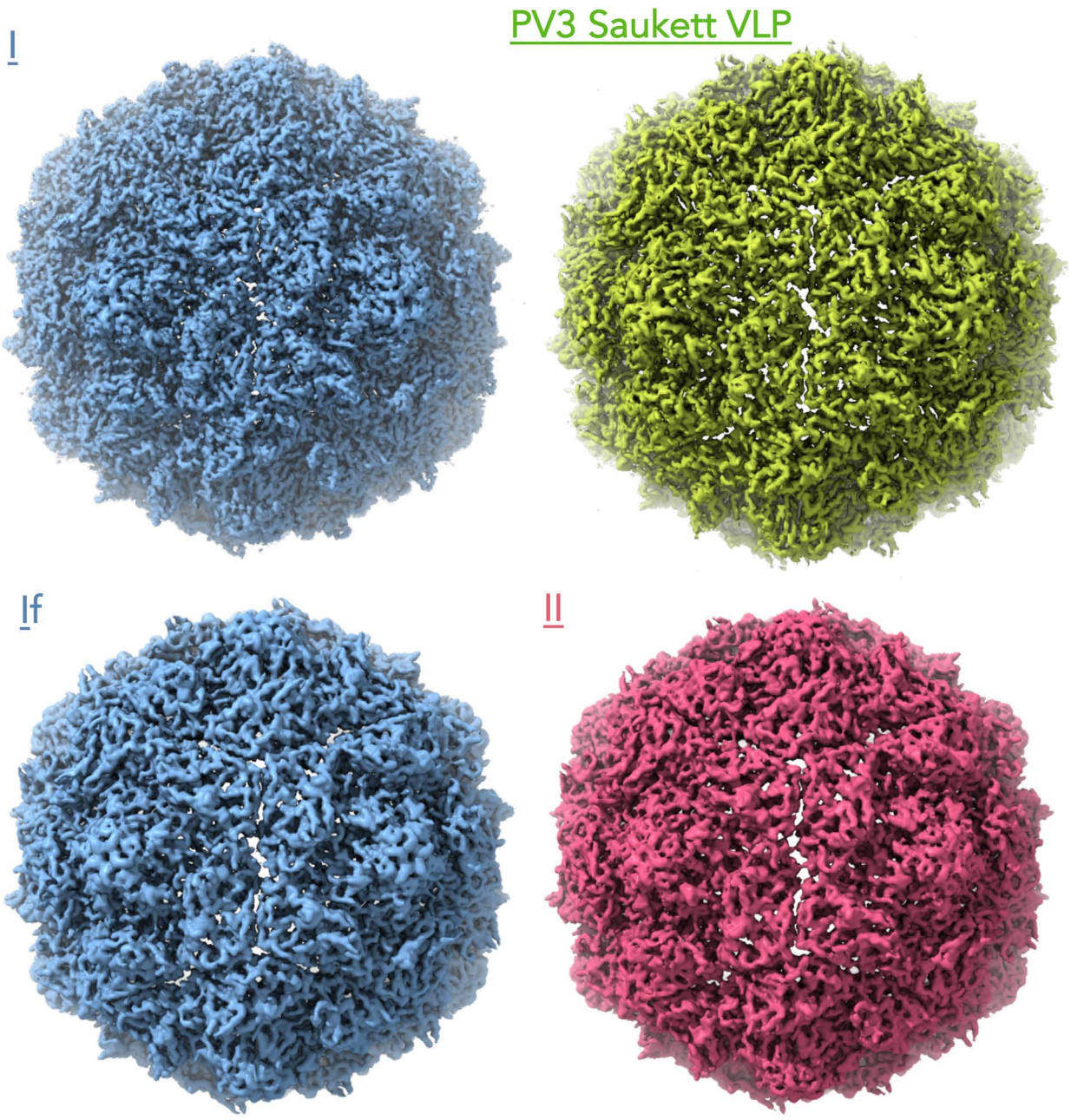


Figure 6.10: Models of particles I (blue) and I filtered, labelled If, the previous plant-derived PV Saukett structure (PDB: 5O5B, green) (Marsian et al., 2017) and II (red).

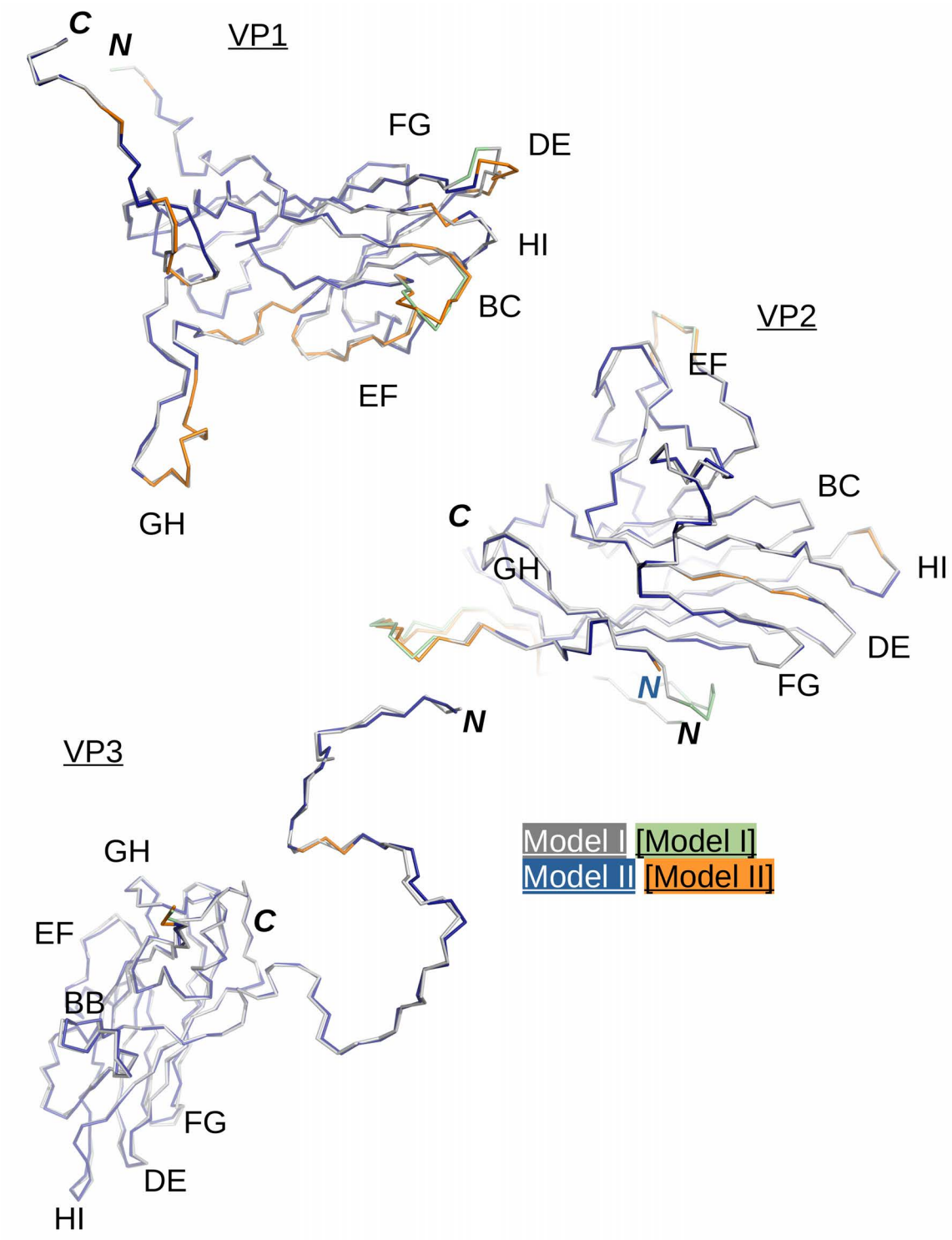


Figure 6.11: Comparison of protomers of real space refined models I (grey) and II (blue) for each VP (as labelled, top to bottom: VP1-3). Regions where map quality was relatively poor are coloured in orange for II and green for I. Loops are marked in black text, and N and C termini in black, bold, italic typeface, with the very different N terminus for II in blue type for VP2.

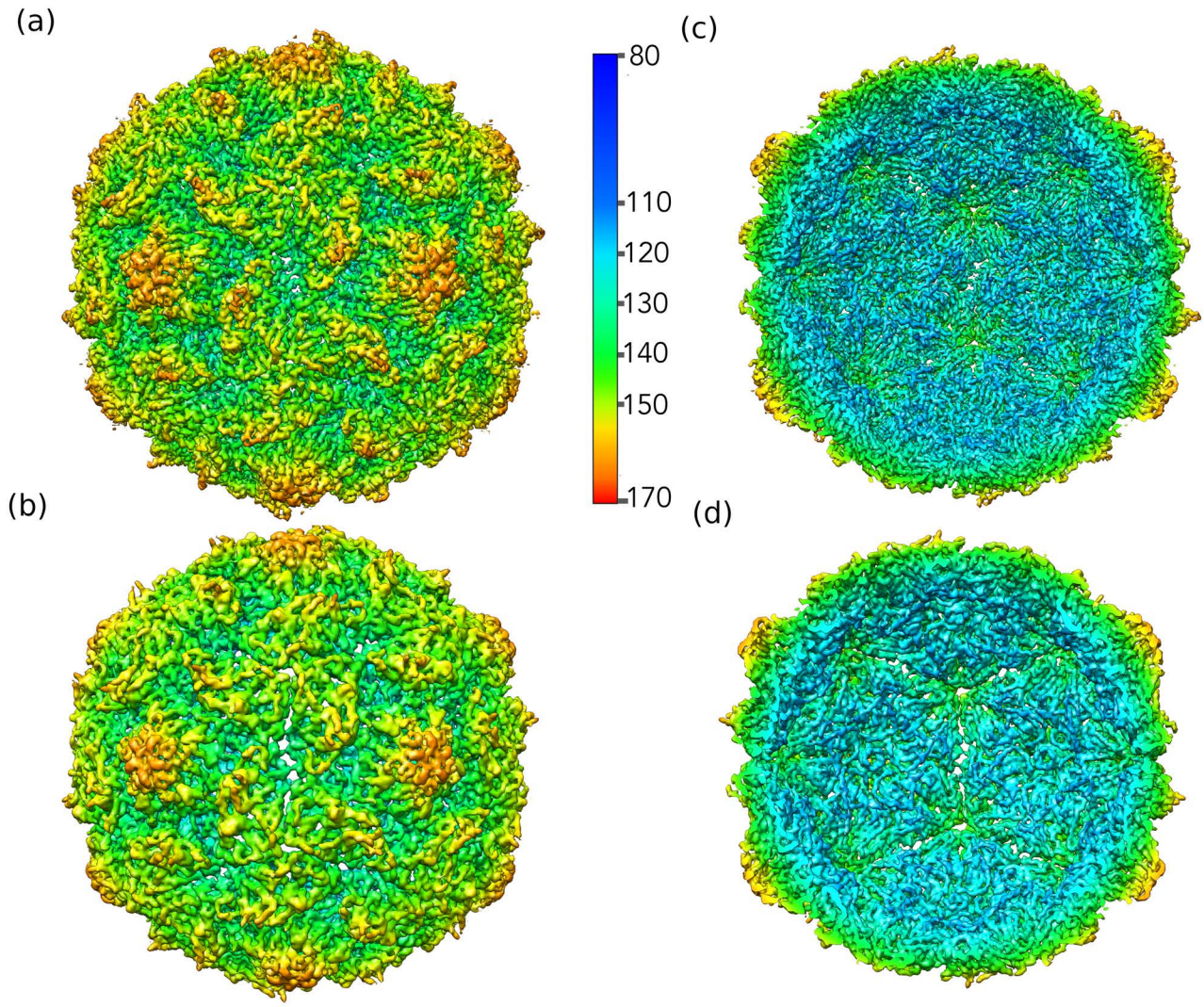


Figure 6.12: Volumes of PV-3 SC8 coloured by radius, as defined by the colour bar adjacent to (a), in Å (a) particle I and its interior (b); (c) particle II with its interior (d).

Table 6.4: Less ordered regions for particles I and II. The conventional '1000-4000' has been omitted from the residue numbering for clarity. * = visible at low contour levels. CT = C-terminus.

	I	II
VP1	1-66	1-68
	94-100	91-101
	142-143	143-149
		160-168
		217-226
		232-237
		251-252
		281-288
		293-295
VP2	1-14 ; 24-29	1-31
	44-56	42-61
		120-124
	164-166	163-169
		241-243
	270-271 (CT)	270-271 (CT)
VP3		16-18
	234-235	234-235 (CT)
VP4	All except 28-41*	All

antigenicity. This is consistent with views at the 5- and 3-fold symmetry axes, where we do not see a pronounced mesa, nor a flattening around the 3-fold (Figures 6.13, 6.14).

In particle I's map, the N-terminal portion of VP2, located in the capsid interior, is disordered up until residue 2014, consistent with plant-derived PV-3 SC8 (Marsian et al., 2017), and observations with FMDV VLPs (see Chapter 5, Section 5.2.1). VP2's N-terminus forms a hydrophobic hairpin that cements inter-pentamer contacts, which is clearly only visible in the map for I, although some patchy density can be observed at lower contour levels for II (Figure 6.13).

Also akin to plant-derived PV-3 SC8, for map I, at lower contour levels, additional density can be seen corresponding to residues 4026-4041 of VP4 that encircles the 5-fold axis (Figure 6.15).

Major antigenic sites are summarised in Figure 6.16. Site 1, containing VP1 BC loop residues 1094-1100, is totally disordered for both I and II Figure 6.14(c). The second major antigenic site, at the VP2 EF loop (residues 2163-2169), is resolvable at lower contour levels, and seems to be similar for both I and II (Figure 6.17).

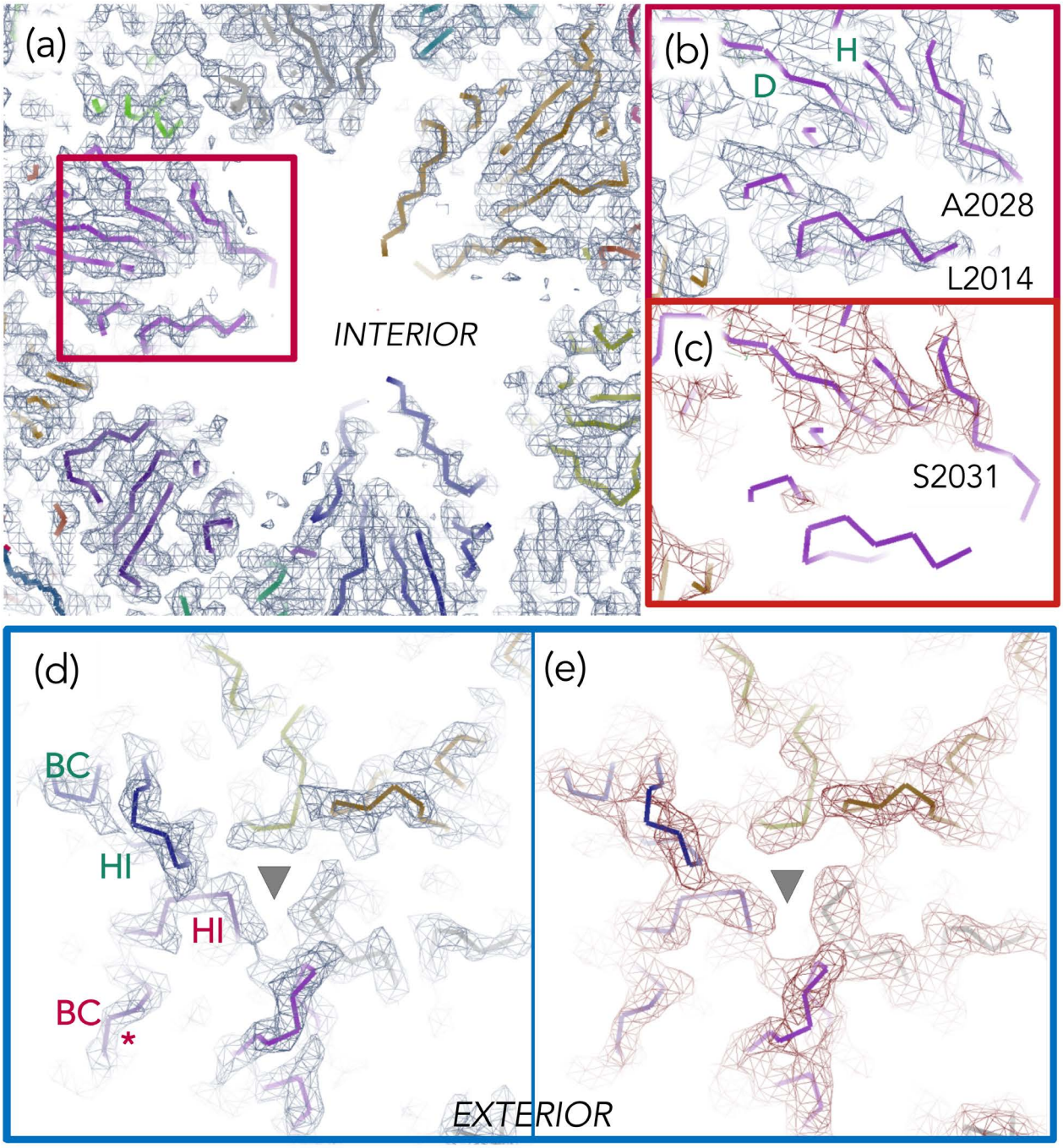


Figure 6.13: (a) View of the base of the 3-fold symmetry axis from the capsid interior with the map and model of I shown. (b) and (c) Zoomed in regions, corresponding to the red-boxed area in (a) showing the density supporting the N-terminal portion of VP2 for I and II, respectively. (d) and (e) view from the top of the 3-fold axis for I and II, respectively. The loop labels are coloured by VP; green VP2, magenta VP3. The '*' site corresponds to part of the third antigenic site.

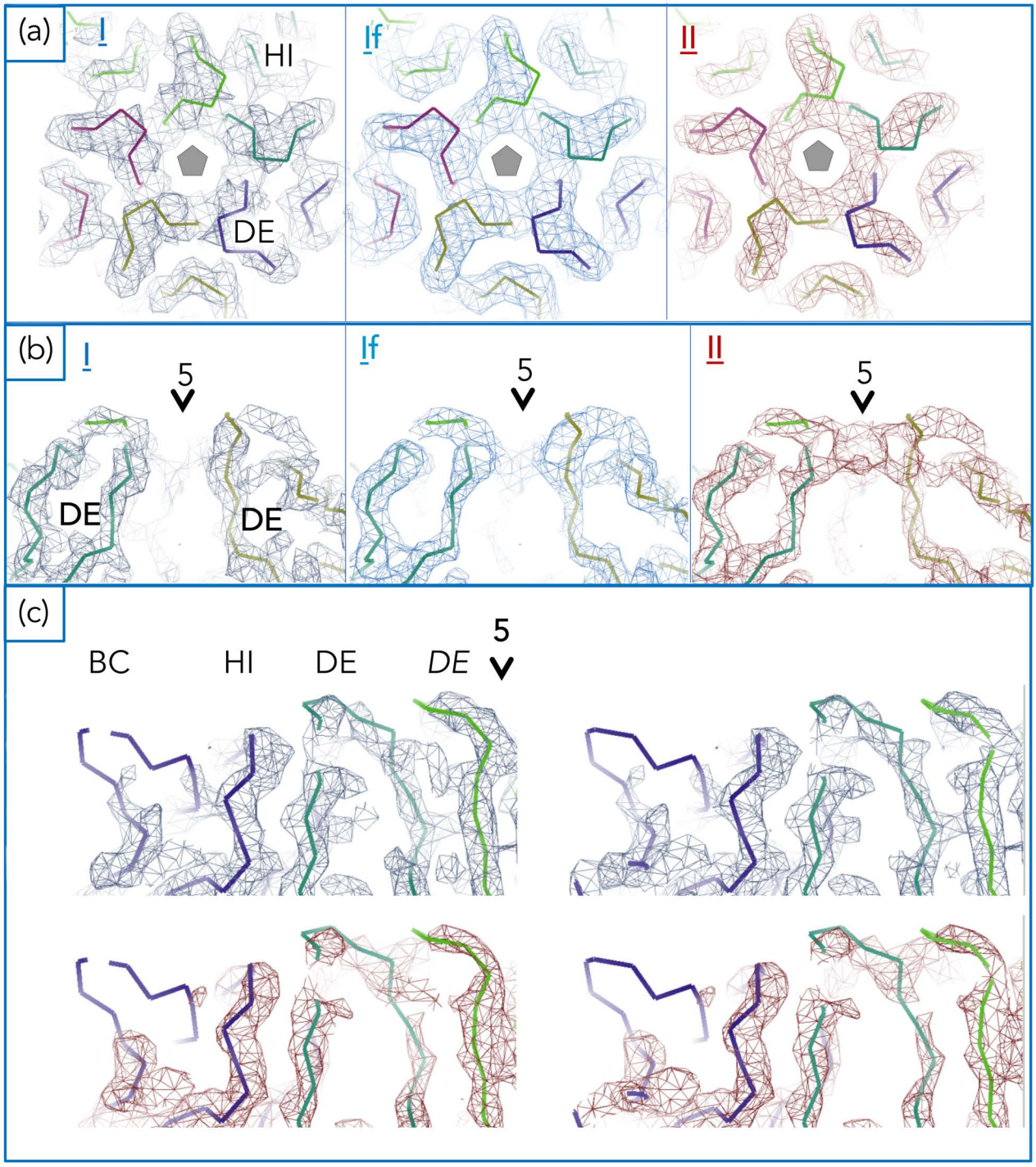


Figure 6.14: Comparison of I and II at the 5-fold axis. Blue-grey maps correspond to I, light blue maps to I-f and red maps correspond to II. (a) Top: Map for (I, If and II, as labelled) with HI and DE loops labelled. The model for I is shown in ribbon representation to aid comparison. The VP1 DE loops from five protomers sit around the 5-fold axis forming a star shape, with VP1 HI loops sitting in between the star tips. (b) Central slice through the 5-fold axis channel for I, If and II. For II, additional density can be seen in the centre, as indicated by the grey arrow. (c) Stereo images of a capsid slice showing an overview of VP1 loops that lead up to the 5-fold axis. The DE loop marked in italics is from the adjacent protomer.

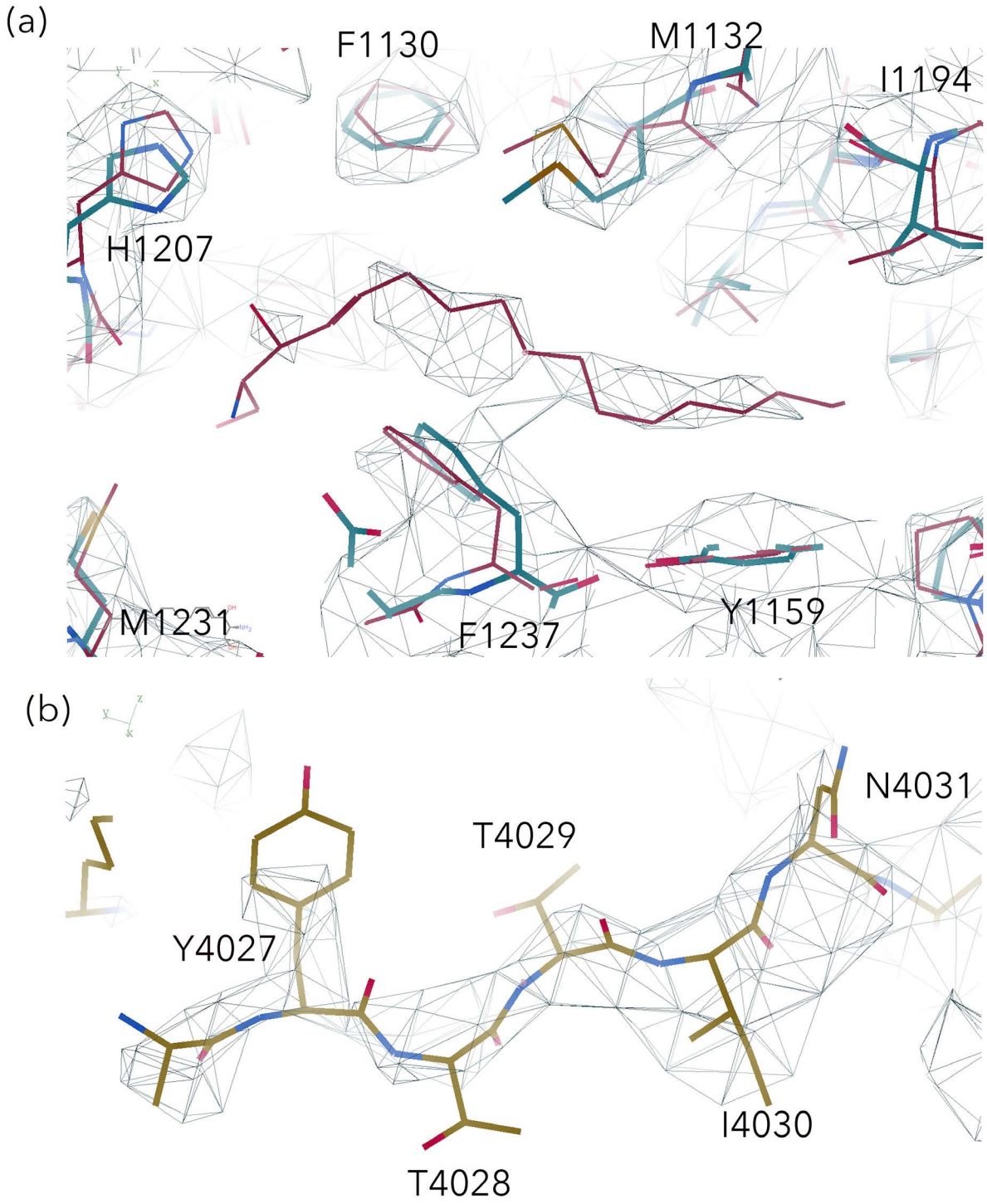


Figure 6.15: (a) Cryo-EM map around particle I's P-site, with PV-3 SC8 model. For comparison, PV-3 Sabin is shown in thin pink sticks (PDB: 1PVC). This was chosen in preference to the PV-3 Sabin model 1VBC for comparative purposes since a 'natural' pocket factor has been modelled in this region (Sayed et al., 1995). (b) Example density from particle I's VP4 site with model in gold stick representation.

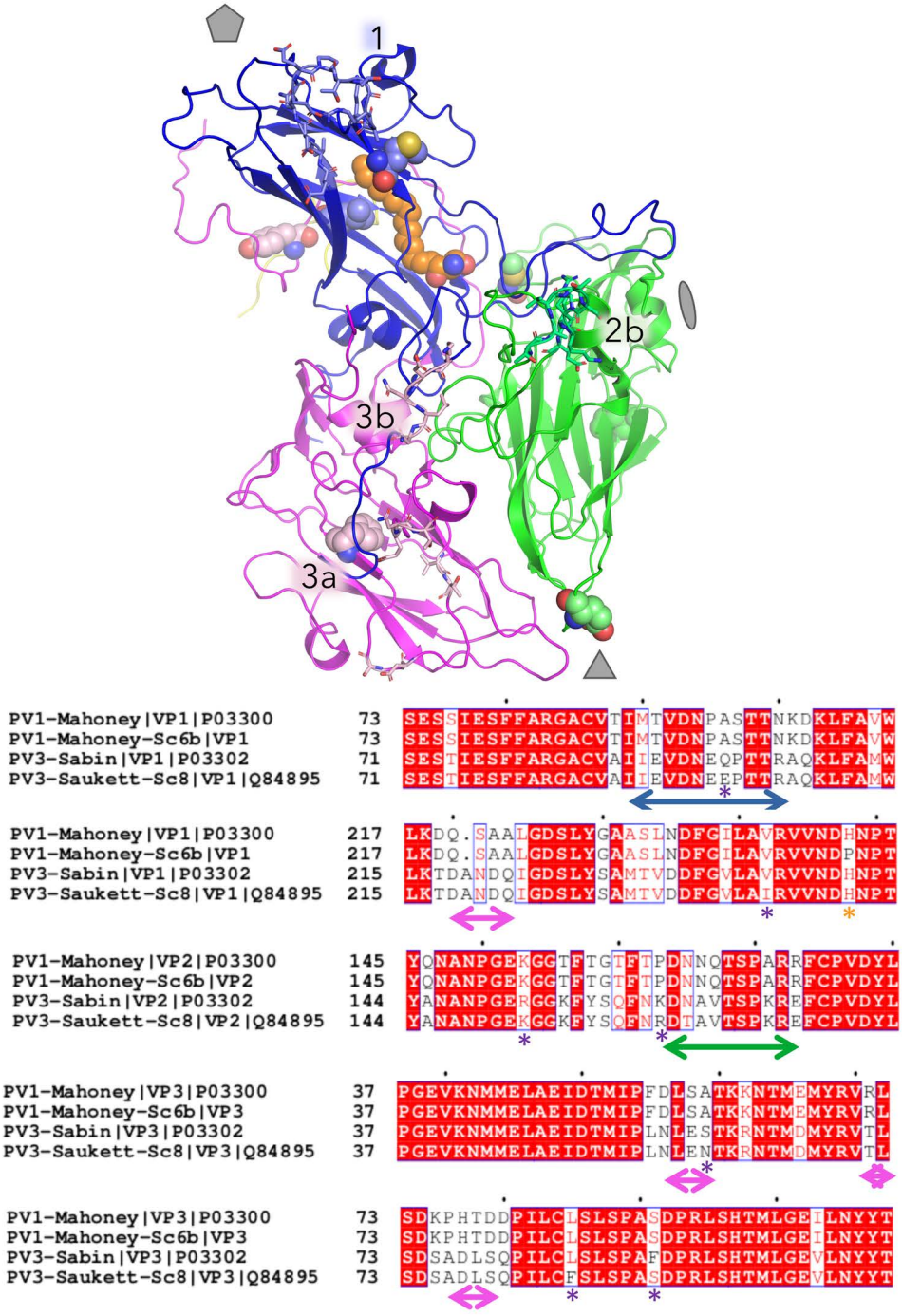


Figure 6.16: Top: PV-3 protomer with major antigenic sites (1-3b) (Minor et al., 1986) in stick representation, coloured by chain and labelled by conventional numbering 1-3b. Grey symbols represent the five-, two- and three-fold axes for orientation. The pocket factor in the P-site is shown as orange spheres, and mutations associated with PV-3 SC8 are shown as spheres coloured by their associated chain. Bottom: Sequence alignments around antigenic sites identified by Minor et al. (1986). VP and acquisition codes, where appropriate, are stated on each line. Note that PV-3 Saukett and Sabin are separate subtypes. Mutations between PV-1 types and between PV-3 types are shown as orange or purple asterisks, respectively. Antigenic sites are marked with coloured arrows (blue site 1; green site 2 and pink site 3). The full sequence alignment can be found in the Appendix.

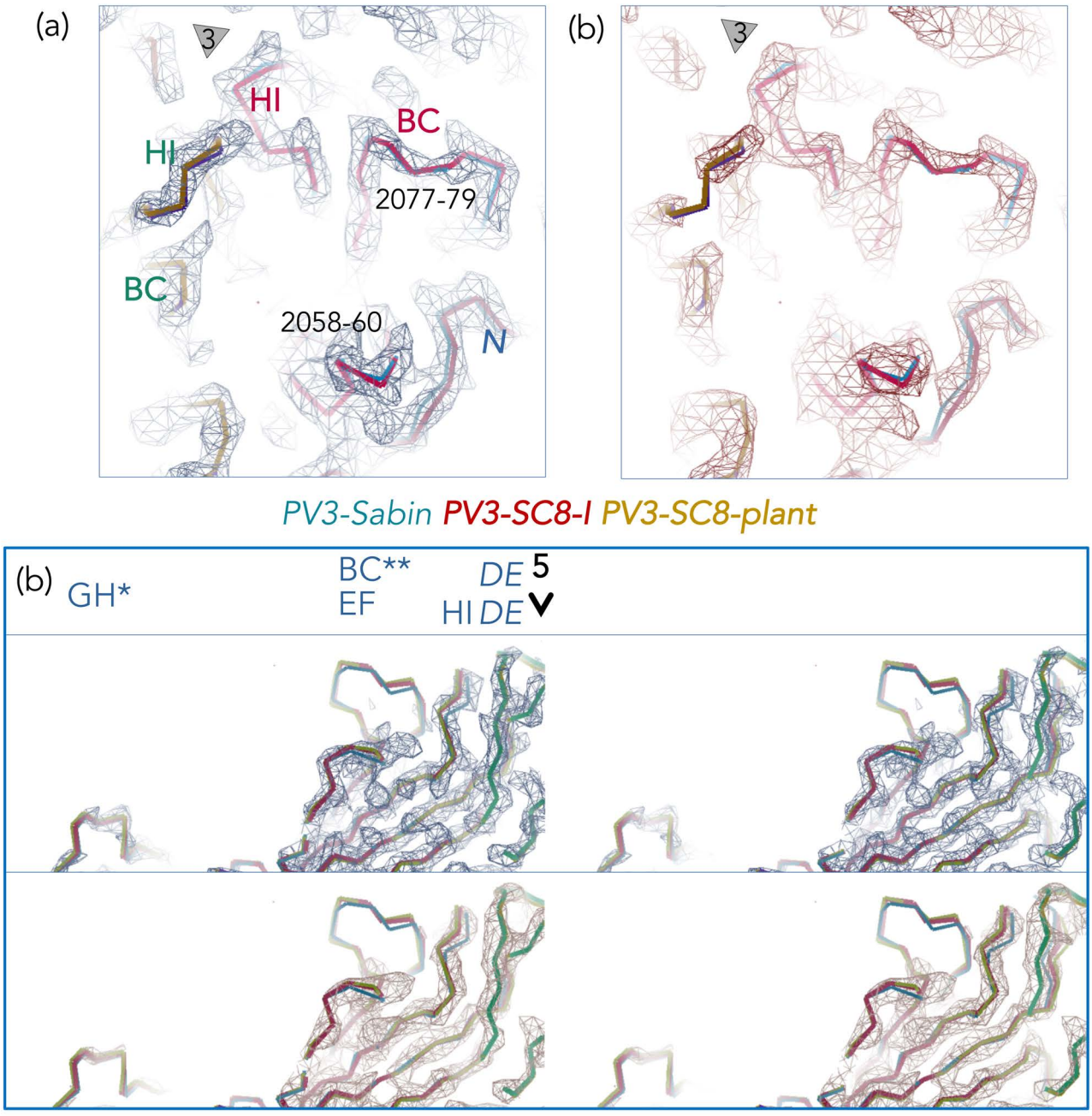


Figure 6.17: (a) Overview of antigenic site 3 for I (left) and II (right). This sits close to the three-fold axis. Loops are labelled and coloured according to protomer (VP3 pink, VP2 green). The N-terminus of VP1 is marked in italics. Models for PV-3 Sabin (PDB: 1HX6) are overlaid in blue and purple ribbon, but the fit is so close that this is barely distinguishable. (b) Wall-eye stereo of capsid slice showing an overview of the surface loops leading up to the 5-fold axis (five fold channel at the right most part of the image, marked with a '5' and arrow for the first stereoview). The loop labels are shown above the first stereo image and correspond to loops vertically below them. "**" denotes BC, where we see no density supporting the modelled loop. Models for PV-3 Sabin (blue) and PV-3 SC8-plant (gold) are also shown overlaid. PV-3 SC8-plant sits in an identical position, but PV-3 Sabin, shown with thinner ribbons, differs slightly at the EF and BC loops.

Particle I versus Existing PV-3 structures

Maps derived from I clearly support the majority of mutations anticipated for PV-3 SC8, shown in Table 6.1 above, relative to that of Sabin strains, Figure 6.18. In addition, as anticipated, particle I shares strong similarities to both the plant-derived PV-3 SC8 VLP model (PDB: 5O5B) and, to a lesser extent, that of native PV-3 Sabin (PDB: 1VBC), RMS between C_α positions of PV-3-I versus PV-3 SC8 Saukett plant and native PV-3 Sabin SHP of 0.85, 0.40 for 728 and 719 equivalences, respectively (Stuart et al., 1979; Grant et al., 1994; Marsian et al., 2017).

In contrast to what was seen for the plant-derived PV-3 SC8, elongated density can be seen at the P-site, sufficient to support at least a myristic acid molecule (Figure 6.19). Fatty acids (FAs) play a major role in yeast cell metabolism. For instance, major FAs for *Saccharomyces cerevisiae*, a member of the same family as *Pichia Pastoris*, includes palmitic acid (C16:0) and myristic acid (C14:0) (Tuller et al., 1999; Viljoen et al., 1986). *N. benthamiana*, from which plant-PV-3 SC8 was derived, contrastingly, has been found to mostly contain unsaturated FAs such as linolenate (50-63 % FA content, C18:3), palmitoleate (13-18 % C16:1), linoleate (5-18 %, C18:2) and hexadecatrienoic acid (2.6-7.7 %, C16:3) (Koiwai et al., 1983). Unsaturated fatty acids have not been reported to bind to the pocket factor site in enteroviruses, a likely consequence of inflexibility and deviation from linearity imparted by main chain double C=C bonds making them less able to enter the P-site and forge favourable contacts with the P-site residues. Thus, it may be that yeast expression systems offer a better provision of stabilising P-site binders.

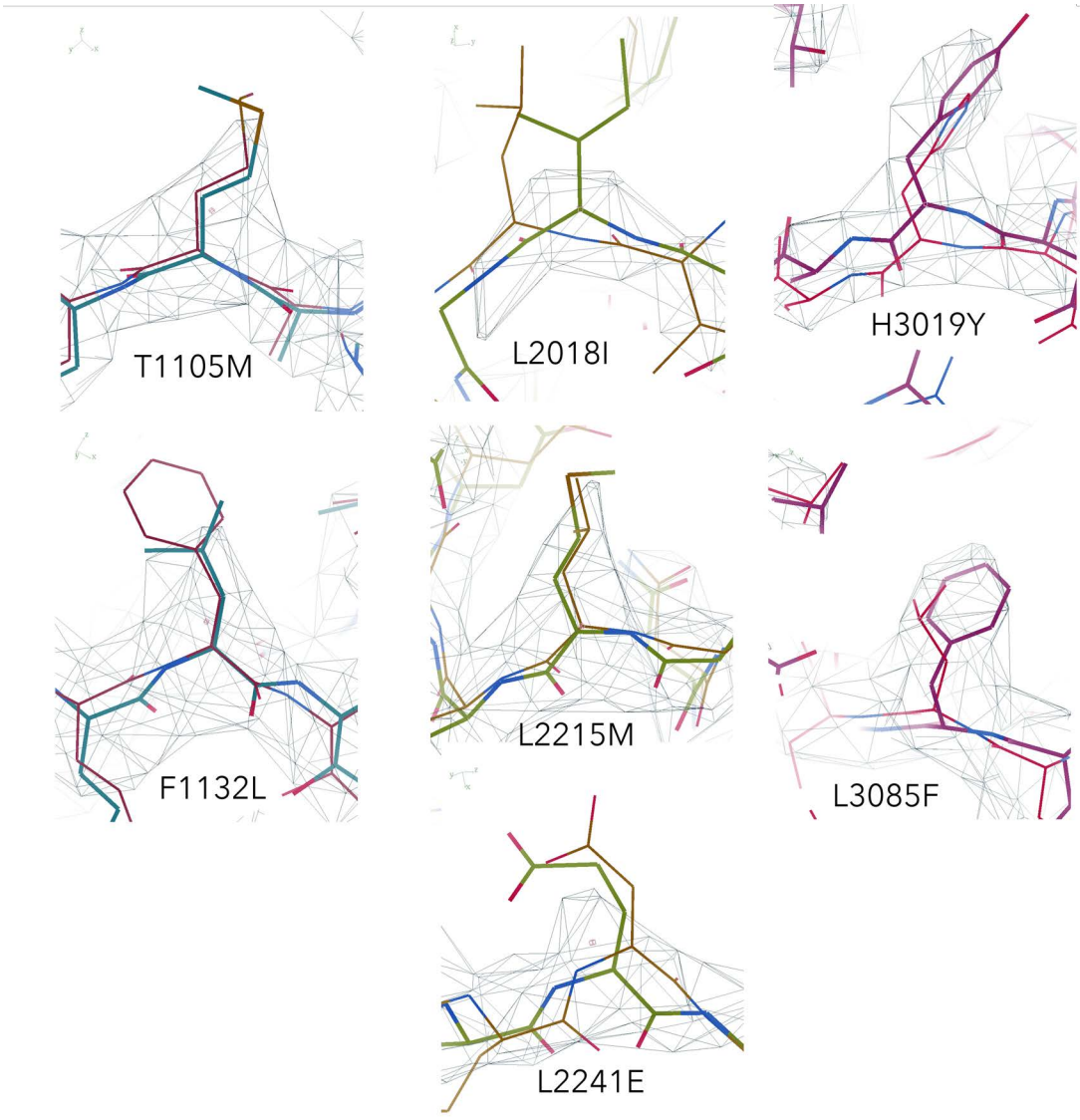


Figure 6.18: Fit of cryo-EM map with the anticipated thermostabilising mutations introduced into Saukett as described in (Fox et al., 2017) shown as thick sticks, coloured by chain. For comparative purposes, PV-3 Sabin is shown in thin lines (PDB: 1VBC, thin lines) (Grant et al., 1994). A final mutation in VP4 is excluded, since there was no visible density.

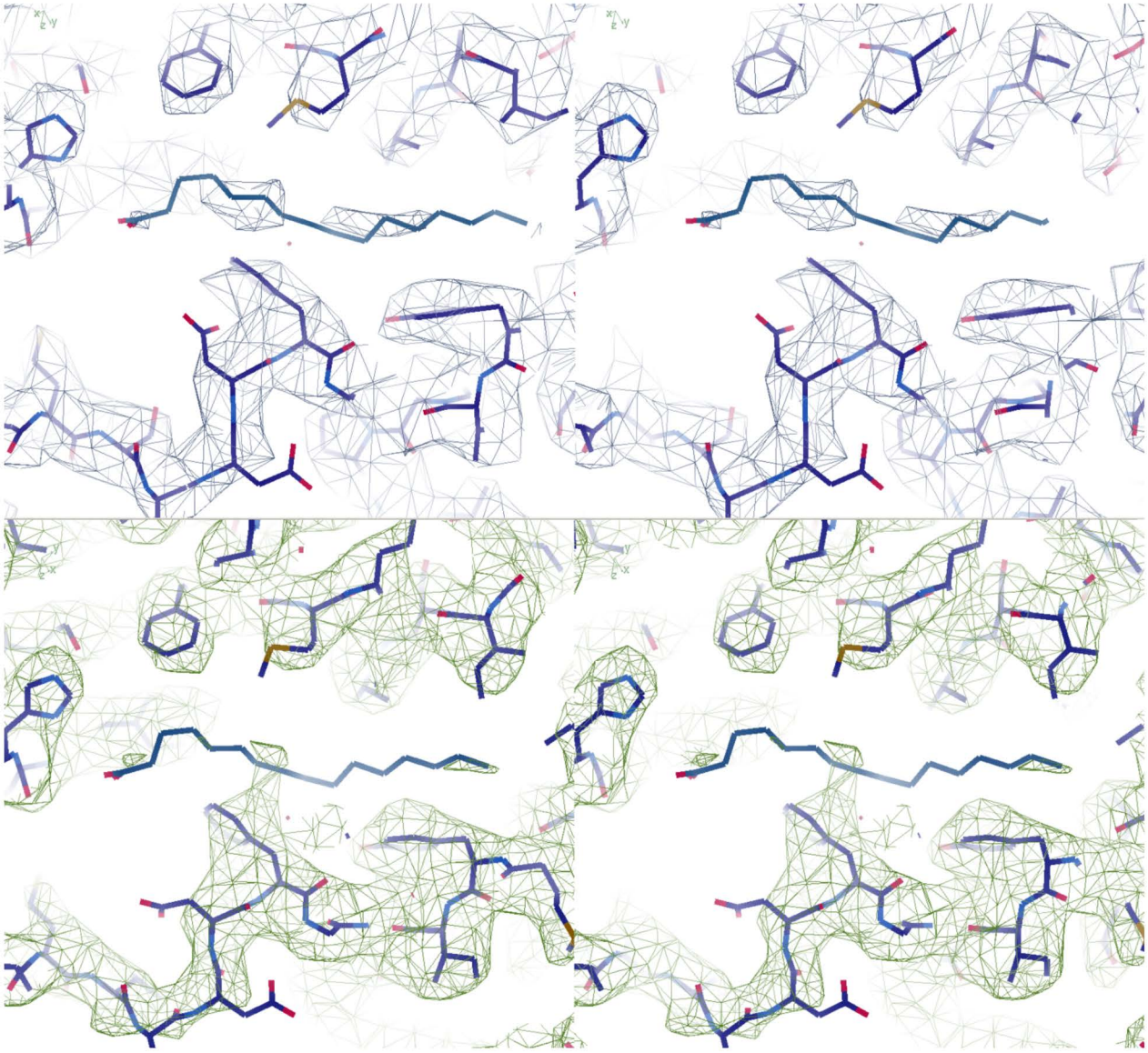


Figure 6.19: Wall-eyed stereo views showing palmitic acid modelled into the P-site for I (blue map, centre) and plant derived PV-3 (PDB: 5O5B green map, bottom) (Marsian et al., 2017).

6.4 Discussion

6.4.1 Rationalising the PV-1 SC6b:PV-3 SC8 C/D antigenic divide

The results presented above suggest that the preparation of yeast expressed PV3-Saukett SC8 VLPs was essentially of completely native (D) antigenic state. The close similarity between the two classes, I and II, may be due to one class representing a subpopulation of particles that have undergone the VP0 maturation cleavage (something that has been observed for some picornavirus VLPs) (Porta et al., 2013a). However, no bands consistent with VP2 and VP4 molecular weights have been observed on SDS-page reducing gels of PV VLP preparations (personal communication, Mohammad Bahar) and thus this latter suggestion seems unlikely.

Contrastingly, for PV-1 Mahoney SC6b VLPs expressed using the same system, purified in a similar way, and picked using a deep-learning algorithm (Tegunov and Cramer, 2018) (this algorithm had initially been trained on FMDV VLPs before retraining on all, round particles seen in micrographs, before being manually filtered to remove non-PV particles), the majority, 70 %, were of C-antigenicity. Both samples were shipped together from Leeds, and PV-1 concentration and grid preparation were performed a week before that of PV-3. The reported C/D antigen content ranged from 390-425 D-antigen units (D-Ag) for three 2 mL fractions of PV-1 and 240-380 D Ag for three 2 mL fractions pooled for PV-3 SC8. Note that D-antigen units cannot be compared across serotypes, and are best used to simply establish the presence of a significant population of D-antigen units.

It would be very attractive to be able to use microscopy as a direct readout of the balance of antigenic states, since C particles are useless as a vaccine component, and present methods of characterisation (ELISA based) rely on having type and state-specific antibodies (Hazendonk et al., 1981). However, we should be aware, as discussed in Chapter 2 Section 2.5.2, that only perhaps one in a thousand particles added to a grid are used in structural analysis, the remainder being lost during the blotting process. The question is therefore whether there is a selection of one particle type over another, thereby invalidating EM as a method for state-quantification.

For the PV VLPs presented here, it was obvious that many particles had an affinity for the carbon support. If the C and D particles show different affinity for carbon, then, for the ultra-thin support used for the PV3 samples, particles would be uniformly distributed, whereas, for C-flat grids, the

particles with higher affinity for carbon would accumulate between grid holes, and thus not be picked for processing (the deep learning algorithm and filtering specifically exclude these since the contrast is much worse). In order to explain the observed predominance of C particles picked for PV-1, D antigenic particles would have a higher carbon affinity than C-antigenic types. This is conceivable, however, the lack of C particles picked for PV-3, where a continuous carbon support was used, and thus no differential carbon-affinity particle selection, would suggest that that there are more C particles in PV-1.

Recently, PV-1, -2 and -3 VLPs with comparable thermo-stabilising mutations (Fox et al., 2017) expressed in baculo infected cells independent of viral proteinase 3CD have been reported, where, rather than a P1 polyprotein, constructs contain VP1, VP3-2A-VP0. This circumvents cytotoxicity issues that have been encountered with 3C proteases, and authors report significant D-antigen content for all three serotypes, even after 24 days (Xu et al., 2019) and this may be a good alternative strategy to try.

Concentration Strategy

Different procedures were employed to concentrate the PV-1 SC6b and PV-3 SC8; a 100 kDa MWCO Amicon filter device for the former and pelleting for the latter. Pelleting was employed in an effort to improve particle densities. It is conceivable, therefore, that D-antigenic particles have a higher affinity for the concentrator membrane, giving rise to a selection bias for C-antigenicity. Alternatively, components on the cellulose membrane itself could be disruptive. Filters reportedly contain trace amounts of glycerol, and, according to manufacturer specifications, a harsh wash with NaOH is sometimes required to fully remove glycerol traces, but it is unlikely that small amounts of glycerol would confer a destabilising effect on picornavirus or picorna-like particles.

6.4.2 Assessing the Significance of Internal Capsid Density

Aside from some apparent density in the capsid centre in central volume slices, there is no evidence for significant packaged RNA. Since yeast is not a natural host and so packaging of yeast RNA will not have been selected against, it is conceivable that the capsids package some host RNA, but this seems extremely unlikely, especially given that the 260/280 readings suggest negligible RNA to be present.

6.4.3 Grid preparation

While lacey grids offer a means of accumulating a sensible particle distribution, inconsistency in lace distribution, and resolution-impeding carbon can seriously limit high resolution data collection in terms of speed of set up, and attainable resolution. For future experiments, particle concentration and density need to be improved, and simple methods such as glow discharging the more robust Quantifoil grids for a shorter period of time, or with different gases, could be investigated.

Chapter 7

Developing tools to probe Aphthovirus Uncoating

7.1 Introduction

FMDV remains one of the world's most impactful livestock diseases for trade, with some outbreaks costing billions of pounds. The costly necessity to quarantine and sometimes eradicate infected animals has had ruinous implications on farming communities, Chapters 1 and 5 Sections 1.2 and 5.1(Knight-Jones and Rushton, 2013; Sinkala et al., 2014). While vaccination serves to help quell the spread of disease (Chapter 5), anti-viral preparations to treat infected animals would help to reduce costly control measures. A greater understanding of the underlying mechanisms of FMDV infection, particularly at the elusive uncoating step, could both inform rational design of anti-virals, and inspire methods to extend the shelf-life of capsid-based vaccine preparations.

As described in Chapter 1 Section 1.1.4, at the first stage of infection, viruses must bind to susceptible cell receptors and, for FMDV, integrin receptors have been identified ($\alpha_v\beta_{1,3,6,8}$ in particular) (Neff et al., 1998; Jackson et al., 2000a,b; Berryman et al., 2005; Burman et al., 2006; O'donnell et al., 2009; Kotecha et al., 2017). The real mystery for aphthoviruses is whether an analogous A-particle-like intermediate exists, and, more importantly, how its precious RNA is translocated across the endosomal membrane into the cytoplasm. It has been shown that FMDV uncoating is not triggered by receptor binding (Mason et al., 1993; Kotecha et al., 2017), and thus it is possible to decouple receptor attachment and priming for genomic release. One way to mimic this integrin

receptor binding event with a cell-free system for cryo-electron tomographic study, is to embed integrin into either nanodiscs or liposomes, which serve to mimic the phospholipid membrane (Bubeck et al., 2005a; Tuthill et al., 2006; Strauss et al., 2013; Lee et al., 2016; Groppelli et al., 2017; Nasr et al., 2017). This cell-free system could then be subjected to different pH conditions in a controlled manner to see if we can capture intermediate FMDV states. To this end, attempts to express full length $\alpha_v\beta_6$ are described.

Time-resolution in cryo-EM studies has typically been achieved using rapid premixing, and spraying techniques that require specialised apparatus (Chen et al., 2015; Frank, 2017). Alternatively, 'caging' of probe molecules, such as with photo-activatable groups, provides a means of on-demand release of probe additives without the need of specialist equipment (see Chapter 2). Activation of a caged proton mixed with FMDV shortly before vitrification may enable one to capture transient uncoating states in a controlled manner. In the case of our studies, we present very provisional investigations into the use of a nitrophenylethyl (NPE) protecting group, which, upon exposure to 280 nm light, is deprotected, via a 2-nitrobenzyl rearrangement, to a nitrosoketone and proton (see right panel, Figure 7.1). This facilitates a straightforward means of on-demand pH drop prior to vitrification of virus-immobilised grids. Figure 7.1 illustrates how we propose to implement this.

The aphthovirus equine rhinitis A virus (ERAV) is genetically and structurally similar to FMDV (Li et al., 1996). ERAV capsids are slightly less featureless than FMDV, however; with an appreciably different VP1 and more extended surface loops, resulting in a slightly more pitted topology akin to cardioviruses (Tuthill et al., 2009). ERAV and FMDV also share similar physicochemical properties, including comparable buoyant density and acid lability, albeit ERAV is somewhat more robust to acidic pHs (Tuthill et al., 2009; Newman et al., 1973).

Such similarities have made ERAV a useful surrogate for understanding FMDV biology, despite recent claims that ERAV infections may be more severe than initially believed; impacting substantially on equine breeding and training (Bażanów et al., 2018). ERAV is also not the perfect virus for structural studies since it does not grow well in tissue culture; yields from purifications are typically low, and samples are unstable in sucrose-free buffers (Tuthill et al., 2009).

Herein efforts to lay the foundations for, and begin to probe, aphthovirus intermediate structures are presented:

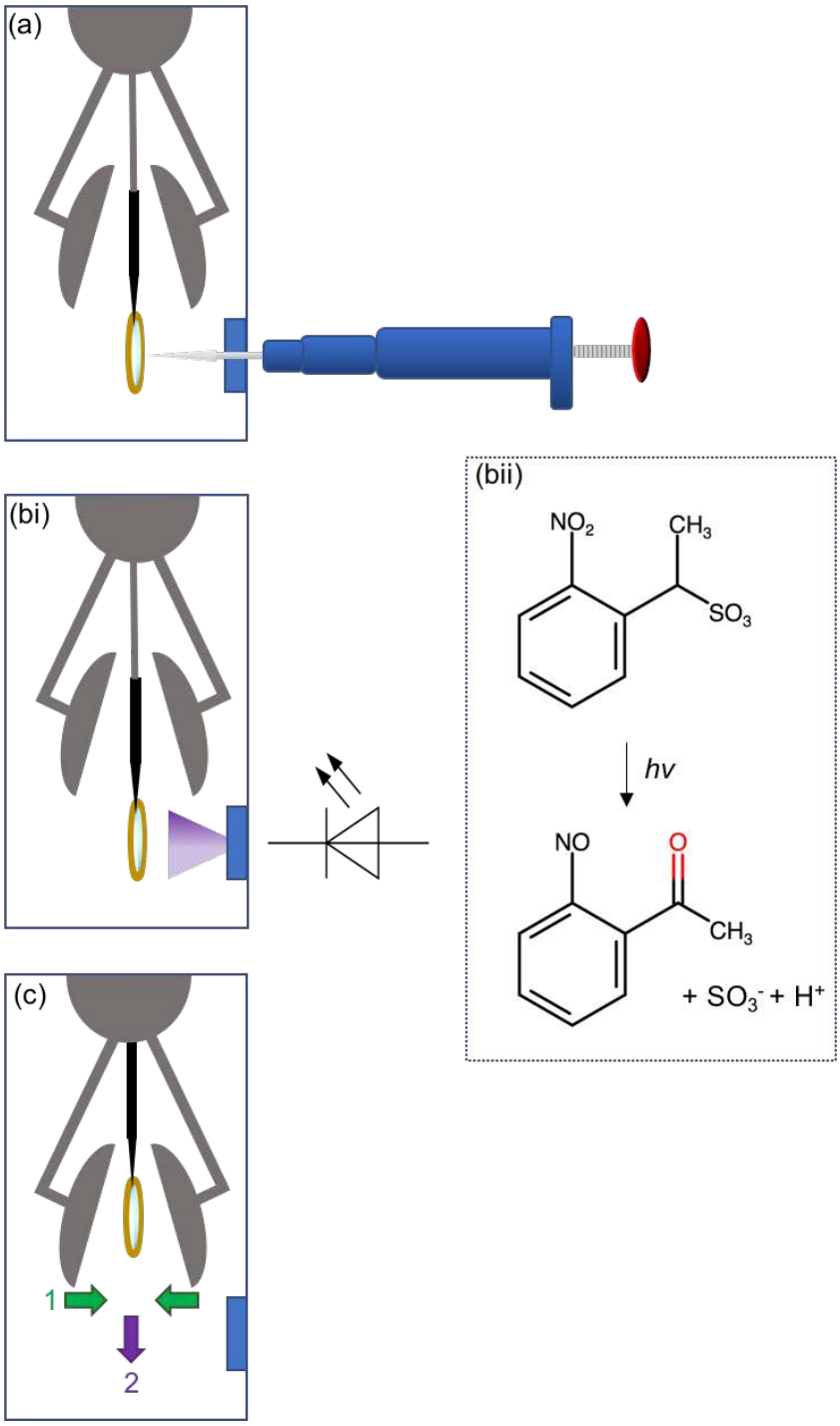


Figure 7.1: Proposed method for pH drop experiments within the vitrobot. (a) A pre-mixed sample containing virus, HEPES buffer, and inactivated caged proton is pipetted onto an EM grid. (bi) The EM grid is subjected to a controlled flash of LED light. This causes photolytic degradation of the NPE cage, leading to proton release, as in (bii). (c) The grid is rapidly 1-blotted, then 2-vitrified.

1. Attempts to over-express full-length α_v β_6 integrin receptor and a truncated α_v β_3 for further investigations into the FMDV-receptor binding event.

2. Development of strategies to capture uncoating intermediate structures:

   (i) Capture through caged proton release.

   (ii) Capture through stabilisation with additives.

3. Study of a low resolution ERAV empty particle.

7.2 Results

7.2.1 Cloning and Expression of $\alpha_v\beta_6$ Integrin Receptor

In an effort to create an *in vitro* model of FMDV receptor binding and uncoating, attempts were made to express full length $\alpha_v\beta_6$ i.e. with the ecto, transmembrane and cytoplasmic domains (Chapters 1, 2; Sections 1.1.4, 2.1.2-2.1.5). To achieve this, I spent some time in the lab of Nicola Abrescia at the CIC bioGUNE, Bilbao.

Constructs of *Bos taurus* α_v and β_6 integrin containing ecto, transmembrane and cytoplasmic domains, cloned into the pHLSec vector containing a C-terminal His_6 tag and N-terminal signal peptide (Aricescu et al., 2006) and either a monovenus (MV) or mRuby2 fluorescent tag (see Chapter 2 Section 2.1.2 for full description) were checked by sequencing, before being used in a triage of transient transfections in mammalian HEK293T cells. Transfection trials were confined to mammalian cell lines to ensure that post-translational modifications, such as glycosylation, that are crucial for integrin receptors (Gu and Taniguchi, 2004; Bellis, 2004; Berman et al., 2003), could occur. Notably, whilst α_v cloning was straightforward, β_6 integrin cloning often failed, even with simple 'cut and paste' ligation-dependent cloning, regardless of ligase and ligation time; requiring repeated attempts before success on each occasion.

For the majority of transfections, β_6 integrin transfected cells showed obvious fluorescence signal one day post-transfection (pt) when transfected in isolation, when co-transfected with α_v constructs and when sequentially transfected (i.e. α_v on one day, β_6 and vice versa).

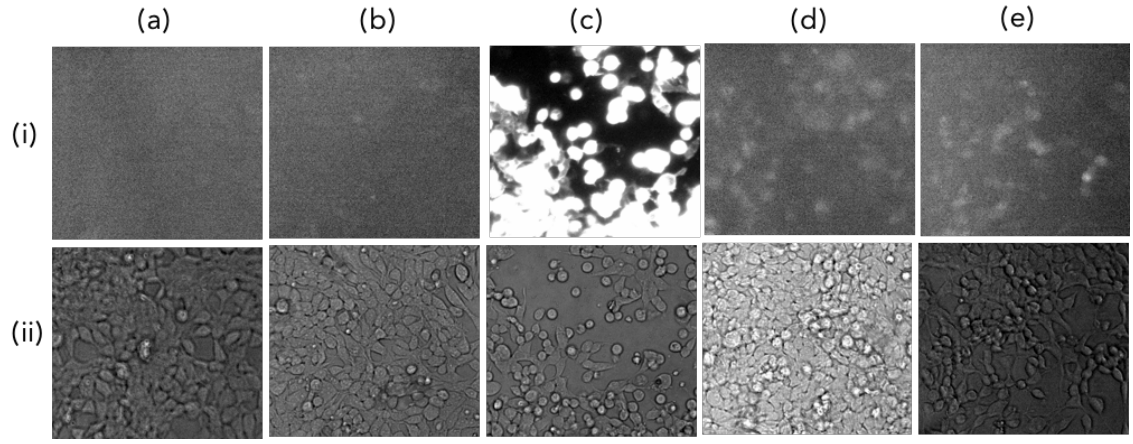


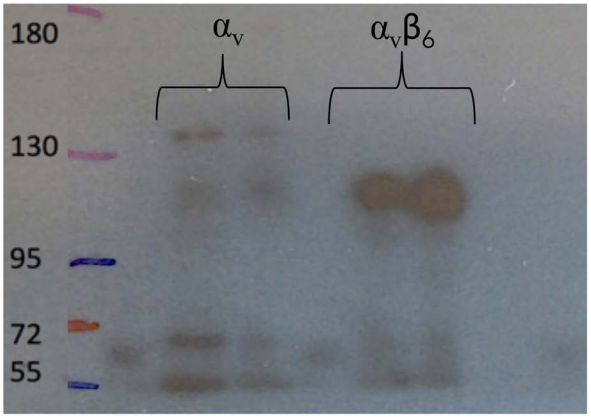
Figure 7.2: Example images from α_v transfection of HEK293T cells with 1 μg DNA and X-tremeGENE (Roche) transfection reagent. Identical contrast and exposure times were used for the fluorescence images. (i) Corresponds to fluorescent signal; (ii) to bright-field images. (a) Negative control, untransfected cells. (b) Negative control, cells transfected with X-tremeGENE (Roche) transfection reagent only. (c) Positive control; construct with mRuby2 fluorescent tag only. (d) α_v pHLSec with mRuby2 tag at a 1:4 ratio with X-tremeGENE transfection reagent. (e) α_v pHLSec with mRuby2 tag 1:2 ratio X-tremeGENE transfection reagent.

Little, if any fluorescence, however, was observed for α_v transfections both with and without co-transfection with β_6. For some HEK293T transfection trials, very faint signal from α_v's fluorescent tag was observed two to three days post-transfection in small scale cell culture, Figure 7.2. This was the case for both the monovenus and mRuby2 tags. western blots (WBs) detecting His tags indicated that a species of similar molecular weight to α_v had been expressed, confirming that the fluorescence observed was not from auto-fluorescence of dying cells (see Figure 7.3(a)). These successes only occurred during my time in Bilbao when X-tremeGENE (Roche), or PEI, transfection agents were used with low-passage cells (pass number below 10), but such results could never be replicated in Oxford.

These difficulties were not entirely unanticipated; some work had been done previously to express integrin in Oxford (Nicola Abrescia, personal communication), to no avail, and, at the start of this project, structural studies using recombinant integrins only used 'truncated' constructs i.e. lacking the cytoplasmic, transmembrane and sometimes a portion of the ectodomains, such as (Xie et al., 2010; Dong et al., 2012; Kotecha et al., 2017). Note that, on this latter point, the main reason for construct truncation was to overcome the inherent flexibility of integrins.

To further confirm the successful expression of the β_6 construct, small scale transfections (175

(a) Anti-penta-His



(b) 8 % SDS-Page

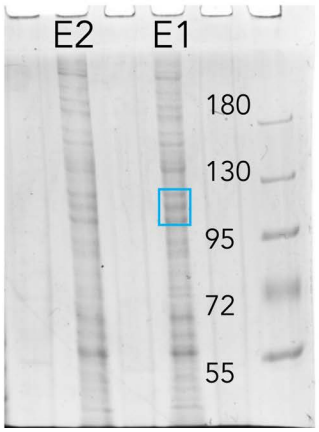


Figure 7.3: (a) Anti-penta-His western blot of an unusually successful mono-α_v transfection, and (right) $\alpha_v\beta_6$ co-transfection. Note that this was exposed following an incubation with the developing film for 2 mins and thus, although bands appear strong, the signal was actually quite weak. (b) 8 % reducing SDS-page gel from $\alpha_v\beta_6$ nickel purification showing eluted fractions 1 and 2. Marker molecular weights (kDa) are indicated on the right hand side. The blue box shows that gel portion excised for mass spectrometric analysis.

mL flasks) of $\alpha_v\beta_6$ were purified using batch nickel-affinity purification (see Chapter 2 Section 1.1.3). Weak bands for the first two eluted fractions (E1, E2) were consistent with the molecular weight of β_6 integrin (117 kDa inclusive of 3C protease cleavage site, MV and His_6 tags). Even after 10-fold concentration of successful fractions with an Amicon Ultra-Centrifugal Filter (Merck, 10 K cut off), likely integrin bands were barely discernible in a reducing SDS page gel (Figure 7.3(b)). Nevertheless, nLC MS/MS (fingerprint mass spectrometry) analysis conducted by the bioGUNE proteomics facility (Bilbao, Spain), confirmed that the region suspected to correspond to $\alpha_v\beta_6$ integrin (light blue box, Figure 7.3(b)) did indeed contain bovine β_6 integrin, but no α_v integrin was detected, nor for the weak band above 130 kDa.

Trouble-shooting Transfection

The major strategies used to improve transient transfection successes are summarised in Table 7.1, and described in the main text below.

Table 7.1: Summary of transfection trials for transient expression of integrin constructs. Note that LF2K = Lipofectamine2000 and XtG refers to X-tremeGENE (Roche).

Condition	Transfection Reagent	Cell line(s)	Success?
Transfection Reagent, 2:1, 3:1, and 4:1 transfection reagent:DNA ratios.	CaPhos	HEK293T	No.
	LF2K	All.	β_6 only.
	PEI	HEK293T	β_6, once or twice at 30 % success for α_v 3 days pt.
	XtG	HEK293T, BHK21	Limited for both.
Cell line	PEI, LF2K, XtG	BHK21	>60 % for β_6
	LF2K	CHO	90 % for β_6
	LF2K	COS7	>60 % for β_6
	PEI,LF2K	HEK293S	90 % for β_6
	PEI,LF2K,XtG, CaPhos	HEK293T	90 % for β_6
	LF2K	HUH7	>60 % for β_6
	LF2K	L	80 % for β_6
	LF2K	WI-38	80 % for β_6
Sequential cotransfection, β_6 first.	PEI, XtG	HEK293T	Some expression for β_6.
Sequential cotransfection, α_v first.	PEI, XtG	HEK293T	Some expression for β_6.
Co-transfection ratios, 5:1,3:1,2:1 Some expression for α_v:β_6, and vice versa.	XtG, PEI	HEK293T	No consistent changes in expression levels.
Incubation at 33 °C	PEI	HEK293T	Low levels of expression.
Pre-transfection incubation at 33 °C	PEI, LF2K, XtG	HEK293T	Similar levels to cells pre-incubated at 37 °C, but to be repeated.
Expression vector: pOPINTTG	PEI	HEK293T	None.
Absent fluorescent tags: α_v-His_6 β_6-His_6	PEI,LF2K,XtG, CaPhos	HEK293T, BHK21	Some anti-His WB signal for β_6 only.
Construct length. α_v, β_6 with ectodomain-transmembrane domain only	PEI	HEK293T, BHK21	None.
Lamp-1 construct	PEI, LF2K	HEK293T	None.

Cell line

Glycosylation is essential for integrin functionality and it seemed possible that the particular glycosylation signature acquired by the α_v construct may hamper expression (Croset et al., 2012), or the cell line used does not permit α_v post-translational cleavage necessary for maturation (Hynes, 1987; Paule et al., 2012). This seemed unlikely, given that β_1 integrin glycoslation is cell phenotype dependent (Bellis, 2004), while, according to the 'Cell Surface Protein Atlas' database (http://www.proteinatlas.org/ and http://wlab.ethz.ch/cspa/, (Bausch-Fluck et al., 2015)), by contrast to the somewhat promiscuous α_v, β_6 integrin, which is not natively expressed in HEK293T/S cells. Previous attempts to express $\alpha_v\beta_6$ head group utilised HEK293S, BHK21 and CHO-K1 cell lines, and thus these were selected for further transfection trials. The more esoteric fibroblast cell lines: 'L' and WI-38 were also tested, Figures 7.4 and 7.5.

For each cell line studied, successful transfection was only detected for β_6 integrin in both mono- and co-transfections with α_v. This was somewhat surprising given that human α_v, with over 90 % sequence similarity to bovine α_v, is endogenously expressed in the cell lines used. Therefore, the successful β_6 integrin expression across all cell lines studied is presumably a consequence of co-association with endogenously expressed α_v integrin. Thus it could be hypothesised that the presence of native α_v integrin may somehow suppress the expression of our constructs, perhaps by saturating the available endogenously expressed enzymes that mediate folding or post translational processing of integrins, such as Pro-protein Convertase 6 (PC6) that post translationally cleaves integrin into functional subunits (Paule et al., 2012). To test this, I attempted to generate a chimera $\beta_6$ectodomain-LAMP1 construct. This was theorised as having the potential to dimerise with native α_v, thereby earmarking it for lysosomal degradation (Conesa et al., 2003). Cells could be transfected with this chimera initially, as before, and then washed and transfected the following day with bovine $\alpha_v\beta_6$ integrin constructs. However, this failed at the initial transfection stage; no $\beta_6$ectodomain-LAMP1 construct could be detected by fluorescence nor anti-His WB, and this strategy was abandoned.

Transfection Conditions

Some additional factor may need to be up-regulated to facilitate $\alpha_v\beta_6$ production. *In vivo*, $\alpha_v\beta_6$ integrin is known to be up-regulated during wound healing, embryogenesis, and in some tumours

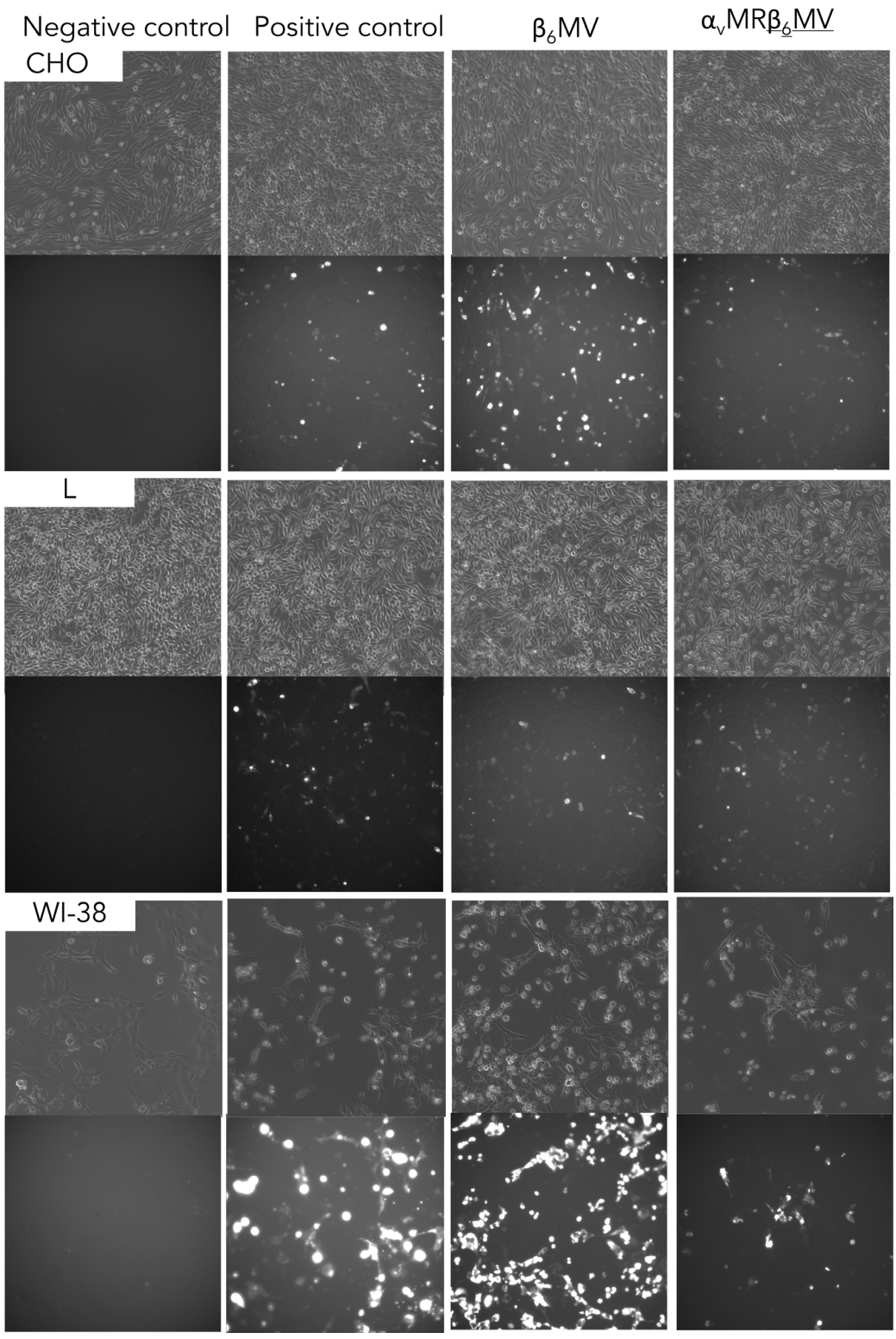


Figure 7.4: Example fluorescence results from transient transfection of difference cell lines. For each image pair, the brightfield light microscopy image is at the top, and the fluorescence image at the bottom. The negative control was un-transfected cells incubated in the same, 2 % DMEM with supplements used for the other conditions (see Chapter 2, Section 2.1.3). Images from the mRuby2 channel were omitted for the co-transfection column (far right) due to the images being completely blank due to absent signal.

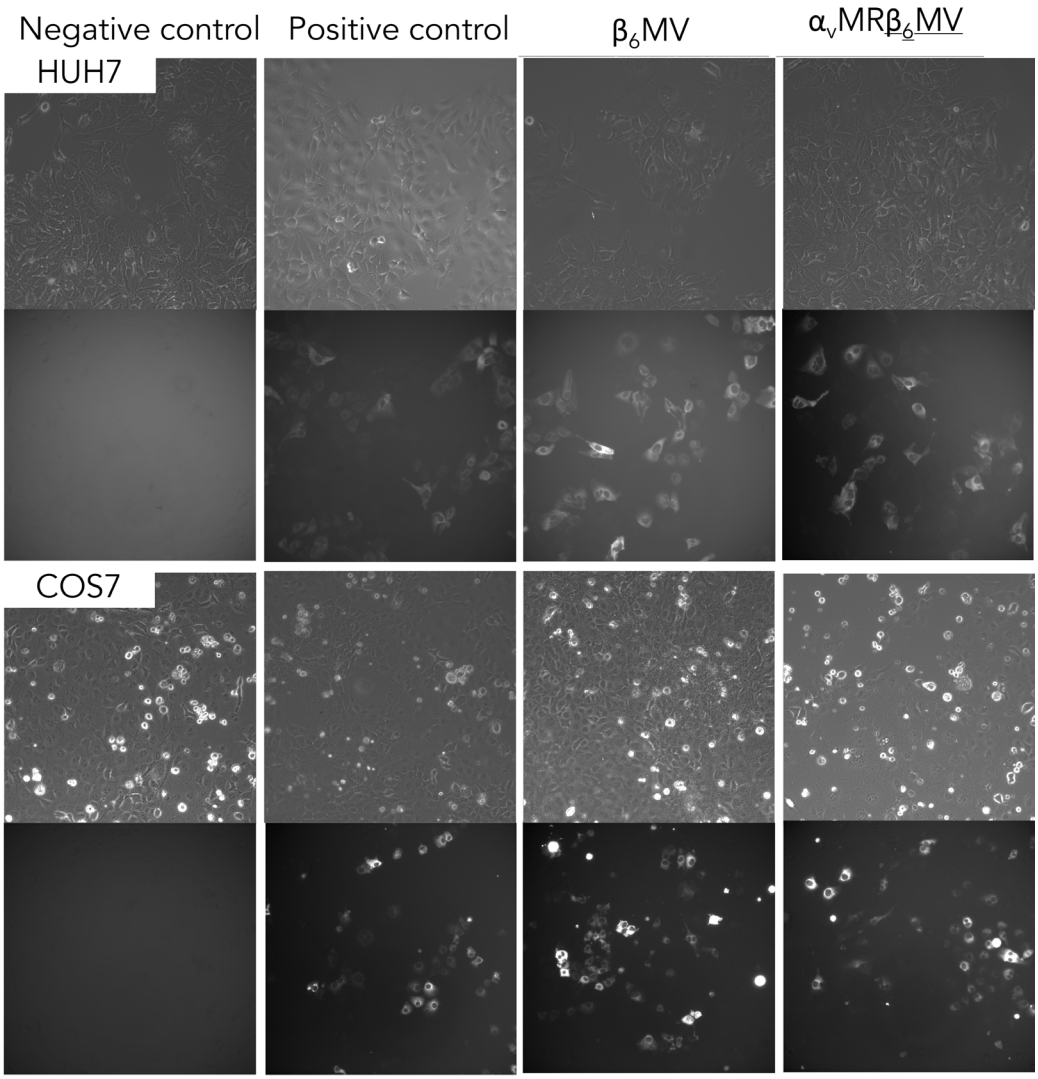


Figure 7.5: Example transfection fluorescence results for HUH and COS7 cell lines, as above. For each image pair, an image of cells under brightfield conditions is at the top and the corresponding fluorescence image for the MV channel at the bottom.

(Bandyopadhyay and Raghavan, 2009). Identification of substrates or conditions that unite these processes may help to induce over-expression of $\alpha_v\beta_6$ constructs.

For instance, major mediators in the inflammation pathway include interleukin-1 (IL-1), the activity of which is mediated by fibronectin attachment (Zhu et al., 1998). Since fibronectin is a key substrate for integrin receptors, it follows that IL-1 expression levels could have an influence on that of $\alpha_v\beta_6$ integrin (Ping et al., 1998; Ostberg et al., 1995). Now IL-1 may inhibit the growth of mammalian cells through arresting G1 and G2 phases of the cell cycle (Murai et al., 2001). If IL-1 indirectly up-regulates integrin over-expression, then emulating IL-1 effects in terms of cell growth may increase $\alpha_v\beta_6$ over-expression. This may be achieved using hypothermic cell culture conditions, such as 30 °C, which have been found to induce blockage at the G1 phase in the cell cycle (Moore et al., 1997).

Hypothermic conditions have also been found to reduce the rate of cell apoptosis (Moore et al., 1997), a feature that may also assist with our α_v expression; given that fluorescence signal was only observed in some α_v transfections three days post-transfection, it is likely that this is slow to express, and cells may be dying before α_v has been properly formed.

Transfections at 30 °C and 33 °C were therefore tested. As before, no signal for α_v was observed in subsequent WB analysis, nor was any fluorescent signal observed.

Construct length

The absence of α_v signal may have been a consequence of cleavage of the C-terminal detection tags, rather than problems with expression alone. Alternatively, α_v integrin's short cytoplasmic tail may somehow inhibit the co-association between our integrin constructs.

It is known that the transmembrane helical region of both subunits is essential for stabilising co-association, and facilitates promiscuity between α- and β-subunits, but the short cytoplasmic tails do not assist with homodimer association; instead mediating intracellular signalling via focal adhesins (FAs) (Sastry and Burridge, 2000; Berman et al., 2003), and should, therefore, be dispensable for the purposes of our investigations. To see if this might be the case, new constructs, lacking the cytoplasmic domain, were cloned. However, no expression was detected for either construct, perhaps due to detection tags interfering with membrane association.

Lenti-viral Expression

Since expression levels were too low for purification, two stable cell lines were established using a lentiviral expression system developed within Strubi (Elegheert et al., 2018). This offers a rapid way of establishing a stable cell line (see Chapter 2, Section 1.1.4). Constructs of full length α_v and β_6 were cloned into bespoke pHR-CMV-TetO2 vectors, and, together with packaging vectors, used to generate lentivirus encoding both constructs using a specialist HEK293-lentiX cell line. Virus stocks were then harvested and used to infect HEK293S cells.

Cell lines were then assessed for fluorescence signal. Strong signal was observed for β_6, but little to no signal could be seen for α_v. Following one passage post-infection, $\alpha_v\beta_6$ cells were then sorted according to fluorescence signal (using a Sony SH800Z Fluorescence activated cell sorter), and cells only showing signs of both fluorescent signals were selected. A population of only 4 % of permissive cells (in terms of shape) were found to be dually fluorescent. Due to concerns over the survival of sorted cells, it was necessary to limit the sort time. The resulting 10,000 cells were seeded, with penicillin-streptomycin antibiotics, onto a single well of a 24-well plate, and repeatedly expanded into a T75 flask before freezing.

However, sorted cells grown in static became increasingly clumped and grew extremely slowly, and this worsened with repeated passages with and without trypsin treatment, see Figure 7.6. This phenotype is consistent with the behaviour that we would expect of integrin, and, although reassuringly indicative of the integrin heterodimer forming, it was impossible to grow cells to sufficient scale for purification. To see if cells faired better in suspension culture with a more rapid rotation speed, a fresh batch of sorted cells were defrosted and used to set up a culture in suspension instead. However, regardless of suspension speed, cells became increasingly clumped and failed to grow, see Figure 7.7, precluding attempts to scale up for purification. Fibronectin-coating of tissue culture plates was also attempted to reduce clumping. Fibronectin is a known substrate of $\alpha_v\beta_6$ and thus may encourage spread across the plate, but no clear difference in clumping was observed when cells were plated on fibronectin-coated or uncoated plates.

Inducible cell lines, whereby expression of target genes are induced once cells are grown to a desired density, can overcome such cytotoxicity issues. A stable cell line, using the lenti-system, as before, was prepared, this time using HEK293S-GnTI$^-$ cells. Constructs were also modified slightly; α_v was cloned into a vector containing a MV tag and cytomegalovirus (CMV) promoter

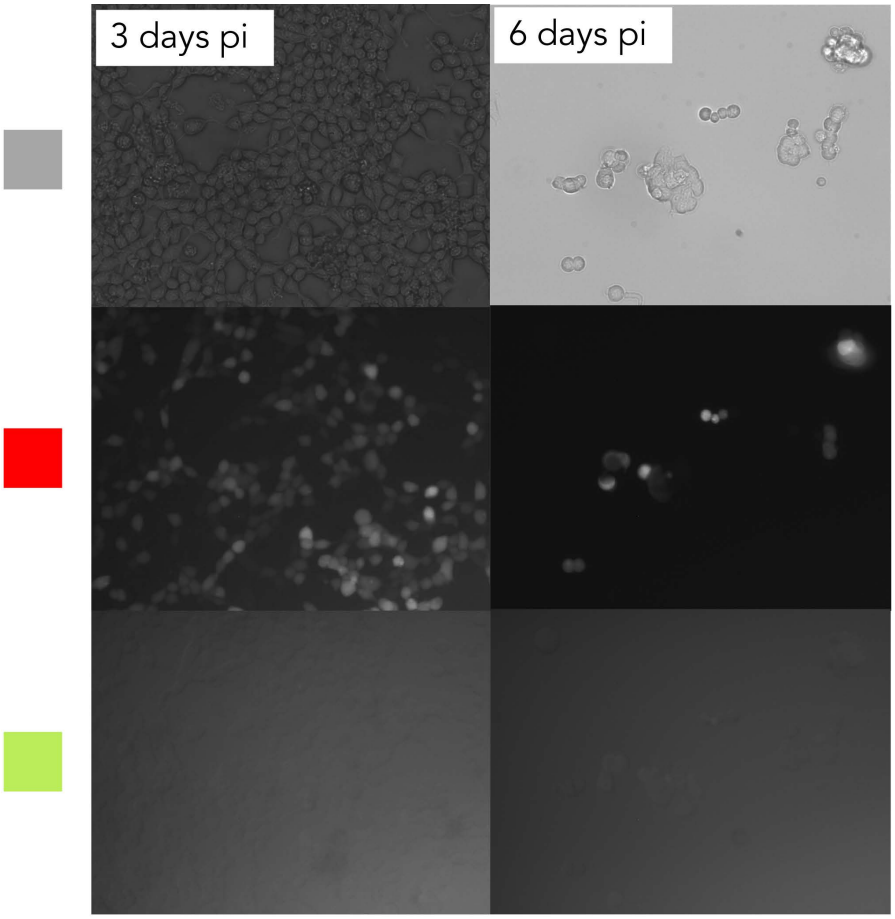


Figure 7.6: Fluorescent microscope images showing the change in cells post infection with $\alpha_v\beta_6$. Boxes on the far left correspond to the colour channel for the adjacent row of images (grey = brightfield). The trio of images on the left show cells 3 days post infection with lentivirus. The red channel shows good expression of the β_6 construct, but, as before, expression was poor for α_v. The right panel shows cells after one passage, 6 days post infection.

(a) (b)

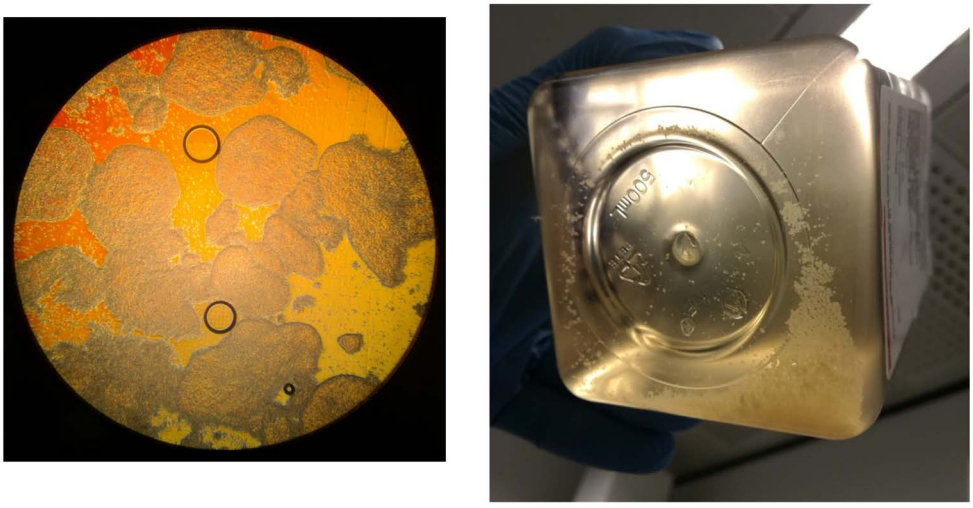


Figure 7.7: Example images of clumped suspension cells (a) aliquot viewed under a light microscope. (b) underside of suspension flask. The cells can be clearly seen as a white slurry.

and β_6 into a vector with a mRuby2 tag and a spleen focus-forming virus (SFFV) promoter. The rationale for using different promoters was that CMV was reputedly more effective for the HEK293S cells, while MV fluorescence tends to be much brighter than mRuby2, and it was hoped that these factors would go some way to equalise the relatively weak expression and signal from α_v.

It has been found previously that the best conditions for induction, in particular, the concentration of Dox (doxycycline, used to induce cells), temperature and histone deacetylase inhibitors (used for stable transgene expression enhancement), is highly construct dependent and so, to find the correct balance for maximal yield of expressed, correctly folded protein, [Dox], additives and temperature were screened for maximal yield. Briefly, stable $\alpha_v\beta_6$ cells were treated with valporic acid (VA), sodium butyrate (NaB) each with 1x Dox, or just 3 x, 1 x or 0.1 x Dox and incubated at either 30 °C or 37 °C. Both VA and NaB are histone deacetylase (HDAC) inhibitors. It is believed that histone hyperacetylation from inhibition of histone deacetylase opens the typically highly compact nucleosome, thereby improving gene accessibility to transcription factors which serves to improve gene expression (Kwaks and Otte, 2006), and has been suggested as a potentially effective additive for inducible cell lines (Elegheert et al., 2018).

After just one day post induction, cells were noticeably clumpy and some detached, consis-

tent with our previously observed phenotype (see earlier). Expression success was monitored by MV and mRuby2 fluorescence (Figure 7.8) and western blotting with Rho-1D4 or Streptactin-HRP antibody (Figure 7.9). Fluorescent signal was exceptionally strong for VA, NaB, and higher concentrations of Dox one day post induction, but the majority of cells were highly clumped and/or detached. Signal also seemed quite strong for the non-induced cells at 37 °C 48 hrs after the other samples had been induced. Cells had been split into plates one day pre-induction, and it is possible that they were still recovering one day post-seeding and expression may become more leaky as cells proliferate. It is therefore conceivable that low temperature conditions may be better for avoiding leaky expression when scaling up cell densities pre-induction, with the caveat that cells will grow slower at lower temperatures.

As may be seen in Figure 7.9(a)), signal is much weaker for the non-induced (negative control), relative to induced samples, consistent with induction having worked, but suggests that there is some 'leaky' expression prior to induction, consistent with observations from fluorescence tests (see above). Such leaky expression has been observed for inducible lines previously, and this may enable FACs enrichment prior to induction (Elegheert et al., 2018). No clear signal could be seen for α_v in the anti-1D4 tag blot for any of the conditions screened, even after long exposure times Figure 7.9(b).

Clearly, induction worked well, and, for future experiments, cells can be grown up in suspension to adequate volumes for purification and then induced. Alternative strategies could include controlling the activation state of integrins using metal ion additives, or down-regulation of integrin binding partners.

7.2.2 Expression of Truncated $\alpha_v\beta_3$

A well established protocol for expression of the head piece (part of the ectodomain, see Chapter 1, Section 1.1.4) of a variety of integrins exists, and has been used previously to generate $\alpha_v\beta_6$ integrin for FMDV binding studies (Dong et al., 2014; Kotecha et al., 2017). Constructs were prepared using this protocol, as described in Chapter 2, Section 2.1.5 and checked by sequencing.

Constructs were then used in a transient co-transfection of 12 2L flasks of HEK293S in suspension (Chapter 2). Following Strep-II affinity and SEC (Figure 7.10(a)) purification steps, the resulting fractions were analysed on a reducing SDS-Page gel and anti-penta-His WB (Figure 7.10(b),(c)).

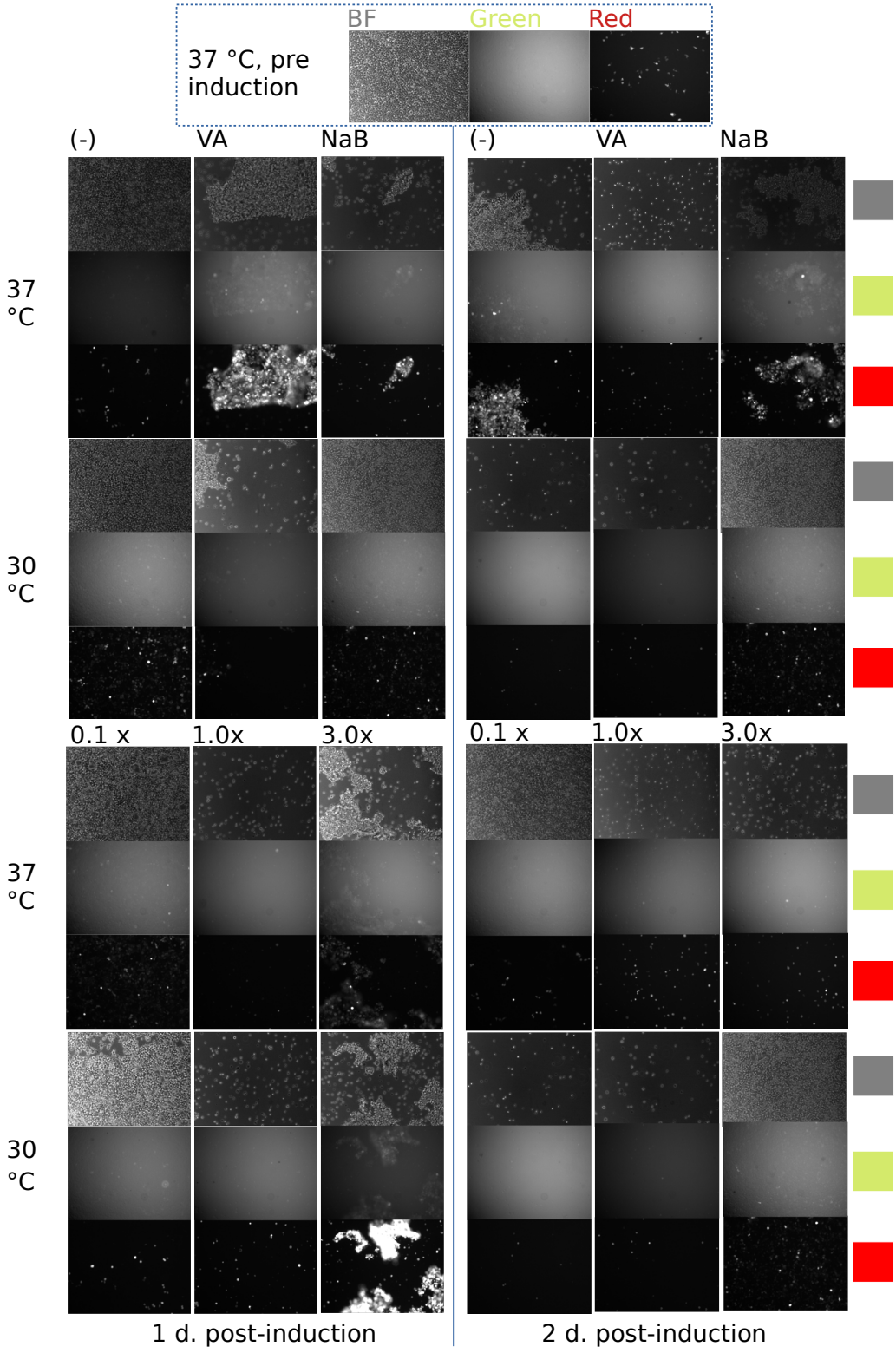


Figure 7.8: Fluorescent microscope images from induction trial. Top blue box shows the cells one day before induction. The different conditions are listed above each row/column (VA = valporic acid; NaB = sodium butyrate; (-) = no induction). Trios of images on the left were taken one day post induction; those on the right, 2 days post induction. Coloured boxes correspond to the imaging channel used: grey is brightfield; green for MV and red for mRuby2. The temperature at which condition row was conducted is listed on the left.

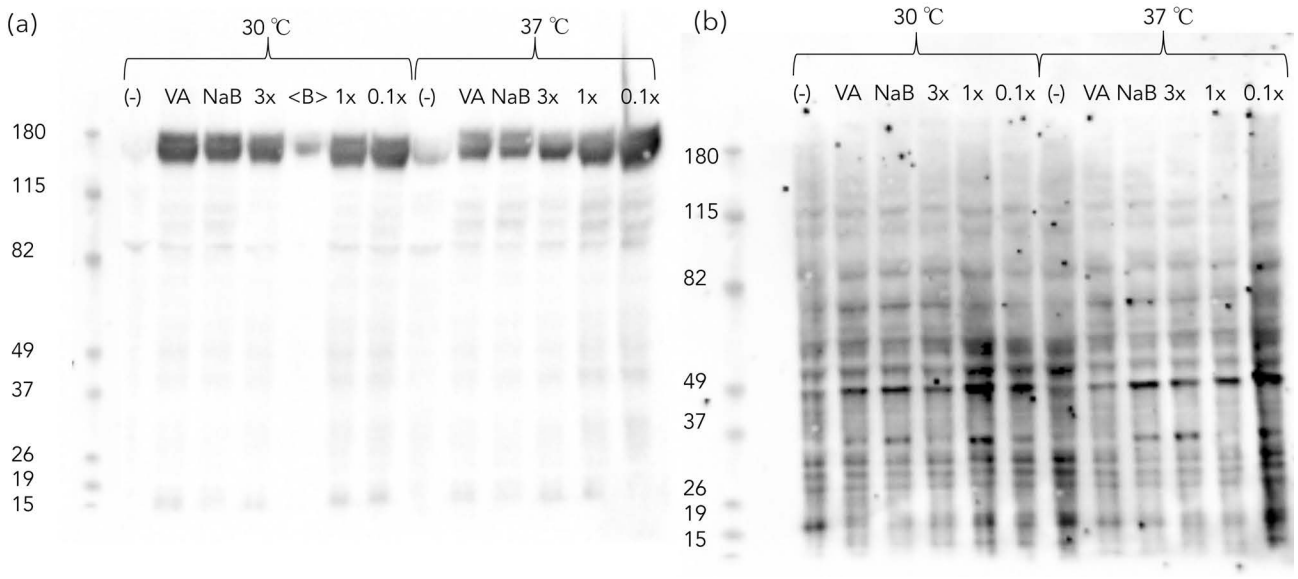


Figure 7.9: Western blots of collected cells from induction trial. Conditions are labelled for each well, with VA = valporic acid; NaB = sodium butyrate and <B> refers to a blank well. Values for both ladders are in kDa, and the ladder displayed (a) has been taken from a longer-exposure image. (a) HRP-conjugated anti-strepII western blot imaged with a 0.5 s exposure time. (b) anti-1D4 western blot after 10 s exposure time.

The presence of bands at ca. 120 and 150 kDa were consistent with previous observations from trials conducted at Pirbright Institute (personal communication, Alison Burman), and represent β_3 and α_v monomers.

Negative stain grids were then prepared to assess whether sample quality may be sufficient for cryoEM Figure 7.10(d), despite the broad SEC trace and multi-band SDS page gel.

Provisional Cryo-EM Analysis

Cryo-electron microscopy samples were prepared to establish whether semi-pure $\alpha_v\beta_3$ binds to FMDV A22cc VLPs. This method was high risk for three reasons: first, to my knowledge, no VLP/receptor binding studies have been performed, second, there are no published FMDV-bound $\alpha_v\beta_3$ structures, albeit evidence does exist that this can be utilised as a receptor (Neff et al., 1998), and, third, the integrin sample was far from pure.

Based on protocols established previously (Kotecha et al., 2017), $\alpha_v\beta_3$ was incubated at room temperature for one hour with A22cc VLPs and 2 mM $MgCl_2$, at final concentrations of 0.236 μM VLP and 70.8 μM $\alpha_v\beta_3$, which should approximately equate to 5 integrins per site, assuming 60 sites

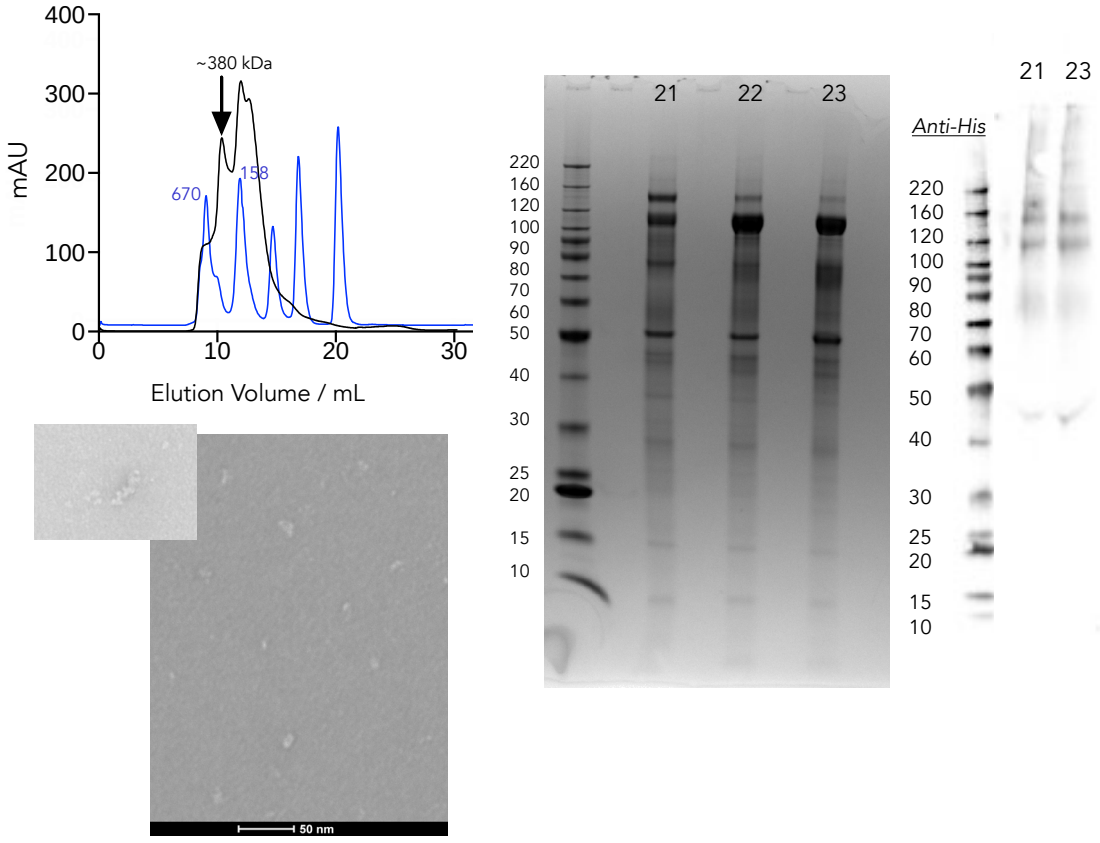


Figure 7.10: (a) SEC trace (black line) from $\alpha_v\beta_3$ purification, with standards and associated MW for two pertinent standards in kDa shown in blue. The region where fractions 21-23 were collected is shown with an arrow. (b) SDS-page gel for fractions 21-23, ladder in kDa. (c) Image from anti-penta-His Western blot (see Chapter 2 for method and antibodies) for fractions 21 and 23. The ladder values are shown on the left in kDa. (d) Negative stain image (uranyl formate stain) from fraction 23, diluted 1/10 , scale bar 50 nm. Inset is an example from another micrograph, showing some potential aggregation.

per VLP. The integrin concentration was determined by Beer-Lambert's law using the ϵ coefficient that was estimated using ExPASy and the absorbance at 280 nm (Gasteiger et al., 2003). Samples were then returned to ice before application to lacey grids and plunge freezing in liquid ethane using the standard plunge freezing protocol outlined in Chapter 2 Section 2.5.2.

Data Collection and Processing

Micrographs, 1302 in total, were collected on a Glacios cryo-electron microscope (Thermo Fisher) at a nominal magnification of 120kX. With hindsight, given the sparsity of VLP/integrin distribution, it might have been sensible to have used a lower magnification, see Figure 7.11(a). The massive excess of integrin also resulted in a relatively high background (Figure 7.11(a)).

A total of 2904 particles were then automatically picked from 1052 of the collected micrographs using Warp (250 of the remaining micrographs were devoid of particles, or had poor lace-distribution), using a training set from the previous A22cc data collection (see Chapter 2, Section 1.5.4). Micrograph stacks were then motion corrected and CTFs fitted as per Chapter 2 Section 1.5.4, and an example micrograph may be seen in Figure 7.11(a). Particles were then filtered by 2D classification into 100 classes, where the E (expectation) step limited to 7 Å, leading to the exclusion of 700 particles that appeared broken, or miss-picked. Good classes may be seen in Figure 7.11(b), and, interestingly, all classes had a white 'cloud' near the particle centre, which may represent bound integrin. This does not necessarily mean, however, that all particles were at least partially decorated, rather it seems likely that bound/unbound particles were not distinguished at the 2D classification stage. Selected classes were then either (i) classified into 5 classes with icosahedral symmetry, or (ii) subjected to an initial refinement with icosahedral symmetry. Particles from (ii) were then classified, and the hopeful looking 72 particles (3.3 % of selected 2D classes) refined with icosahedral symmetry (with symmetry, equating to effectively 4320 particles) to a reported 20.3 Å resolution. For strategy (i), 312 particles were selected from 3D classes (14.2 % of selected 2D classes) showing some density on the outer rim of the particles, and were refined, also with icosahedral symmetry, to only 34.9 Å resolution.

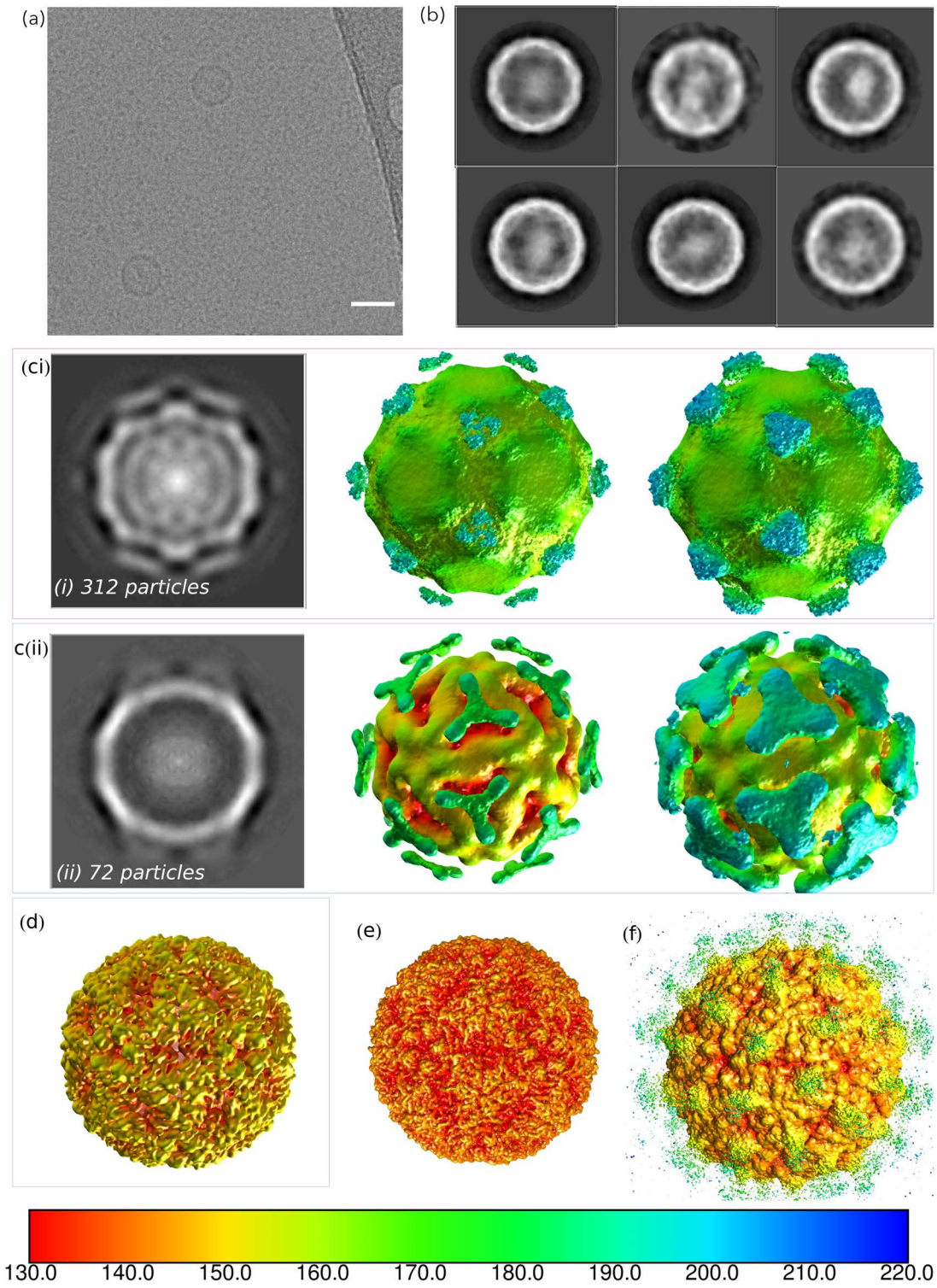


Figure 7.11: (a) Example section from a motion corrected micrograph, scale bar = 30 nm. (b) Example classes from 2D classification. (ci) and (cii) Two processing groups, from left to right: projection of final refinement, and 3D reconstructions at two contour levels coloured by radius. Processing route (i) and (ii). (d) Initial refinement from all 2D classes, reclassified during route (ii). (e) and (f) Icosahedrally averaged reconstruction from FMDV O1M $\alpha_v\beta_6$, as published in Kotecha et al. (2017). Scale bar (bottom) represents radial distance in Å.

Comparison with Native FMDV

The initial reconstruction, prior to classification for strategy (i), shown in Figure 7.11)(d) appears slightly expanded, with a minor opening at the 2-fold axis, relative to the published FMDV O1M $\alpha_v\beta_6$ structure at lower contour levels, Figure 7.11(e). This is consistent with observations outlined in Chapter 5 and it is not entirely clear whether this represents a mixed population of expanded/native states, and/or whether integrin preferentially binds non-expanded particles.

Reconstructions at two contour levels, following classification employing strategies (i) and (ii), may be seen in Figure 7.11(ci), (cii), respectively. For strategy (i), the reconstruction shows possible integrin density at the vertices of each pentameric unit. For the lower resolution Figure 7.11(cii) reconstruction, capsids seem to be adorned with propellers at each of the pentameric unit vertices which may represent density from three adjacent integrins. This seems to be in agreement with the positioning observed in the published FMDV O1M $\alpha_v\beta_6$ structure Figure 7.11(f), where integrin attached to FMDV's GH-loop. However, this should not be read into too closely; the overall topology of the bound particles differs greatly from what we would expect for FMDV, and is presumably a consequence of the only small population of particles averaged and very low resolution.

The low percentage of decorated particles selected in each strategy may well be due to preferential binding to a small population of non-expanded VLPs, but, given that, for previous studies using $\alpha_v\beta_6$, only 75 or 38 % of sub-particles for PanAsia and O1M, respectively, were decorated, it is therefore possible that much of the integrin signal is simply 'washed out' from icosahedral averaging (Kotecha et al., 2017).

7.2.3 Investigating the use of caged protons for on-grid pH drop

To probe the utility of NPE-caged protons as a means of on-demand on-grid acidification, a series of control experiments were conducted to probe the kinetics of proton release.

FMDV is typically purified using a HEPES or Tris buffer (usually at 50 mM concentration) at a pH of 7.5-8.0, to stabilise particles against degradative acidic environs. Although it is possible to dialyse samples prior to grid preparation, minimal potentially disruptive handling is desired. Thus, it was necessary to establish a suitable buffer concentration that was permissible for proton release within a ms time scale.

To establish an optimal ratio of (i) [HEPES pH 7.5] : [caged-proton], and (ii) LED-exposure time for efficient proton release, control experiments, tracking the change in bromothymol blue (BTB) indicator absorbance with pH, were performed using a micro-spectrophotometer (Andor Solis, Andor Technology) at Diamond Light Source, UK. An in-line camera was used to check that the sample filled the entire well before absorbance readings were taken, Figure 7.12(a)(ii). To ensure that the sample depth was consistent, and comparable to the volumes on a grid, it was necessary to design a sample holder plate with grid-sized sample wells, Figure 7.12(b).

First, standard curves of absorbances for known pH values were recorded using stock solutions of BTB, and Succinic Acid Sodium Dihydrogen Phosphate and Glycine buffer (SPG) solutions, across a range of wavelengths. As may be seen in Figure 7.13, peak BTB absorbance measurements occurred at 445-450 nm and 615-620 nm, consistent with literature (Klotz et al., 2011; Kara and Alkan, 2000). The first, *circa* 450 nm, peak corresponds to the proportion of protonated indicator in solution, and is therefore higher for more acidic solutions. The second peak corresponds to the deprotonated indicator form.

As may be seen in Figure 7.13(b), the concentration of BTB has a substantial impact on absorbance intensity measurements. To eliminate this variable, we took the ratio of absorbances between the two peak wavelengths for each pH. To derive a usable standard curve for pH estimation, the log of the ratio between the two peak wavelengths was taken for each pH, and a straight line was fitted. In order to establish the time dependence of acidification, solutions of NPE-cage and HEPES buffer were exposed, in a controlled fashion using a fast-shutter, to LED light. Absorbances as a function of the peak wavelength were then plotted, and pHs were estimated using the equation derived from Figure 7.14, equation (7.1).

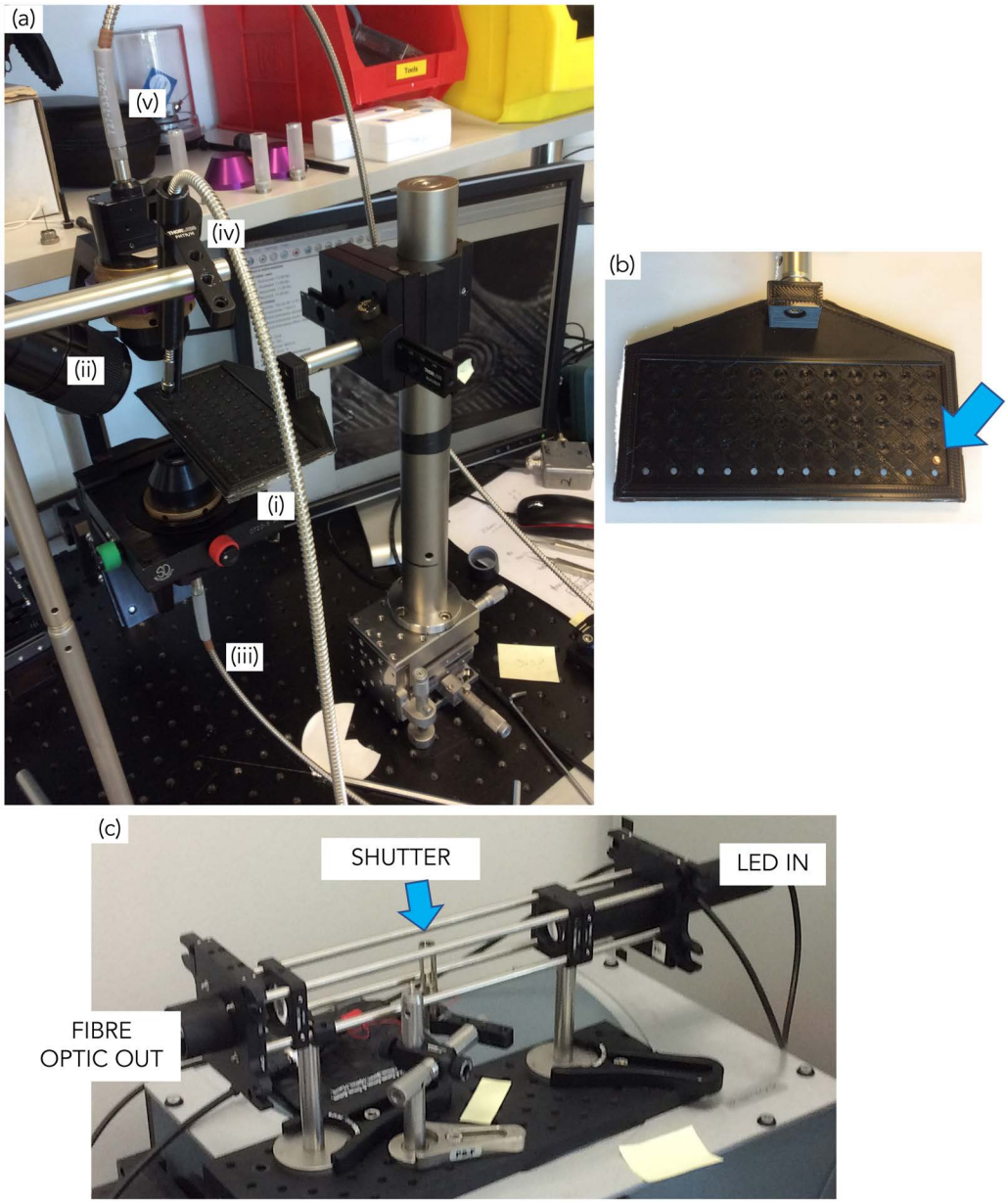


Figure 7.12: (a) General set up of the microspectrometer. (i) Sample stage; (ii) Camera focussed on well, image can be seen on the screen in the background. (iii) is the fibre optic from light source. This was connected to a 300 nm filter to minimise caged proton excitation. (iv) Fibre optic to LED for cage excitation. This is linked to the shutter set up in inset (c). (v) fibre optic to detector. (b) Sample stage, with grid (indicated by blue arrow) for comparison. The bottom of the wells was coated with clear tape for each set of measurements. (c) Shutter for probe-LED source, and lens set-up. Shutter control box is omitted.

$$0.9716 \cdot pH - 6.869 = log_{10}\frac{A_{\lambda 450}}{A_{\lambda 620}} \tag{7.1}$$

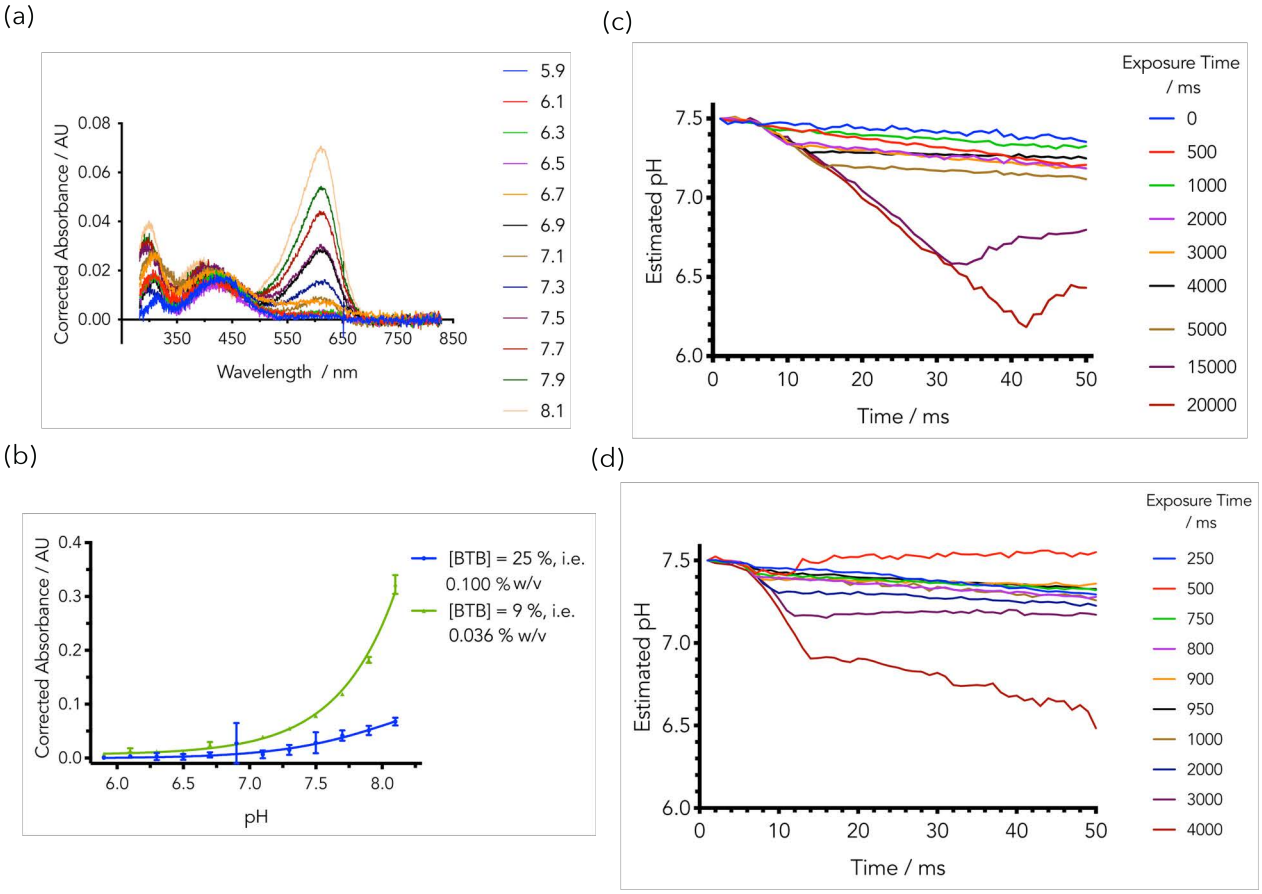


Figure 7.13: (a) Comparison of absorbance profiles for BTB/SPG solutions at different pHs over a range of wavelengths. pH is indicated in the key on the right. The 'fuzzy' line at wavelengths below 400 nm is a consequence of the filter implemented to reduce premature photo-excitation in caged-proton-containing samples. This filter was included in this experiment to allow for any potential dampening effects that this may have on signal. (b) Corrected absorbances for BTB/SPG solutions, at pHs 5.9 to 8.1. Concentrations of BTB solutions are indicated in the key. Fitted lines are sigmoidal. For (c and d), recording time is plotted along the x axis, while lines are coloured by exposure time i.e. the length of time the samples were illuminated with an LED. (c) pH drops observed for a variety of time points with 1 mM cage compound and 0.05 mM HEPES solution (20:1 ratio cage:HEPES), buffered to pH 7.5. pH values were estimated using the linear regression line on standard curve, above. Slight recovery in pH after longer exposure times could be due to the small quantity of buffer in the mix. (d) pH drops observed for a variety of time points with 5 mM cage compound and 0.05 mM HEPES solution (100:1 ratio cage:HEPES), buffered to pH 7.5. pH values were estimated using the linear regression line on standard curve, above. LED exposure started after 5 s.

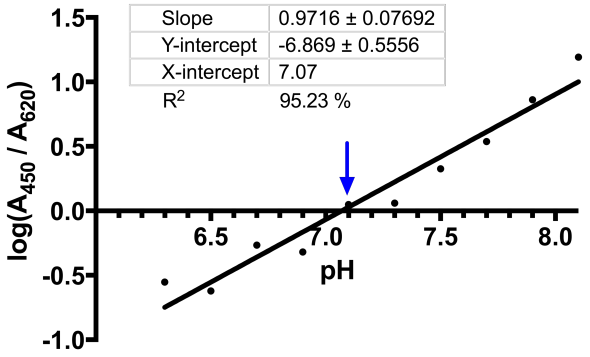


Figure 7.14: Standard curve for BTB indicator, with a straight line fitted using GraphPadPrism. The likely pKa is indicated by a blue arrow, and is consistent with literature.

Mixtures of BTB, HEPES and cage were then prepared, and monitored using the micro-spectrometer set up described earlier to discern how quickly a pH drop to pH 6.5 could be achieved. It was noted that the caged proton showed some photosensitivity in daylight, and, even greater sensitivity from the light source. In an attempt to combat the former issue, initial experiments were conducted under a tent made by cardboard boxes, but this was too disruptive to the experimental set up to warrant pursuing and the decision was made instead to keep the laboratory lights switched off. To minimise premature excitation from the spectrometer probe light source, it was necessary to introduce a filter at approximately 400 nm, at the cost of signal.

With a 20:1 v/v ratio of [cage] : [HEPES pH 7.5], a pH drop just below pH 6.3 was achieved after a 20 s exposure, Figure 7.13(c), while a final pH of 6.5 was achieved with a 100-fold excess of cage over HEPES, and, given that this line keeps to a downward trend, see Figure 7.13(c), this seemed sufficient to overcome the buffering capacity of 0.05 mM HEPES. As anticipated, the pH drop was also sharper with higher concentrations of NPE cage (Figure 7.13(d)). Strangely, recovery to neutral pH was observed with the 20:1 caged proton sample for exposure times exceeding 10 s. Given that an identical concentration of HEPES was used and a lower pH was recorded for the 20:1 sample after a long exposure time, this discrepancy in pH recovery must be related to either the higher concentration of cage somehow interfering with the BTB signal, or heating from the longer exposure time causing a protonation reaction with one of the side products (no obvious reactions can be determined based on inspection of the anticipated products). This is not a major concern, however, provided the acidic environment is maintained long enough to trigger particle

uncoating, but does somewhat limit precise control over exposure to acidic pHs.

Collectively, these results indicate that, with a traditional LED light of the correct wavelength, it is not possible to achieve a controlled drop in pH below 6.5 on the ms timescales desired for our pH drop experiments, while high concentrations of cage proton will be required to achieve a sufficient pH drop. The light sensitivity of NPE cages is also a concern, and certainly limits the control we would have over low pH exposure.

7.2.4 Additive Stabilisation of FMDV intermediates

Stabilisation of otherwise too transient FMDV states could offer an alternative, or, in the case of the caged proton experiments, complimentary, method to capture uncoating intermediates, serving to increase the permissible time scale over which intermediate states may be captured.

Previously, it was shown that incubation of acid treated ERAV with 0.1 mg/mL BSA and 2 mM $CaCl_2$ enabled the isolation of empty ERAV particles (Tuthill et al., 2009). Such additives are also known known to stabilise PV A-particles (Tuthill et al., 2006)). Given FMDV's similarities with ERAV, it seemed conceivable that such conditions may facilitate stabilisation of intermediate FMDV states.

The thermostability and acid lability of inactivated FMDV with and without BSA and 2 mM $CaCl_2$ was provisionally tested by thermofluor using the genome detection dye SYT09 (Chapter 2, Section 2.7). Melting temperatures, reflecting the midpoint of RNA release curves, were determined using both five-parameter and Boltzmann sigmoidal fits. The latter method assumes the distribution to be symmetric. The five-parameter fit estimate of Tm = 60 °C which, is an overestimate and does not agree with assessment of the first derivative plot alone, perhaps due to the curve plateauing, rather than decreasing at elevated temperatures. Nevertheless, as may be seen in Figure 7.15, capsid thermostability improved substantially in the presence of BSA/ $CaCl_2$ by as much as 10 °C (see Tables 7.2 and 7.3).

We also see, albeit to a much lesser extent, a slight delay in the midpoint of RNA release signal in conditions containing EDTA (ca. 1 °C). For this condition, 1 mM final concentration of EDTA was dialysed against SAT2 the day before for 2 hrs 15 mins before concentration by resting the dialysis membrane on solid PEG 20K, before overnight storage at 4 °C. EDTA has been shown to interfere with some thermofluor dyes, such as SYPRO Orange (Kroeger et al., 2017). However, the curve profile does not look unusual, and, perhaps a more credible explanation is that some smaller

Table 7.2: Boltzmann sigmoidal fit of first derivative dataset from thermofluor. Least squares regression line, fitted between temperatures 25-65 °C.

	SAT2 only	SAT2 + BSA/$CaCl_2$	SAT2 + Myristic acid	SAT2 + EDTA
Tm / °C	46.6	55.5	40.1	47.7
Confidence (+/-) / °C	0.28	0.09	0.46	0.37
R^2	98.0	99.7	96.3	96.1

Table 7.3: 5-parameter fit of first derivative dataset from thermofluor, fitted between temperatures 25-65 °C.

Sample	SAT2	SAT2+BSA/$CaCl_2$	SAT2 + Myristic acid	SAT2 EDTA
Tm / °C	46.9	60.7	40.4	47.9
+/-	0.29	0.22	0.43	0.41
R^2	98.3	98.7	97.5	96.2

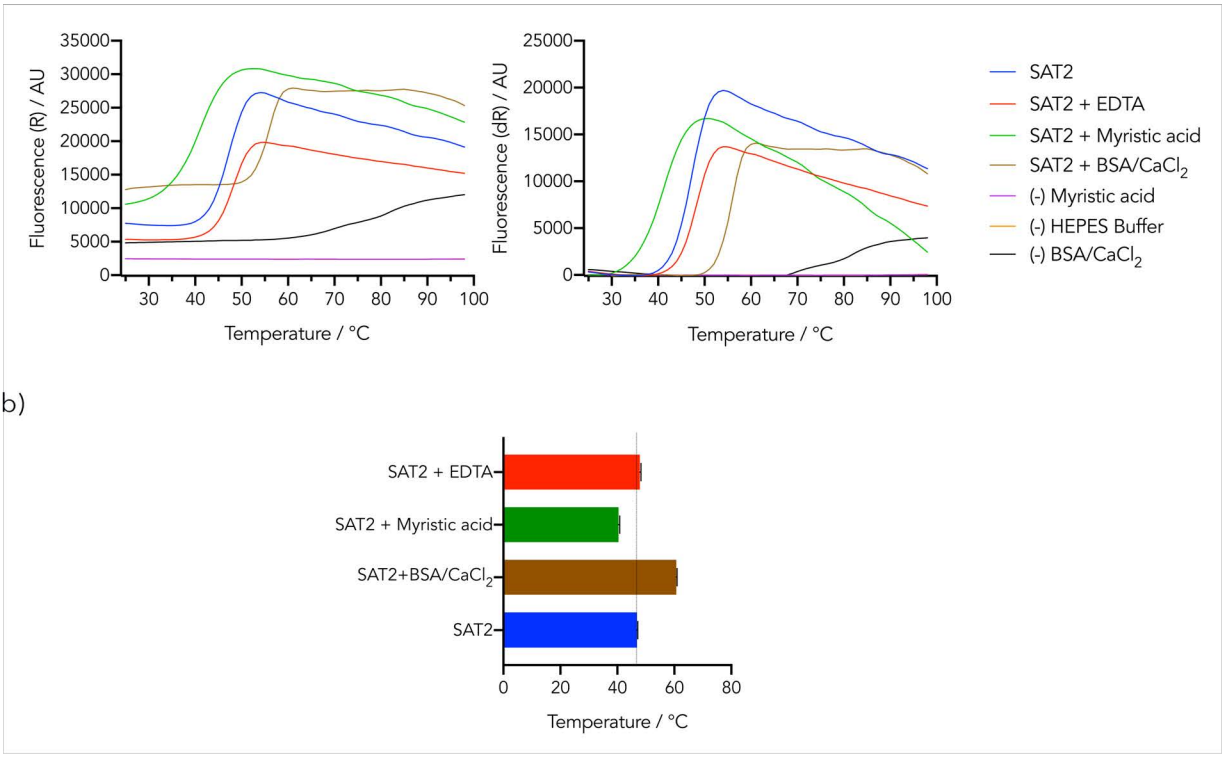
PEG products leached into the system during concentration, and may have served to stabilise the sample. The concentration of EDTA used for dialysis may have been too low, especially since no change in A260/280 readings were observed after 2 hrs dialysis (still 1.72), but this dialysis time was sufficient for Rowlands et al. (1975).

Perhaps notably, 2 μM myristic acid dissolved in 50 mM HEPES, 200 mM NaCl at pH 8.0 caused a much earlier peak in RNA signal (ca. 6 °C lower than the control), and given the stabilisation effect of BSA/$CaCl_2$, it would be interesting to see if this condition combined with BSA and or 2 mM $CaCl_2$ could stabilise against myristic acid.

To corroborate these results, SAT2 FMDV samples heated with and without BSA, 2 mM $CaCl_2$ additives were immobilised on cryo-EM grids and screened on a Glacios (Thermo Fisher) cryo-TEM with a FIII detector. As shown in Figure 7.15(c), a slight, but not complete, protection was conferred against elevated temperatures when the additives were present. It would be interesting to see whether the particles in the heated additive sample are slightly expanded.

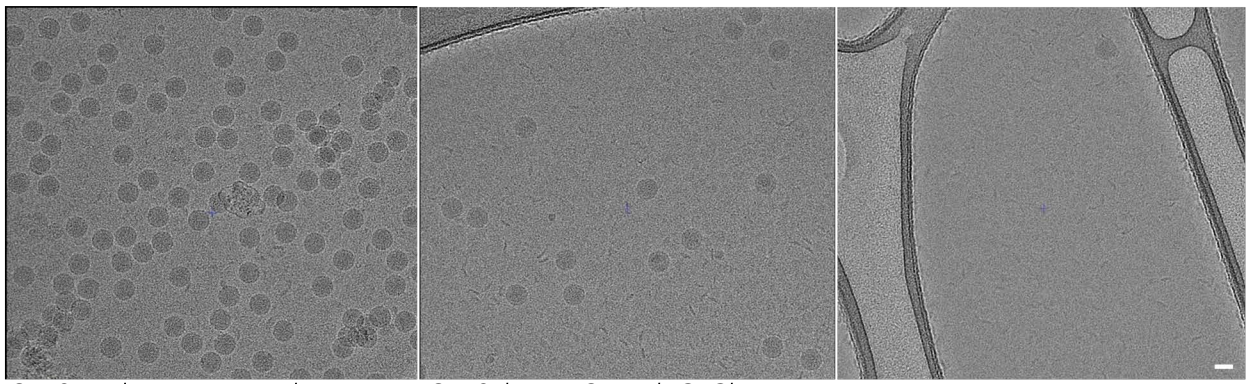
As mentioned previously (see Chapter 1), Ca^{2+} ions are known to stabilise FMDV capsids somewhat against elevated temperatures (Kotecha et al., 2016). Further work is required to establish whether BSA plays a role, if at all, in stabilising the FMDV capsid.

(a)



(b)

(c)



SAT2, no heat, untreated *SAT2, heat, BSA with $CaCl_2$* *SAT2, heat, untreated*

Figure 7.15: (a) Raw and first derivative curves from thermofluor experiments investigating the effects of additives on SAT2. Negative controls, indicated by a (-) did not have any virus sample present. (b) Bar chart showing Tms determined from a five-parameter fit of the curves in (a), coloured as per the key shown in (a). The dotted line shows the Tm for SAT2 only, and, confidence bars are shown in black. (c) Cryo-EM images of SAT2 treated with and without heat and additives, as stated below each micrograph. The scale bar is shown in white on the rightmost image panel and corresponds to 30 nm.

7.2.5 A Low Resolution ERAV Expanded Particle

In an effort to produce sufficient ERAV material for crystallographic metal ion experiments, as in 4 and Chapters 5, a number of ERAV preparations were performed using previously developed methods (Plummer and Kerry, 1962; Tuthill et al., 2009; Caridi et al., 2016). HeLa or Vero cells were infected with stocks of ERAV, provided by Toby Tuthill, Pirbright Institute where I learnt the protocol for virus growth and purification (see Chapter 2, Section 2.2.1).

Preparations typically resulted in comparatively low particle yields, such that no discernible band was observed at the fractionation step, thus 0.5 mL fractions were sequentially taken and the absence/presence of virus determined by SDS-page gel. These exceptionally low virus concentrations allowed limited crystallisation trials that were unsuccessful (not covered in this thesis). Faint bands were observed on SDS-page gels for earlier fractions consistent with the presence of VP0, the precursor to VP4/VP2 that is cleaved upon maturation for most picornavirus particles (Chapter 1 Section 1.1.2, Figure 7.16(a)), consistent with the presence of empty capsids (no bands could ever be seen for these fractions on a 15-45 % sucrose gradient, even for more concentrated mature virus). Empty capsids have never been seen in ERAV purifications (personal communication, Toby Tuthill), at least not at a sufficiently high concentration to be resolved on a SDS-page gel.

Empty ERAV samples were pooled, desalted and concentrated to approximately 1 mg/mL and plunge frozen suspended on c-flat 200-mesh grids at 4 °C, 100 % humidity. Grids were initially screened on a T12 microscope (Tecnai), before being clipped and loaded onto a Titan Krios (Thermo Fisher) for data collection.

Due to multiple handling steps, grids were of very poor quality, with numerous broken holes and extreme contamination. Nevertheless, sufficient particles could be seen to warrant a quick data collection in linear mode. This was set up with a nominal magnification of 59 kX, corresponding to a calibrated pixel size of 1.36 Å/pix (calibrated with a rotavirus sample collected at a similar time, personal communication, Jingshan Ren). A total of 1297 micrographs were collected, from which 2068 particles were manually selected using Relion's 'manual pick' GUI (Scheres, 2012) to ensure selection of particles outside of broken holes, and to avoid miss-selection of contamination. Particles were then filtered by 2D and 3D classification with icosahedral symmetry in Relion to only 233 particles (Scheres, 2012). The reason for this dramatic culling of particles seemed to be due

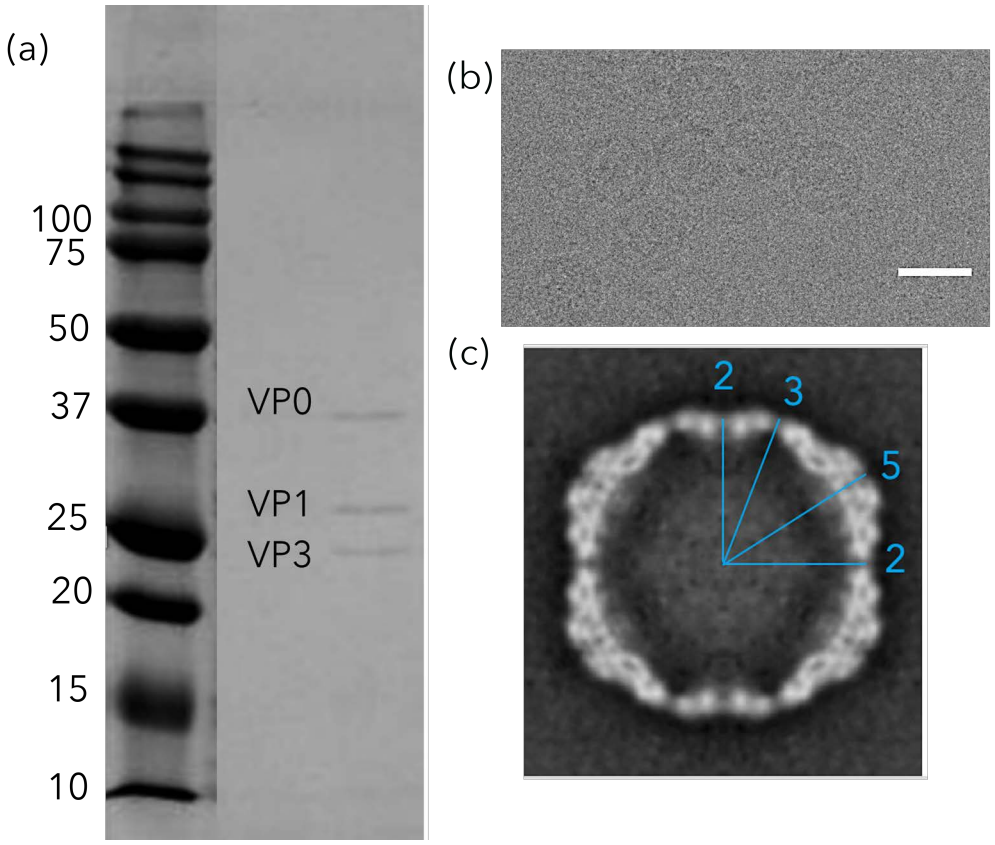


Figure 7.16: (a) Reducing SDS-page gel confirming the presence of empty particles produced from infected vero cells, virus proteins and ladder (kDa) marked. Note that the gel has been truncated to the ladder for clarity. (b) Example electron micrograph from collection, scale bar shown as a white rectangle, 50 nm. (c) Central section from refined ERAV empty particles. 2, 3 and 5-fold axes are marked in blue.

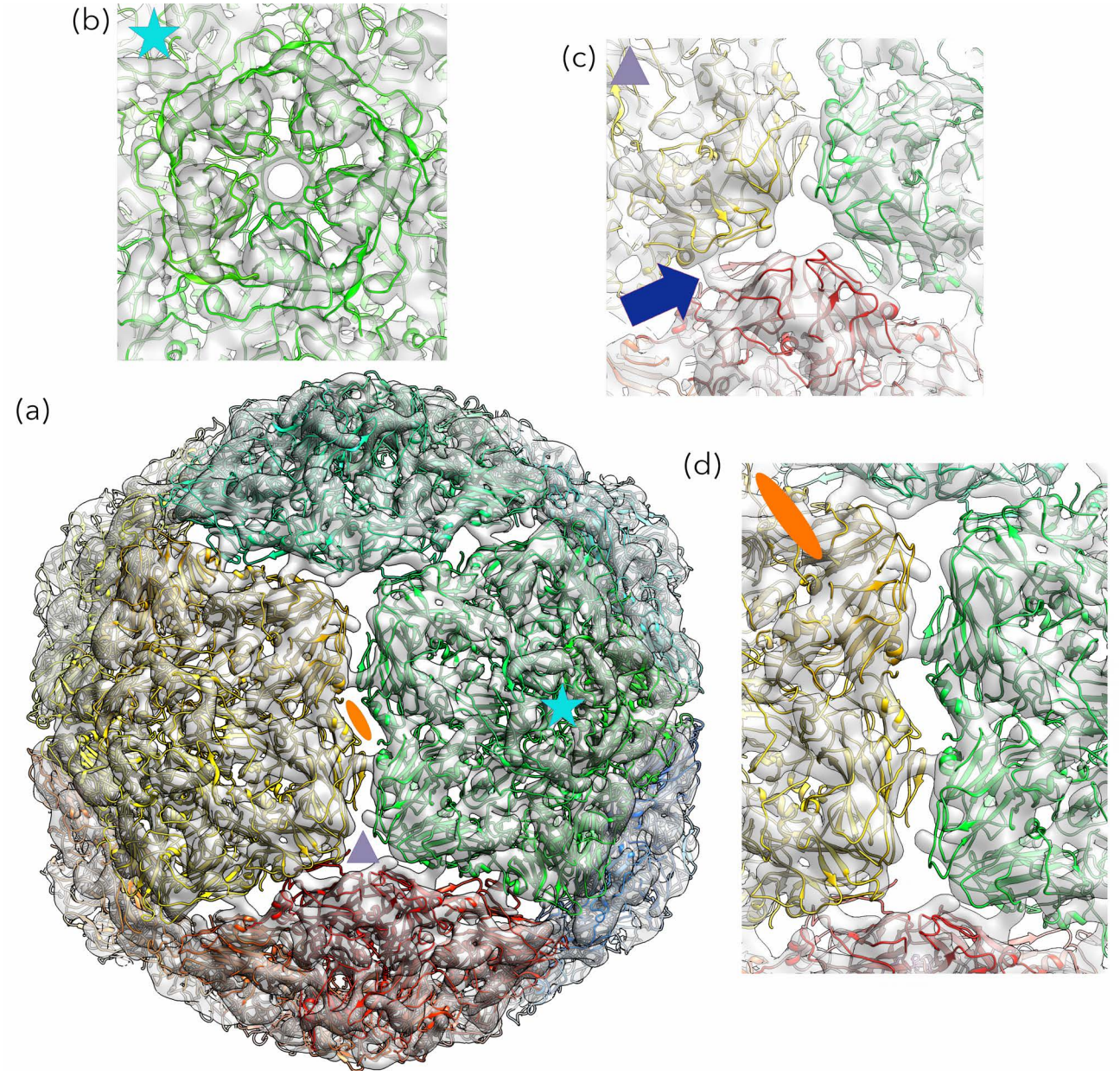


Figure 7.17: (a) Rigid body fit of ERAV (PDB: 2WFF) into our expanded ERAV map (pale grey) (Tuthill et al., 2009). The two-fold axis is marked with an ellipse, three-fold with a triangle and five-fold with a star and pentamers individually coloured for clarity. (b) Zoomed in view of five-fold axis. (c) Cropped view of three-fold, with an interesting bridge of density indicated with a blue arrow. (d) View of 2-fold. Note that, for (b-d), the centre of the image corresponds to the axis centre.

Table 7.4: Correlation coefficient values for rigid body fit of pentamers from native ERAV (2WFF) into our low resolution empty ERAV map, as determined by Phenix comprehensive validation (Tuthill et al., 2009; Adams et al., 2010).

CC	Value
mask	0.62
mainchain	0.82
box	0.86
volume	0.60
peaks	0.57

to a substantial proportion of particles deviating from icosahedral symmetry (many projections appeared to be slightly non-circular), but should be investigated further. Data were processed in Relion (Zivanov et al., 2018), using the same protocol as before (see Chapter 2. Imposing icosahedral symmetry resulted in a 9.71 Å (FSC = 0.143) reconstruction (see Figure 7.17(a), 7.18(a)). As can be seen in Figure 7.16(c), there is slight shading in the capsid interior, but far less than we would usually observe with a genome-containing virus. In addition, small pores can be seen at the 2- and 3-fold axes (Figure 7.16(c), 7.17(a)).

To aid with density interpretation, an existing ERAV crystallographic structure (PDB: 2WFF) (Tuthill et al., 2009), arranged into a pentamer was initially rigid body-fitted into the resulting reconstruction using Coot (see Figure 7.17) (Emsley and Cowtan, 2004; Emsley et al., 2010), before individual rigid body fitting of individual protomers. The CC values determined by Phenix are shown in Table 7.4 (Adams et al., 2010, 2011).

Provisional Structural Analysis

As may be seen in Figure 7.17(b), the five fold axis seems to be non-expanded, contrasting the three-fold axis, where we see a substantial expansion relative to native virions, with the spherical pore opened to a trefoil-type shape, held together by 'threads' between VP2 and VP3 units (blue arrow, Figure 7.17(c)). The distance between C_α atoms at the F2118 residues that sit at each of the three corners of the three fold axis is 11.4 Å in native ERAV, contrasting an 18.2 Å separation in our expanded form.

The map is poor around the modelled VP3 close to the three-fold, especially at the 3192-3201 loop that typically sits at the top of the three-fold axis lip, and it is possible that remodelling of this region and the lower 3115-3120 E strand facilitates the formation of the bridges that we see

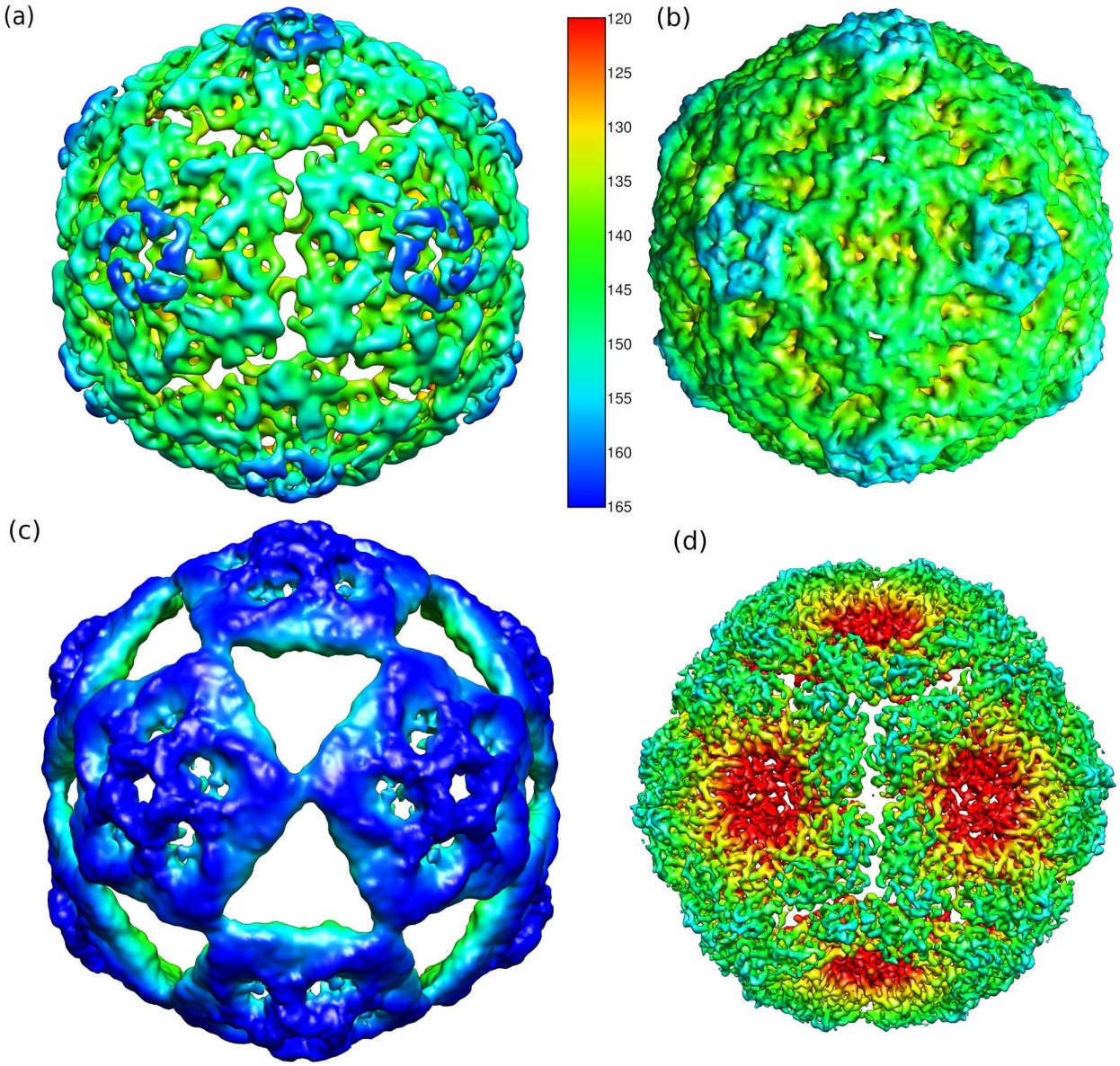


Figure 7.18: Comparison of an expanded ERAV empty capsid with existing PDB structures. All structures are coloured by radius, according to the scale bar shown adjacent to panel (a), where distances are listed in Å. (a) Reconstruction of ERAV empty. (b) ERAV native (PDB = 2WFF) (Tuthill et al., 2009). The map was generated using the 'molmap' functionality in Chimera, with a 9 Å resolution limit to aid comparison (Pettersen et al., 2004). (c) Expanded ERAV, PDB = 4CTG, (Bakker et al., 2014) reported to be 17 Å resolution. (d) An inside-out miss-formed FMDV particle (PDB=5OWX) at 5.2 Å reported resolution. (Malik et al., 2017)

at each propeller of the three-fold trefoil (Figure 7.17(c)). These observations are commensurate with reports of an FMDV A1061 pentamer by Malik et al. (2017), where one of the most notable differences between individual pentameric units and the mature virion was a 4.6 ° shift of the bulk VP3. It may be speculated, therefore, that this shift serves to stabilise genomic-free capsids. This interpretation should be treated with extreme caution, given the low resolution of our map and limited modelling performed.

Expansion is also clear at the 2-fold axis, where a substantial opening can be seen. The distance between modelled C_αs of two E2096 residues (located in a position central to the 2-fold axis) from protomers on either side of the 2-fold axis, is double (12.7 Å) that of the native model (ca. 6 Å Figure 7.17 (d)).

Comparison with existing Aphthovirus structures

As may be seen in Figure 7.18(a,b), relative to the native ERAV structure, our empty particles are expanded 'outwards' for each pentameric unit, with no obvious rotation.

The arrangement of individual pentameric units and extent of expansion for our structure Figure 7.18(a) is vastly different to that of a previous ERAV empty intermediate (Figure 7.18(c)). This intermediate reported in Bakker et al. (2014) was reconstructed from micrographs of purified native particles, and this has been subsequently proposed to be a product of reassembled pentamers (Malik et al., 2017). In contrast to our structure, this intermediate was reported to be held together by the F2118 residues that typically sit at each vertex of the three-fold axis triangle and individual pentamer units are substantially rotated (Figure 7.18(c)).

7.3 Discussion

The journey towards creating a reliable expression system for further structural study of $\alpha_v\beta_6$ integrin was certainly not straightforward. Many of the difficulties encountered with integrin expression likely derive from my focus on expression of 'native' full length construct, and it may be useful to investigate constructs lacking the cytoplasmic domain further, perhaps with a more extensive panel of flexible linkers. Nevertheless, recent results with an inducible cell line are encouraging and present a viable way forward, if it is possible to expand cells to sufficient quantities before

clumping from leaky expression occurs.

The apparent binding of truncated $\alpha_v\beta_3$ to FMDV VLP sample was unprecedented, and offers two exciting avenues; first, truncated integrin may be incorporated into a model system to capture FMDV in the act of uncoating (*vide infra*). It will be necessary to initially establish whether sufficient discrimination can be attained, possibly through capture of VLPs on grids functionalised with our integrin constructs via nickel-chelating head groups, or through use of an ELISA. First, however, to avoid needing the large quantity of costly media and incubator space needed to expand cells for transfection to sufficient quantities to mitigate low expression levels, a stable cell line should first be established. This should be trivial with the lentivirus expression system used for the full-length $\alpha_v\beta_6$ constructs. For future microscopy studies with truncated $\alpha_v\beta_3$, particle decoration can be improved by using Mn^{2+} instead of Mg^{2+} as the integrin activator additive, as this was previously found to be more effective at pushing the integrin conformational equilibrium to the 'active' state that has an affinity for FMDV capsids (Kotecha et al., 2017).

The use of truncated integrin constructs may be a more attainable strategy for our ultimate goal to characterise the FMDV uncoating process in a cell-free system (Bubeck et al., 2005a; Tuthill et al., 2006; Strauss et al., 2013; Lee et al., 2016; Groppelli et al., 2017). It should be possible to link the His-tag of $\alpha_v\beta_3$ constructs to nanodiscs or liposomes functionalised with a nickel-chelating head group, in an analogous manner to Nasr et al. (2017), where authors report the linkage of a His-tagged PV receptor (PVr, CD155) to nanodiscs via a nitrilotriacetic acid (NTA) nickel-chelating head group (Bubeck et al., 2005a; Bilek et al., 2007; Strauss et al., 2013; Lee et al., 2016; Groppelli et al., 2017).

This integrin-liposome/nanodisc system could be combined with the use of caged protons to trigger an on-grid pH drop for uncoating intermediate capture. However, given that microspectrometer results clearly indicate that, even within the small volumes of 4 μL on EM grids, pH drops from 7.5 to below 6.5 could only be achieved after relatively long exposure times (above 4 s), a different system might be required for controlled proton release, or a more powerful excitation source. There have been numerous recent reports of alternative ways of time-resolved cryo-EM, such as on-grid mixing of acid and virus sample using microfluidics as an alternative solution to controlling uncoating events (Feng et al., 2017; Kaledhonkar et al., 2018).

Provisional tests of additive stabilisation of FMDV with BSA/$CaCl_2$ indicate that this could be a

potential means of stabilising FMDV uncoating intermediates. This may be unwise, given the fact that calcium ions would likely tip the equilibrium away from the active integrin state that engages with FMDV capsids. For certain, however, additive stabilisation of VLP samples (see Chapter 5) should be investigated.

Although at low resolution, our empty ERAV structure certainly supports the idea that a metastable expanded form of aphthoviruses can exist. Our structure is, however, vastly different from the empty ERAV structures described in Bakker et al. (2014), with a global movement of pentameric units 'outwards' relative to the native virion and the reason this expanded state is unclear. The buffers used for purification were the same as had been used for previous ERAV preparations (Tuthill et al., 2009), but the empty particle sample were stored in a ca. 30 % sucrose/PBS buffer at 4 °C for a month at the Pirbright Institute before shipment, on dry ice, to Oxford. This long storage time could, therefore, cause empty particles to expand (as may be the case for FMDV VLPs, Chapter 5). Alternatively, it is conceivable that the membrane of the Amicon Ultra Centrifugal Filter (Merck) used to concentrate the sample caused expansion, as was also hypothesised to be the cause for PV VLP antigenic switching in Chapter 6.

Chapter 8

Conclusions and Future Directions

Picornaviral infections are a global threat to human and animal health and to effectively control an infection outbreak, a two-pronged approach is needed; attack (anti-virals, either small molecule or a biological) and defence (vaccines). To rationally design tools to fight infection, it is important to understand infectious agents and their intermediate forms at a molecular level. Structures of picornaviruses are now rather numerous (498 in the Protein Data Bank as of 29th September 2019), weaving a rich tapestry of members of the picornavirus family. However, we still understand little of the picornaviral infection process.

Enteroviruses are a major cause of human disease globally, with infections associated with an array of symptoms ranging from the benign to life-threatening in severity. It is often convenient to use a less pathogenic surrogate for structural study, such as when searching for therapeutics, or probing function using chimeric enterovirus capsids (Royston et al., 2018). I have characterised EV-F3, a surrogate for more pathogenic enteroviruses, describing mature, natural empty, and expanded uncoating intermediates. Structures of native and natural empty EV-F3 were solved to sub-3 Å resolution using room-temperature crystallography and have been shown to share similarities with its closest relative, EV-E1, excepting generally more truncated surface loops, supporting the relatively recent classification of these two bovine enteroviruses as distinct species.

Anomalous densities from data collected close to the potassium and calcium K-edges support the presence of four potassium ions per icosahedral asymmetric unit, located around the 2-fold symmetry axes within EV-F3's capsid interior. Extension of analyses to other enteroviruses to see whether we see any trends in preferences for bound metal ions would be informative, albeit with

complicating considerations for viruses requiring containment. CVA-24v and HRV16, for instance, would be interesting targets, especially since metal ions have been modelled for both structures (Oliveira et al., 1993; Zocher et al., 2014).

Natural empty EV-F3 particles, solved to 2.7 Å resolution, are shown to be almost structurally identical to their mature counterpart. It may be informative to repeat analyses of these natural empty particles using cryo-EM to see if we can resolve the missing section of polypeptide chain (4070-4071; 2001-2008) surrounding the scissile bond. By contrast, 'natural' ERAV empty particles, determined to a reported 9.7 Å resolution, are shown to be expanded, perhaps due to destabilisation by the cellulose membrane (Amicon concentrator device) to concentrate sample. Additional data need to be collected to resolve the exact basis for this expansion. Further refinement should be performed on the ERAV model, with rigid body refinement of protomer domains to complete the analysis of the current structure.

I present three cryo-EM structures of heat induced EV-F3 conformational switching. These, at resolutions better than 4.3 Å, show characteristics consistent with other expanded enterovirus A- and B-particles and an intermediate 'I' state showing characteristics of both. All particles appear to be expanded by ca. 4 %, and show the characteristic widening at the two-fold axis and changes in the VP3 GH-loop. Before publication, additional calibrations should be performed of the magnification to be absolutely sure of the extent of expansion, and more modelling performed to carefully tease out differences between A, I and B particles. Additional processing might also be performed, for instance symmetry relaxation approaches, whereby particles are oriented using icosahedral symmetry before relaxation to C1 symmetry to identify features that do not adhere to icosahedral symmetry, such as have been applied to the CV-B3 A-particle captured in nanodiscs (Lee et al., 2016). It might also be interesting to do a more robust analysis of EV-F3's thermo-induced uncoating intermediates, such as a shorter time point to see if we can enrich the population of A-particles, and a prolonged heating cycle to see if we can further break-down the 'B' particles into pentamers.

Provisional thermofluor results indicate that Mg^{2+} and Ca^{2+} reduce the temperature at which signal from SYTO9 nucleic acid dye by 1 °C, relative to untreated virions (confidence limits at less than 0.4 °), and should be followed up with sucrose density gradient and/or microscopy analysis of virus sample heated with Mg^{2+} or Ca^{2+}, to establish whether and how metal ions impact on

capsid stability.

I have initiated several lines of enquiry to move towards understanding how FMDV's long (even in the most compact form, FMDV's genome is over 2 μm in length), tightly packed (packing density in the capsid is of the order of 600 mg/mL), single-stranded RNA genome is delivered into the cytoplasm of a cell to initiate infection. To this end, I have described efforts to create a model system for FMDV infection, which would allow potentially the study of receptor attachment through to uncoating and transmission of the genome through a membrane. My attempts to express full length $\alpha_v\beta_6$ integrin have shown that over-expression of native constructs is highly cytotoxic, but may be achieved through an inducible cell line. Head-group $\alpha_v\beta_3$ integrin can be successfully expressed using existing protocols and I show, at very low resolution (25 Å reported resolution) that semi-pure $\alpha_v\beta_3$ can bind a small population of FMDV A22cc VLPs. The next step will be to establish a stable cell line to produce a large quantity of $\alpha_v\beta_3$. Once integrin can be reliably produced in large quantities, the intention is to move towards creating a model system of FMDV infection, investigating attachment of $\alpha_v\beta_3$ to nanodiscs to mimic a phospholipid membrane, as described in, for instance, Strauss et al. (2013); Lee et al. (2016); Nasr et al. (2017). Receptor-containing nanodiscs could then be incubated with FMDV and visualised with cryo-electron tomography under different pH conditions to understand the two distinct FMDV cellular entry processes of receptor attachment and uncoating. Ultimately, if a pH drop can be synchronised using a laser (see below), the goal would be to follow RNA transport through the nanodisc following laser-triggered uncoating.

I have also demonstrated the potential for caged protons to enable a pH drop within a small sample volume, but limited by the low power of the LED used. A much more powerful uncaging source, such as the PORTO laser (movable femtosecond laser, tuneable to wavelengths ranging from UV to near-IR) recently installed at Diamond Light Source (Harwell, UK) could be used to induce the pH drop in a controlled fashion and an investigation is planned. If successful, time-course experiments using SAXS and cryo-EM could be performed. Briefly, a series of SAXS measurements at different time coordinates could be taken of acid or EDTA-treated particles. Structures of acid or EDTA-treated FMDV could then be characterised using cryo-EM, and SAXS scattering profiles determined from different structures and compared (Aramayo et al., 2005). The cryo-EM structures could then be sorted along a time coordinate through comparison of the predicated

and measured SAXS scattering profiles. It would be also informative to apply such methods to corroborate findings that support FMDV's capsid breathing (see Chapter 1, Section 1.1.3), where broadly cross-reactive monoclonal antibodies were shown to bind to FMDV's usually internal, VP2 N-terminus and VP4 (personal communication, Toby Tuthill), possibly by applying gentle heat to speed up the breathing process (Lin et al., 2012; Panjwani et al., 2014; Roivainen et al., 1993; Pulli et al., 1998) and capturing complexes with antibodies or Fabs.

Existing antivirals against enteroviruses tend to focus on the P-site, but, due to the RNA polymerase inaccuracy, resistant mutants often rapidly develop, such that, for robust therapy (by analogy with multi-drug therapy for HIV, (Egger et al., 2002) the development of antivirals targeting novel sites would be useful. Receptor binding sites are such potential targets. Structures of EV-F3 bound to 3' and 6'-linked sialyllactoses solved to 1.92 and 2.32 Å resolutions, respectively, show a binding epitope similar to enteroviruses such as EV-A71. A competition-soaked structure to 1.92 Å resolution indicates a higher affinity for 3' linkages. Preliminary plaque assays suggest a dependence on sialic acid on entry, but these should be repeated and additional tests, such as pre-incubation with sialylactoses should be performed to confirm this. Collectively, these findings suggest that EV-F3 may utilise sialylated sugars during the entry process, but, whether this is a consequence of tissue culture adaptation, and whether sialylated sugars are the primary receptor, remain unknown.

Crystallographic structures of GSH-bound to 1.69 Å, CG-bound to 1.61 Å and a competition soak to 2.32 Å resolution demonstrate the existence of a site at the inter-protomer surface (in this thesis, termed the 'G-site') to which GSH preferentially binds relative to its degradation product CG, and confers a concentration-dependent stabilising effect. Thermofluor and infectivity assays indicate that EV-F3 is not only dependent on glutathione for infectivity, but that glutathione also confers a stabilising effect, making this a potential druggable site with glutathione providing a start point for drug development. Analyses should be extended in future to other enteroviruses to see if this site is well conserved, and, more importantly, be used for structure-based design for antivirals. Shorter term, it would be interesting to do competition binding experiments with 'Compound 17' to see if glutathione has a higher affinity (Abdelnabi et al., 2019).

Most existing picornaviral vaccine preparations comprise live or attenuated virus, demanding costly containment facilities and mis-preparations can sometimes cause infection. Preparations

are also often unstable and, in the case of FMDV vaccines, require a costly cold-chain, making administration difficult, particularly in remote and warm equatorial regions. VLPs offer safer, cheaper and customisable vaccine candidates, but the inherent instability of natural empty particles necessitates rational engineering to stabilise VLPs against pH, heat and local ionic strength. Stabilisation may also be achieved by using additives, and, our findings for GSH and more potent additives should be explored as stabilisers for vaccine preparations. Alternatives could also be explored; certain enteroviruses (CV-B3, PV), Kobuviruses (Aichi) and Cardioviruses (mengovirus) appear to be stabilised by bacteria (Aguilera et al., 2019).

By taking advantage of the potential for resolving heterogeneity during cryo-EM analysis, FMDV and yeast-derived PV-1 Mahoney and PV-3 Saukett with thermostabilising mutations (SC6b, SC8) have been reported showing different states of expansion. For FMDV VLPs, 45 % of picked AsiaShcc particles showed density supporting a S-S bridge, (solved to 2.9-3.5 Å resolution, three structures). DTT treatment greatly disrupts particles, but, with milder 1 mM treatments, we show that the disulphide bridge can remain intact (60 % of picked particles, 3.1 Å reported resolution). At high DTT concentrations (10 mM), capsids fall appart into pentamers, from which a 3.9 Å cryo-EM reconstruction was determined. This medium-resolution pentamer structure shows the switch in antigenicity from an assembled VLP, accounting for why degraded vaccine preparations are ineffective. Anomalous crystallographic data indicate that calcium ions do not seem to bind to VLP capsids, even following soaks with 2 mM $CaCl_2$.

PV-1 Mahoney and PV-3 Saukett with thermostabilising mutations (SC6b and SC8) yielded two reconstructions each of 5.8 Å, 6.8 Å and 3.4 Å, 4.3 Å reported resolutions. Reconstructions showed that, for PV-1, only 30 % of PV-1 SC6b particles show D-antigenicity. By contrast, for PV-3 SC8, both reconstructions showed characteristics consistent with D antigenic types. These ratios do not match those estimated by ELISA, for reasons that are not yet clear.

All VLP cryo-EM structures need to be further analysed to truly rationalise the motions that take place following DTT treatment, and, the results of these analyses could inspire new potential thermostabilising mutants of FMDV and PV VLPs. Part of this further study should involve reprocessing of the existing cryo-EM data with symmetry relaxation to see whether we can resolve any additional changes in capsid architecture that are diluted during icosahedral averaging.

Chapter 9

Publications

A copy of published papers listed below may be found in the Appendix.

1. Duyvesteyn, H.M., Ginn, H.M., Pietilä, M.K., Wagner, A., Hattne, J., Grimes, J.M., Hirvonen, E., Evans, G., Parsy, M.L., Sauter, N.K. and Brewster, A.S., 2018. Towards in cellulo virus crystallography. Scientific reports, 8(1), p.3771.

2. Duyvesteyn, H.M., Kotecha, A., Ginn, H.M., Hecksel, C.W., Beale, E.V., de Haas, F., Evans, G., Zhang, P., Chiu, W. and Stuart, D.I., 2018. Machining protein microcrystals for structure determination by electron diffraction. Proceedings of the National Academy of Sciences, 115(38), pp.9569-9573.

3. Duyvesteyn, H.M., Ren, J., Walter, T.S., Fry, E.E. and Stuart, D.I., 2020. Glutathione facilitates enterovirus assembly by binding at a druggable pocket. Communications biology, 3(1), pp.1-7.

4. Watanabe, Y., Vasiljevic, S., Allen, J.D., Seabright, G.E., Duyvesteyn, H.M., Doores, K.J., Crispin, M. and Struwe, W.B., 2018. Signature of antibody domain exchange by native mass spectrometry and collision-induced unfolding. Analytical chemistry, 90(12), pp.7325-7331.

5. Ebrahim, A., Moreno-Chicano, T., Appleby, M.V., Chaplin, A.K., Beale, J.H., Sherrell, D.A., Duyvesteyn, H.M., Owada, S., Tono, K., Sugimoto, H. and Strange, R.W., 2019. Dose-resolved serial synchrotron and XFEL structures of radiation-sensitive metalloproteins. IUCrJ, 6(4).

6. Moreno-Chicano, T., Ebrahim, A., Axford, D., Appleby, M.V., Beale, J.H., Chaplin, A.K., Duyvesteyn, H.M., Ghiladi, R.A., Owada, S., Sherrell, D.A. and Strange, R.W., 2019. High-throughput structures of protein-ligand complexes at room temperature using serial femtosecond crystallography. IUCrJ, 6(6).

Bibliography

Aarthi, D., Rao, K. A., Robinson, R. and Srinivasan, V. (2004), 'Validation of binary ethyleneimine (bei) used as an inactivant for foot and mouth disease tissue culture vaccine', *Biologicals* **32**(3), 153–156.

Abdel-Moneim, A., El-Senousy, W. M., Abdel-Latif, M. and Khalil, R. G. (2018), 'Association between antioxidant enzyme activities and enterovirus-infected type 1 diabetic children', *Medical Principles and Practice* **27**(1), 86–91.

Abdel-Rahman, S. M. and Kearns, G. L. (1998), 'Single-dose pharmacokinetics of a pleconaril (vp63843) oral solution and effect of food', *Antimicrobial agents and chemotherapy* **42**(10), 2706–2709.

Abdelnabi, R., Delang, L. and Neyts, J. (2017), 'Glutathione is a highly efficient thermostabilizer of poliovirus sabin strains', *Vaccine* **35**(10), 1370–1372.

Abdelnabi, R., Geraets, J. A., Ma, Y. et al. (2019), 'A novel druggable interprotomer pocket in the capsid of rhino-and enteroviruses', *PLoS Biology* **17**(6), e3000281.

Acharya, R., Fry, E., Stuart, D. et al. (1989), 'The three-dimensional structure of foot-and-mouth disease virus at 2.9 a resolution', *Nature* **337**(6209), 709–716.

Adams, P. D., Afonine, P. V., Bunkóczi, G. et al. (2010), 'Phenix: a comprehensive python-based system for macromolecular structure solution', *Acta Crystallographica Section D: Biological Crystallography* **66**(2), 213–221.

Adams, P. D., Afonine, P. V., Bunkóczi, G. et al. (2011), 'The phenix software for automated determination of macromolecular structures', *Methods* **55**(1), 94–106.

Adams, S. R. and Tsien, R. Y. (1993), 'Controlling cell chemistry with caged compounds', *Annual review of physiology* **55**(1), 755–784.

Adeyemi, O. O., Nicol, C., Stonehouse, N. J. and Rowlands, D. J. (2017), 'Increasing type 1 poliovirus capsid stability by thermal selection', *Journal of virology* **91**(4), e01586–16.

Adrian, M., Dubochet, J., Lepault, J. and McDowall, A. W. (1984), 'Cryo-electron microscopy of viruses', *Nature* **308**(5954), 32.

Adxv – A Program to Display X-ray Diffraction Images, http://www.scripps.edu/ arvai/adxv.html (2013).

Aebi, U. and Pollard, T. D. (1987), 'A glow discharge unit to render electron microscope grids and other surfaces hydrophilic', *Journal of electron microscopy technique* **7**(1), 29–33.

Aggarwal, N. and Barnett, P. (2002), 'Antigenic sites of foot-and-mouth disease virus (fmdv): an analysis of the specificities of anti-fmdv antibodies after vaccination of naturally susceptible host species', *Journal of General Virology* **83**(4), 775–782.

Agol, V. I. (1991), The 5'-untranslated region of picornaviral genomes, *in* 'Advances in virus research', Vol. 40, Elsevier, pp. 103–180.

Aguilera, E. R., Nguyen, Y., Sasaki, J. and Pfeiffer, J. K. (2019), 'Bacterial stabilization of a panel of picornaviruses', *mSphere* **4**(2), e00183–19.

Airaksinen, A. (2000), 'The vp1 intracapsid hook and uncoating of enteroviruses'.

Airaksinen, A., Roivainen, M. and Hovi, T. (2001), 'Coxsackievirus a9 vp1 mutants with enhanced or hindered a particle formation and decreased infectivity', *Journal of virology* **75**(2), 952–960.

Airaksinen, A., Roivainen, M., Stanway, G. and Hovi, T. (1999), 'Site-saturation mutagenesis of the paltavetg motif in coxsackievirus a9 capsid protein vp1 reveals evidence of conservation of a periodic hydrophobicity profile', *Journal of general virology* **80**(8), 1919–1927.

Alexandersen, S., Zhang, Z. and Donaldson, A. I. (2002), 'Aspects of the persistence of foot-and-mouth disease virus in animals—the carrier problem', *Microbes and infection* **4**(10), 1099–1110.

Alexandrov, A. I., Mileni, M., Chien, E. Y., Hanson, M. A. and Stevens, R. C. (2008), 'Microscale fluorescent thermal stability assay for membrane proteins', *Structure* **16**(3), 351–359.

Anderson, K. and Bond, C. W. (1987), 'Biological properties of mengovirus: characterization of avirulent, hemagglutination-defective mutants', *Archives of virology* **93**(1-2), 31–49.

Andréll, J. and Tate, C. G. (2013), 'Overexpression of membrane proteins in mammalian cells for structural studies', *Molecular membrane biology* **30**(1), 52–63.

Ansardi, D. C., Porter, D. C., Anderson, M. J. and Morrow, C. D. (1996), 'Poliovirus assembly and encapsidation of genomic rna', *Advances in Virus Research* pp. 1–68.

Aramayo, R., Merigoux, C., Larquet, E. et al. (2005), 'Divalent ion-dependent swelling of tomato bushy stunt virus: a multi-approach study', *Biochimica et Biophysica Acta (BBA)-General Subjects* **1724**(3), 345–354.

Aricescu, A. R., Lu, W. and Jones, E. Y. (2006), 'A time-and cost-efficient system for high-level protein production in mammalian cells', *Acta Crystallographica Section D: Biological Crystallography* **62**(10), 1243–1250.

Arita, M., Koike, S., Aoki, J., Horie, H. and Nomoto, A. (1998), 'Interaction of poliovirus with its purified receptor and conformational alteration in the virion', *Journal of virology* **72**(5), 3578–3586.

Arnold, E. and Rossmann, M. G. (1990), '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**(4), 763–801.

Arnold, S. A., Albiez, S., Bieri, A. et al. (2017), 'Blotting-free and lossless cryo-electron microscopy grid preparation from nanoliter-sized protein samples and single-cell extracts', *Journal of structural biology* **197**(3), 220–226.

Arnold, S. A., Albiez, S., Opara, N. et al. (2016), 'Total sample conditioning and preparation of nanoliter volumes for electron microscopy', *ACS nano* **10**(5), 4981–4988.

Asfor, A. S., Upadhyaya, S., Knowles, N. J. et al. (2014), 'Novel antibody binding determinants on the capsid surface of serotype o foot-and-mouth disease virus', *The Journal of general virology* **95**(Pt 5), 1104.

Auffinger, P., Bielecki, L. and Westhof, E. (2003), 'The mg2+ binding sites of the 5s rrna loop e motif as investigated by molecular dynamics simulations', *Chemistry & biology* **10**(6), 551–561.

Bachmann, M. F. and Jennings, G. T. (2010), 'Vaccine delivery: a matter of size, geometry, kinetics and molecular patterns', *Nature Reviews Immunology* **10**(11), 787.

Bakker, S. E., Groppelli, E., Pearson, A. R. et al. (2014), 'Limits of structural plasticity in a picornavirus capsid revealed by a massively expanded equine rhinitis a virus particle', *Journal of virology* **88**(11), 6093–6099.

Bandyopadhyay, A. and Raghavan, S. (2009), 'Defining the role of integrin $\alpha v\beta 6$ in cancer', *Current drug targets* **10**(7), 645–652.

Bansal, A. and Simon, M. C. (2018), 'Glutathione metabolism in cancer progression and treatment resistance', *The Journal of cell biology* **217**(7), 2291–2298.

Barczyk, M., Carracedo, S. and Gullberg, D. (2010), 'Integrins', *Cell and tissue research* **339**(1), 269.

Barrera, J., Brake, D. A., Schutta, C. et al. (2018), 'Versatility of the adenovirus-vectored foot-and-mouth disease vaccine platform across multiple foot-and-mouth disease virus serotypes and topotypes using a vaccine dose representative of the adta24 conditionally licensed vaccine', *Vaccine* **36**(48), 7345–7352.

Basavappa, R., Filman, D., Syed, R. et al. (1994), '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**(10), 1651–1669.

Bausch-Fluck, D., Hofmann, A., Bock, T. et al. (2015), 'A mass spectrometric-derived cell surface protein atlas', *PloS one* **10**(4), e0121314.

Baxt, B. and Bachrach, H. L. (1980), 'Early interactions of foot-and-mouth disease virus with cultured cells', *Virology* **104**(1), 42–55.

Baxt, B., Vakharia, V., Moore, D., Franke, A. and Morgan, D. (1989), 'Analysis of neutralizing antigenic sites on the surface of type a12 foot-and-mouth disease virus.', *Journal of virology* **63**(5), 2143–2151.

Bayer, N., Prchla, E., Schwab, M., Blaas, D. and Fuchs, R. (1999), 'Human rhinovirus hrv14 uncoats from early endosomes in the presence of bafilomycin', *Febs Letters* **463**(1-2), 175–178.

Bażanów, B., Frącka, A., Jackulak, N. et al. (2018), 'Viral, serological, and antioxidant investigations of equine rhinitis a virus in serum and nasal swabs of commercially used horses in poland', *BioMed Research International* **2018**.

Bellis, S. L. (2004), 'Variant glycosylation: an underappreciated regulatory mechanism for β1 integrins', *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1663**(1), 52–60.

Belnap, D. M., Filman, D. J., Trus, B. L. et al. (2000), 'Molecular tectonic model of virus structural transitions: the putative cell entry states of poliovirus', *Journal of virology* **74**(3), 1342–1354.

Benschop, K., van der Avoort, H., Duizer, E. and Koopmans, M. (2015), 'Antivirals against enteroviruses: a critical review from a public-health perspective', *Antivir Ther* **20**(2), 121–30.

Berman, A., Kozlova, N. and Morozevich, G. (2003), 'Integrins: structure and signaling', *Biochemistry (Moscow)* **68**(12), 1284–1299.

Berryman, S., Clark, S., Monaghan, P. and Jackson, T. (2005), 'Early events in integrin αvβ6-mediated cell entry of foot-and-mouth disease virus', *Journal of virology* **79**(13), 8519–8534.

Bhat, S., Saravanan, P., Hosamani, M. et al. (2013), 'Novel immunogenic baculovirus expressed virus-like particles of foot-and-mouth disease (fmd) virus protect guinea pigs against challenge', *Research in veterinary science* **95**(3), 1217–1223.

Bilek, G., Kremser, L., Wruss, J., Blaas, D. and Kenndler, E. (2007), 'Mimicking early events of virus infection: capillary electrophoretic analysis of virus attachment to receptor-decorated liposomes', *Analytical chemistry* **79**(4), 1620–1625.

Bostina, M., Levy, H., Filman, D. J. and Hogle, J. M. (2011), 'Poliovirus rna is released from the capsid near a twofold symmetry axis', *Journal of virology* **85**(2), 776–783.

Brandenburg, B., Lee, L. Y., Lakadamyali, M. et al. (2007), 'Imaging poliovirus entry in live cells', *PLoS biology* **5**(7), e183.

Brennan, S. and Cowan, P. (1992), 'A suite of programs for calculating x-ray absorption, reflection, and diffraction performance for a variety of materials at arbitrary wavelengths', *Review of scientific instruments* **63**(1), 850–853.

Brittain, H. G., Richardson, F. S. and Martin, R. B. (1976), 'Terbium (iii) emission as a probe of calcium (ii) binding sites in proteins', *Journal of the American Chemical Society* **98**(25), 8255–8260.

Brooksby, J. (1982), 'Portraits of viruses: foot-and-mouth disease virus', *Intervirology* **18**(1-2), 1–23.

Brown, F. and Cartwright, B. (1961), 'Dissociation of foot-and-mouth disease virus into its nucleic acid and protein components', *Nature* **192**(4808), 1163.

Brune, K. D., Leneghan, D. B., Brian, I. J. et al. (2016), 'Plug-and-display: decoration of virus-like particles via isopeptide bonds for modular immunization', *Scientific reports* **6**, 19234.

Brunger, A. T. (2007), 'Version 1.2 of the crystallography and nmr system', *Nature protocols* **2**(11), 2728–2733. LR: 20080324; JID: 101284307; ppublish.

Bubeck, D., Filman, D. J., Cheng, N. et al. (2005b), '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**(12), 7745–7755.

Bubeck, D., Filman, D. J. and Hogle, J. M. (2005a), 'Cryo-electron microscopy reconstruction of a poliovirus-receptor-membrane complex', *Nature structural & molecular biology* **12**(7), 615–618.

Burman, A., Clark, S., Abrescia, N. G. et al. (2006), 'Specificity of the vp1 gh loop of foot-and-mouth disease virus for αv integrins', *Journal of virology* **80**(19), 9798–9810.

Busk, M., Pytela, R. and Sheppard, D. (1992), 'Characterization of the integrin alpha v beta 6 as a fibronectin-binding protein.', *Journal of Biological Chemistry* **267**(9), 5790–5796.

Butan, C., Filman, D. J. and Hogle, J. M. (2014), 'Cryo-electron microscopy reconstruction shows poliovirus 135s particles poised for membrane interaction and rna release', *Journal of virology* **88**(3), 1758–1770.

Cai, J., Chen, Y., Seth, S. et al. (2003), 'Inhibition of influenza infection by glutathione', *Free Radical Biology and Medicine* **34**(7), 928–936.

Cao, Y., Lu, Z. and Liu, Z. (2016), 'Foot-and-mouth disease vaccines: progress and problems', *Expert review of vaccines* **15**(6), 783–789.

Cao, Y., Lu, Z., Sun, J. et al. (2009), 'Synthesis of empty capsid-like particles of asia i foot-and-mouth disease virus in insect cells and their immunogenicity in guinea pigs', *Veterinary microbiology* **137**(1-2), 10–17.

Cao, Z., Ding, Y., Cao, L. et al. (2017), 'Isochlorogenic acid c prevents enterovirus 71 infection via modulating redox homeostasis of glutathione', *Scientific reports* **7**(1), 16278.

Caridi, F., Cañas-Arranz, R., Vázquez-Calvo, A., Sobrino, F. and Martín-Acebes, M. A. (2016), 'Equine rhinitis a virus mutants with altered acid resistance unveil a key role of vp3 and intra-subunit interactions in the control of the ph stability of aphthovirus capsid', *Journal of virology* pp. JVI–01043.

Carrillo, C., Tulman, E., Delhon, G. et al. (2005), 'Comparative genomics of foot-and-mouth disease virus', *Journal of virology* **79**(10), 6487–6504.

Carrillo-Tripp, M., Shepherd, C. M., Borelli, I. A. et al. (2008), 'Viperdb2: an enhanced and web api enabled relational database for structural virology', *Nucleic acids research* **37**(suppl_1), D436–D442.

Carson, S. D., Hafenstein, S. and Lee, H. (2017), 'Mops and coxsackievirus b3 stability', *Virology* **501**, 183–187.

Cartwright, B., Chapman, W. and Brown, F. (1980), 'Serological and immunological relationships between the 146s and 12s particles of foot-and-mouth disease virus', *Journal of General Virology* **50**(2), 369–375.

Cartwright, B., Chapman, W. and Sharpe, R. (1982), 'Stimulation by heterotypic antigens of foot-and-mouth disease virus antibodies in vaccinated cattle', *Research in veterinary science* **32**(3), 338–342.

Casasnovas, J. M. and Springer, T. A. (1994), 'Pathway of rhinovirus disruption by soluble intercellular adhesion molecule 1 (icam-1): an intermediate in which icam-1 is bound and rna is released.', *Journal of virology* **68**(9), 5882–5889.

Chaudhary, S., Pak, J. E., Gruswitz, F., Sharma, V. and Stroud, R. M. (2012), 'Overexpressing human membrane proteins in stably transfected and clonal human embryonic kidney 293s cells', *Nature protocols* **7**(3), 453.

Cheah, K.-C., Leong, L. and Porter, A. G. (1990), 'Site-directed mutagenesis suggests close functional relationship between a human rhinovirus 3c cysteine protease and cellular trypsin-like serine proteases.', *Journal of Biological Chemistry* **265**(13), 7180–7187.

Chen, B., Kaledhonkar, S., Sun, M. et al. (2015), 'Structural dynamics of ribosome subunit association studied by mixing-spraying time-resolved cryogenic electron microscopy', *Structure* **23**(6), 1097–1105.

Cheng, M.-L., Weng, S.-F., Kuo, C.-H. and Ho, H.-Y. (2014), 'Enterovirus 71 induces mitochondrial reactive oxygen species generation that is required for efficient replication', *PLoS One* **9**(11), e113234.

Cho, H.-J., Hyun, J.-K., Kim, J.-G. et al. (2013), 'Measurement of ice thickness on vitreous ice embedded cryo-em grids: investigation of optimizing condition for visualizing macromolecules', *Journal of Analytical Science and Technology* **4**(1), 7.

Chow, M., Newman, J., Filman, D. et al. (1987), 'Myristylation of picornavirus capsid protein vp4 and its structural significance', *Nature* **327**(6122), 482–486.

Chung, Y.-C., Huang, J.-H., Lai, C.-W. et al. (2006), 'Expression, purification and characterization of enterovirus-71 virus-like particles', *World journal of gastroenterology: WJG* **12**(6), 921.

Cifuente, J. O., Lee, H., Yoder, J. D. et al. (2013), 'Structures of the procapsid and mature virion of enterovirus 71 strain 1095', *Journal of virology* **87**(13), 7637–7645.

Cifuente, J. O. and Moratorio, G. (2019), 'Evolutionary and structural overview of human picornavirus capsid antibody evasion', *Frontiers in Cellular and Infection Microbiology* **9**.

Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. and Sayers, E. W. (2015), 'Genbank', *Nucleic acids research* **44**(D1), D67–D72.

Clarke, B., Brown, A., Currey, K. et al. (1987), 'Potential secondary and tertiary structure in the genomic rna of foot and mouth disease virus', *Nucleic acids research* **15**(17), 7067–7079.

Conesa, M., Annik, P., MARVALDI, J. et al. (2003), 'Down-regulation of alphav/beta3 integrin via misrouting to lysosomes by overexpression of a beta3lamp1 fusion protein', *Biochemical Journal* **370**(2), 703–711.

Conn, G. L., Gittis, A. G., Lattman, E. E., Misra, V. K. and Draper, D. E. (2002), 'A compact rna tertiary structure contains a buried backbone–k+ complex', *Journal of molecular biology* **318**(4), 963–973.

Consortium, U. (2018), 'Uniprot: a worldwide hub of protein knowledge', *Nucleic acids research* **47**(D1), D506–D515.

Cottam, E. M., Wadsworth, J., Shaw, A. E. et al. (2008), 'Transmission pathways of foot-and-mouth disease virus in the united kingdom in 2007', *PLoS pathogens* **4**(4), e1000050.

Coyne, C. B., Kim, K. S. and Bergelson, J. M. (2007), 'Poliovirus entry into human brain microvascular cells requires receptor-induced activation of shp-2', *The EMBO journal* **26**(17), 4016–4028.

Croset, A., Delafosse, L., Gaudry, J.-P. et al. (2012), 'Differences in the glycosylation of recombinant proteins expressed in hek and cho cells', *Journal of biotechnology* **161**(3), 336–348.

Crotty, S., Maag, D., Arnold, J. J. et al. (2000), 'The broad-spectrum antiviral ribonucleoside ribavirin is an rna virus mutagen', *Nature medicine* **6**(12), 1375.

Crowther, J., Farias, S., Carpenter, W. and Samuel, A. (1993), 'Identification of a fifth neutralizable site on type o foot-and-mouth disease virus following characterization of single and quintuple monoclonal antibody escape mutants', *Journal of General Virology* **74**(8), 1547–1553.

Curry, S., Abrams, C. C., Fry, E. et al. (1995), 'Viral rna modulates the acid sensitivity of foot-and-mouth disease virus capsids.', *Journal of virology* **69**(1), 430–438.

Curry, S., Fry, E., Blakemore, W. et al. (1996), 'Perturbations in the surface structure of a22 iraq foot-and-mouth disease virus accompanying coupled changes in host cell specificity and antigenicity', *Structure* **4**(2), 135–145.

Curry, S., Fry, E., Blakemore, W. et al. (1997), '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**(12), 9743–9752.

Danev, R., Yanagisawa, H. and Kikkawa, M. (2019), 'Cryo-electron microscopy methodology: Current aspects and future directions', *Trends in biochemical sciences* .

D'Argenio, D. A. and Wilson, C. B. (2010), 'A decade of vaccines: Integrating immunology and vaccinology for rational vaccine design', *Immunity* **33**(4), 437–440.

De Palma, A. M., Vliegen, I., De Clercq, E. and Neyts, J. (2008), 'Selective inhibitors of picornavirus replication', *Medicinal research reviews* **28**(6), 823–884.

de Sena, J. and Mandel, B. (1977), 'Studies on the in vitro uncoating of poliovirus ii. characteristics of the membrane-modified particle', *Virology* **78**(2), 554–566.

Delageniere, S., Brenchereau, P., Launer, L. et al. (2011), 'Ispyb: an information management system for synchrotron macromolecular crystallography', *Bioinformatics (Oxford, England)* **27**(22), 3186–3192.

DeLano, W. L. et al. (2002), 'Pymol: An open-source molecular graphics tool', *CCP4 Newsletter on protein crystallography* **40**(1), 82–92.

DiCara, D., Burman, A., Clark, S. et al. (2008), 'Foot-and-mouth disease virus forms a highly stable, edta-resistant complex with its principal receptor, integrin $\alpha v\beta 6$: implications for infectiousness', *Journal of virology* **82**(3), 1537–1546.

Dickeson, S. and Santoro, S. (1998), 'Ligand recognition by the i domain-containing integrins', *Cellular and Molecular Life Sciences CMLS* **54**(6), 556–566.

Dill, V., Zimmer, A., Beer, M. and Eschbaumer, M. (2019), 'Investigation of cell culture conditions for optimal foot-and-mouth disease virus production', *BMC Biotechnology* **19**(1), 33.

Doel, T. and Baccarini, P. (1981), 'Thermal stability of foot-and-mouth disease virus', *Archives of virology* **70**(1), 21–32.

Doel, T. and Chong, W. (1982), 'Comparative immunogenicity of 146s, 75s and 12s particles of foot-and-mouth disease virus', *Archives of virology* **73**(2), 185–191.

Domingo, E., Pariente, N., Airaksinen, A. et al. (2005), Foot-and-mouth disease virus evolution: exploring pathways towards virus extinction, *in* 'Foot-and-Mouth Disease Virus', Springer, pp. 149–173.

Dong, X., Hudson, N. E., Lu, C. and Springer, T. A. (2014), 'Structural determinants of integrin β-subunit specificity for latent tgf-β', *Nature structural & molecular biology* **21**(12), 1091.

Dong, X., Mi, L.-Z., Zhu, J. et al. (2012), 'αvβ3 integrin crystal structures and their functional implications', *Biochemistry* **51**(44), 8814–8828.

Dong, Y., Liu, Y., Jiang, W. et al. (2017), 'Antibody-induced uncoating of human rhinovirus b14', *Proceedings of the National Academy of Sciences* **114**(30), 8017–8022.

Dorner, A., Dorner, L., Larsen, G., Wimmer, E. and Anderson, C. (1982), 'Identification of the initiation site of poliovirus polyprotein synthesis.', *Journal of virology* **42**(3), 1017–1028.

Drulyte, I., Johnson, R. M., Hesketh, E. L. et al. (2018), 'Approaches to altering particle distributions in cryo-electron microscopy sample preparation', *Acta Crystallographica Section D: Structural Biology* **74**(6).

Dunne, H., Huang, C. and Lin, W. J. (1974), 'Bovine enteroviruses in the calf: an attempt at serologic, biologic, and pathologic classification', *Journal of the American Veterinary Medical Association* .

Egger, M., May, M., Chêne, G. et al. (2002), 'Prognosis of hiv-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies', *The Lancet* **360**(9327), 119–129.

Elad, N., Bellapadrona, G., Houben, L., Sagi, I. and Elbaum, M. (2017), 'Detection of isolated protein-bound metal ions by single-particle cryo-stem', *Proceedings of the National Academy of Sciences* **114**(42), 11139–11144.

Elegheert, J., Behiels, E., Bishop, B. et al. (2018), 'Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins', *Nature protocols* **13**(12), 2991.

Ellard, F. M., Drew, J., Blakemore, W. E., Stuart, D. I. and King, A. M. (1999), 'Evidence for the role of his-142 of protein 1c in the acid-induced disassembly of foot-and-mouth disease virus capsids', *Journal of General Virology* **80**(8), 1911–1918.

Emsley, P. and Cowtan, K. (2004), 'Coot: model-building tools for molecular graphics', *Acta crystallographica.Section D, Biological crystallography* **60**(Pt 12 Pt 1), 2126–2132.

Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. (2010), 'Features and development of coot', *Acta Crystallographica Section D: Biological Crystallography* **66**(4), 486–501.

Engels, J. and Schlaeger, E. J. (1977), 'Synthesis, structure, and reactivity of adenosine cyclic 3', 5'-phosphate-benzyltriesters', *Journal of medicinal chemistry* **20**(7), 907–911.

Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N. and Nordlund, P. (2006), 'Thermofluor-based high-throughput stability optimization of proteins for structural studies', *Analytical biochemistry* **357**(2), 289–298.

Ermantraut, E., Wohlfart, K. and Tichelaar, W. (1998), 'Perforated support foils with pre-defined hole size, shape and arrangement', *Ultramicroscopy* **74**(1-2), 75–81.

Feng, Q., Chen, X., Ma, O. et al. (2003), 'Serotype and vp1 gene sequence of a foot-and-mouth disease virus from hong kong (2002)', *Biochemical and biophysical research communications* **302**(4), 715–721.

Feng, X., Fu, Z., Kaledhonkar, S. et al. (2017), 'A fast and effective microfluidic spraying-plunging method for high-resolution single-particle cryo-em', *Structure* **25**(4), 663–670.

Fenwick, M. L. and Cooper, P. D. (1962), 'Early interactions between poliovirus and erk cells: some observations on the nature and significance of the rejected particles', *Virology* **18**(2), 212–223.

Ferguson, M., Wood, D. J. and Minor, P. D. (1993), 'Antigenic structure of poliovirus in inactivated vaccines', *Journal of general virology* **74**(4), 685–690.

Ferrer-Orta, C., Ferrero, D. and Verdaguer, N. (2015), 'Rna-dependent rna polymerases of picornaviruses: from the structure to regulatory mechanisms', *Viruses* **7**(8), 4438–4460.

Filman, D. J., Syed, R., Chow, M. et al. (1989), 'Structural factors that control conformational transitions and serotype specificity in type 3 poliovirus.', *The EMBO journal* **8**(5), 1567–1579.

Filman, D., Wien, M., Cunningham, J., Bergelson, J. and Hogle, J. (1998), 'Structure determination of echovirus 1', *Acta Crystallographica Section D: Biological Crystallography* **54**(6), 1261–1272.

Fox, H., Knowlson, S., Minor, P. D. and Macadam, A. J. (2017), 'Genetically thermo-stabilised, immunogenic poliovirus empty capsids; a strategy for non-replicating vaccines', *PLoS pathogens* **13**(1), e1006117.

Frank, J. (2017), 'Time-resolved cryo-electron microscopy: Recent progress', *Journal of Structural Biology* .

Fricks, C. E. and Hogle, J. M. (1990), 'Cell-induced conformational change in poliovirus: externalization of the amino terminus of vp1 is responsible for liposome binding.', *Journal of virology* **64**(5), 1934–1945.

Fry, E., Acharya, R. and Stuart, D. (1993), 'Methods used in the structure determination of foot-and-mouth disease virus', *Acta crystallographica.Section A, Foundations of crystallography* **49**, 45–55.

Fry, E. E., Grimes, J. and Stuart, D. I. (1999a), 'Virus crystallography', *Molecular biotechnology* **12**(1), 13–23.

Fry, E. E., Knowles, N. J., Newman, J. W. et al. (2003), 'Crystal structure of swine vesicular disease virus and implications for host adaptation', *Journal of virology* **77**(9), 5475–5486.

Fry, E. E., Lea, S. M., Jackson, T. et al. (1999b), 'The structure and function of a foot-and-mouth disease virus–oligosaccharide receptor complex', *The EMBO Journal* **18**(3), 543–554.

Fry, E. E., Newman, J. W., Curry, S. et al. (2005), 'Structure of foot-and-mouth disease virus serotype a1061 alone and complexed with oligosaccharide receptor: receptor conservation in the face of antigenic variation', *Journal of general virology* **86**(7), 1909–1920.

Fry, E. E., Tuthill, T. J., Harlos, K. et al. (2010), 'Crystal structure of equine rhinitis a virus in complex with its sialic acid receptor', *Journal of General Virology* **91**(8), 1971–1977.

Garriga, D., Pickl-Herk, A., Luque, D. et al. (2012), 'Insights into minor group rhinovirus uncoating: the x-ray structure of the hrv2 empty capsid', *PLoS pathogens* **8**(1), e1002473.

Gasteiger, E., Gattiker, A., Hoogland, C. et al. (2003), 'Expasy: the proteomics server for in-depth protein knowledge and analysis', *Nucleic acids research* **31**(13), 3784–3788.

Geerlof, A., Brown, J., Coutard, B. et al. (2006), 'The impact of protein characterization in structural proteomics', *Acta Crystallographica Section D: Biological Crystallography* **62**(10), 1125–1136.

George, M., Venkataramanan, R., Gurumurthy, C. B. and Hemadri, D. (2001), 'The non-structural leader protein gene of foot-and-mouth disease virus is highly variable between serotypes', *Virus Genes* **22**(3), 271–278.

Ginn, H. M., Messerschmidt, M., Ji, X. et al. (2015), 'Structure of cpv17 polyhedrin determined by the improved analysis of serial femtosecond crystallographic data', *Nature communications* **6**, 6435.

Giranda, V. L., Heinz, B. A., Oliveira, M. A. et al. (1992), 'Acid-induced structural changes in human rhinovirus 14: possible role in uncoating', *Proceedings of the National Academy of Sciences* **89**(21), 10213–10217.

Glaeser, R. M. (2007), 'Electron crystallography of biological macromolecules'.

Glasel, J. (1995), 'Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios.', *Biotechniques* **18**(1), 62–63.

Goddard, T. D., Huang, C. C., Meng, E. C. et al. (2018), 'Ucsf chimerax: Meeting modern challenges in visualization and analysis', *Protein Science* **27**(1), 14–25.

Goodfellow, I. G., Sioofy, A. B., Powell, R. M. and Evans, D. J. (2001), 'Echoviruses bind heparan sulfate at the cell surface', *Journal of Virology* **75**(10), 4918–4921.

Gouet, P., Courcelle, E., Stuart, D. I. and M√© toz, F. (1999), 'Espript: analysis of multiple sequence alignments in postscript.', *Bioinformatics (Oxford, England)* **15**(4), 305–308.

Grant, R. A., Hiremath, C. N., Filman, D. J. et al. (1994), 'Structures of poliovirus complexes with anti-viral drugs: implications for viral stability and drug design', *Current Biology* **4**(9), 784–797.

Grant, T., Rohou, A. and Grigorieff, N. (2018), 'cistem, user-friendly software for single-particle image processing', *Elife* **7**, e35383.

Grazioli, S., Fallacara, F. and Brocchi, E. (2013), 'Mapping of antigenic sites of foot-and-mouth disease virus serotype asia 1 and relationships with sites described in other serotypes', *Journal of General Virology* **94**(3), 559–569.

Grimm, R., Typke, D., Bärmann, M. and Baumeister, W. (1996), 'Determination of the inelastic mean free path in ice by examination of tilted vesicles and automated most probable loss imaging', *Ultramicroscopy* **63**(3-4), 169–179.

Groppelli, E., Levy, H. C., Sun, E. et al. (2017), 'Picornavirus rna is protected from cleavage by ribonuclease during virion uncoating and transfer across cellular and model membranes', *PLoS pathogens* **13**(2), e1006197.

Groppelli, E., Tuthill, T. J. and Rowlands, D. J. (2010), 'Cell entry of the aphthovirus equine rhinitis a virus is dependent on endosome acidification', *Journal of virology* **84**(12), 6235–6240.

Grubman, M. J. and Baxt, B. (2004), 'Foot-and-mouth disease', *Clinical microbiology reviews* **17**(2), 465–493.

Gu, J. and Taniguchi, N. (2004), 'Regulation of integrin functions by n-glycans', *Glycoconjugate journal* **21**(1), 9–15.

Gu, W., Xu, Y., Xie, X. et al. (2014), 'The role of rna structure at 5′ untranslated region in microrna-mediated gene regulation', *Rna* **20**(9), 1369–1375.

Guan, H., Tian, J., Qin, B. et al. (2017), 'Crystal structure of 2c helicase from enterovirus 71', *Science advances* **3**(4), e1602573.

Guo, H.-C., Sun, S.-Q., Jin, Y. et al. (2013), 'Foot-and-mouth disease virus-like particles produced by a sumo fusion protein system in escherichia coli induce potent protective immune responses in guinea pigs, swine and cattle', *Veterinary research* **44**(1), 48.

Guy, M., Chilmonczyk, S., Cruciere, C., Eloit, M. and Bakkali-Kassimi, L. (2009), 'Efficient infection of buffalo rat liver-resistant cells by encephalomyocarditis virus requires binding to cell surface sialic acids', *Journal of General Virology* **90**(1), 187–196.

Hadfield, A. T., Diana, G. D. and Rossmann, M. G. (1999), 'Analysis of three structurally related antiviral compounds in complex with human rhinovirus 16', *Proceedings of the National Academy of Sciences* **96**(26), 14730–14735.

Hadfield, A. T., Lee, W.-m., Zhao, R. et al. (1997), 'The refined structure of human rhinovirus 16 at 2.15 å resolution: implications for the viral life cycle', *Structure* **5**(3), 427–441.

Hamdi, F., Tueting, C., Semchonok, D. A. et al. (2019), '2.7 å cryo-em structure of vitrified m. musculus h-chain apoferritin from 200 kev" screening microscope"', *bioRxiv* p. 738724.

Handing, K. B., Niedzialkowska, E., Shabalin, I. G. et al. (2018), 'Characterizing metal-binding sites in proteins with x-ray crystallography', *Nature protocols* **13**(5), 1062.

Harrison, S. C. (1992), 'Viruses: Current opinion in structural biology 1992, 2: 293–299', *Current Opinion in Structural Biology* **2**(2), 293–299.

Hayden, F. G., Turner, R. B., Gwaltney, J. M. et al. (2003), 'Phase ii, randomized, double-blind, placebo-controlled studies of ruprintrivir nasal spray 2-percent suspension for prevention and treatment of experimentally induced rhinovirus colds in healthy volunteers', *Antimicrobial agents and chemotherapy* **47**(12), 3907–3916.

Hazendonk, A. et al. (1981), 'D-antigen determination in polio vaccine production: comparison of gel diffusion and elisa methods.', *Developments in biological standardization* **47**, 101–108.

Heikkilä, O., Susi, P., Stanway, G. and Hyypiä, T. (2009), 'Integrin $\alpha v\beta 6$ is a high-affinity receptor for coxsackievirus a9', *Journal of General Virology* **90**(1), 197–204.

Heikkilä, O., Susi, P., Tevaluoto, T. et al. (2010), 'Internalization of coxsackievirus a9 is mediated by β2-microglobulin, dynamin, and arf6 but not by caveolin-1 or clathrin', *Journal of virology* **84**(7), 3666–3681.

Henderson, R., Sali, A., Baker, M. L. et al. (2012), 'Outcome of the first electron microscopy validation task force meeting', *Structure* **20**(2), 205–214.

Hendry, E., Hatanaka, H., Fry, E. et al. (1999), 'The crystal structure of coxsackievirus a9: new insights into the uncoating mechanisms of enteroviruses', *Structure* **7**(12), 1527–1538.

Herod, M. R., Adeyemi, O. O., Ward, J. et al. (2019), 'The broad-spectrum antiviral drug arbidol inhibits foot-and-mouth disease virus genome replication.', *The Journal of general virology* .

Herzenberg, L. A., De Rosa, S. C., Dubs, J. G. et al. (1997), 'Glutathione deficiency is associated with impaired survival in hiv disease', *Proceedings of the National Academy of Sciences* **94**(5), 1967–1972.

Hewat, E. A., Neumann, E. and Blaas, D. (2002), 'The concerted conformational changes during human rhinovirus 2 uncoating', *Molecular cell* **10**(2), 317–326.

Heymann, J. B. (2001), 'Bsoft: image and molecular processing in electron microscopy', *Journal of structural biology* **133**(2-3), 156–169.

Heymann, J. B. and Belnap, D. M. (2007), 'Bsoft: image processing and molecular modeling for electron microscopy', *Journal of structural biology* **157**(1), 3–18.

Hird, T. R. and Grassly, N. C. (2012), 'Systematic review of mucosal immunity induced by oral and inactivated poliovirus vaccines against virus shedding following oral poliovirus challenge', *PLoS pathogens* **8**(4).

Ho, H.-Y., Cheng, M.-L., Weng, S.-F. et al. (2008), 'Glucose-6-phosphate dehydrogenase deficiency enhances enterovirus 71 infection', *Journal of General Virology* **89**(9), 2080–2089.

Hogle, J., Chow, M. and Filman, D. (1985), 'Three-dimensional structure of poliovirus at 2.9 a resolution', *Science* **229**(4720), 1358–1365.

Hovi, T. and Roivainen, M. (1993), 'Peptide antisera targeted to a conserved sequence in poliovirus capsid vp1 cross-react widely with members of the genus enterovirus.', *Journal of Clinical Microbiology* **31**(5), 1083–1087.

Huang, Y., Hogle, J. M. and Chow, M. (2000), 'Is the 135s poliovirus particle an intermediate during cell entry?', *Journal of virology* **74**(18), 8757–8761.

Hull, R. (1977), 'The stabilization of the particles of turnip rosette virus and of other members of the southern bean mosaic virus group', *Virology* **79**(1), 58–66.

Hummeler, K., Hamparian, V. V. et al. (1958), 'Studies on the complement fixing antigens of poliomyelitis. i. demonstration of type and group specific antigens in native and heated viral preparations.', *Journal of Immunology* **81**(6), 499–505.

Hynes, R. O. (1987), 'Integrins: a family of cell surface receptors', *cell* **48**(4), 549–554.

Hynes, R. O. (2002), 'Integrins: bidirectional, allosteric signaling machines', *Cell* **110**(6), 673–687.

Jackson, T., Blakemore, W., Newman, J. W. et al. (2000b), 'Foot-and-mouth disease virus is a ligand for the high-affinity binding conformation of integrin $\alpha 5\beta 1$: influence of the leucine residue within the rgdl motif on selectivity of integrin binding', *Journal of General Virology* **81**(5), 1383–1391.

Jackson, T., Ellard, F. M., Ghazaleh, R. A. et al. (1996), 'Efficient infection of cells in culture by type o foot-and-mouth disease virus requires binding to cell surface heparan sulfate.', *Journal of Virology* **70**(8), 5282–5287.

Jackson, T., Sharma, A., Ghazaleh, R. A. et al. (1997), 'Arginine-glycine-aspartic acid-specific binding by foot-and-mouth disease viruses to the purified integrin alpha (v) beta3 in vitro.', *Journal of virology* **71**(11), 8357–8361.

Jackson, T., Sheppard, D., Denyer, M., Blakemore, W. and King, A. M. (2000a), 'The epithelial integrin $\alpha v\beta 6$ is a receptor for foot-and-mouth disease virus', *Journal of virology* **74**(11), 4949–4956.

Jain, T., Sheehan, P., Crum, J., Carragher, B. and Potter, C. S. (2012), 'Spotiton: a prototype for an integrated inkjet dispense and vitrification system for cryo-tem', *Journal of structural biology* **179**(1), 68–75.

Jamal, S. M. and Belsham, G. J. (2013), 'Foot-and-mouth disease: past, present and future', *Veterinary research* **44**(1), 116.

Jang, S. K., Kräusslich, H., Nicklin, M. et al. (1988), 'A segment of the 5'nontranslated region of encephalomyocarditis virus rna directs internal entry of ribosomes during in vitro translation.', *Journal of virology* **62**(8), 2636–2643.

Jiang, P., Liu, Y., Ma, H.-C., Paul, A. V. and Wimmer, E. (2014), 'Picornavirus morphogenesis', *Microbiol. Mol. Biol. Rev.* **78**(3), 418–437.

Jiménez-Clavero, M. A., Escribano-Romero, E., Mansilla, C. et al. (2005), 'Survey of bovine enterovirus in biological and environmental samples by a highly sensitive real-time reverse transcription-pcr', *Appl. Environ. Microbiol.* **71**(7), 3536–3543.

Jimenez-Clavero, M., Douglas, A., Lavery, T., Garcia-Ranea, J. and Ley, V. (2000), 'Immune recognition of swine vesicular disease virus structural proteins: novel antigenic regions that are not exposed in the capsid', *Virology* **270**(1), 76–83.

Kabsch, W. (2010*a*), 'Integration, scaling, space-group assignment and post-refinement', *Acta Crystallographica Section D: Biological Crystallography* **66**(2), 133–144.

Kabsch, W. (2010*b*), 'Xds', *Acta Crystallographica Section D: Biological Crystallography* **66**(2), 125–132.

Kaledhonkar, S., Fu, Z., White, H. and Frank, J. (2018), Time-resolved cryo-electron microscopy using a microfluidic chip, *in* 'Protein Complex Assembly', Springer, pp. 59–71.

Kaplan, C. (1963), 'The influence of some metal ions and ph on the inactivation of vaccinia virus by heat', *Microbiology* **31**(2), 311–314.

Kaplan, J. H. (1990), 'Special topic: caged compounds in cellular physiology', *Annual Review of Physiology* **52**(1), 853–855.

Kaplan, J. H., Forbush III, B. and Hoffman, J. F. (1978), 'Rapid photolytic release of adenosine 5'-triphosphate from a protected analog: utilization by the sodium: potassium pump of human red blood cell ghosts', *Biochemistry* **17**(10), 1929–1935.

Kaplan, J. and Somlyo, A. (1989), 'Flash photolysis of caged compounds: new tools for cellular physiology', *Trends in neurosciences* **12**(2), 54–59.

Kara, D. and Alkan, M. (2000), 'Determination of acidity constants of acid–base indicators by second-derivative spectrophotometry', *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **56**(14), 2753–2761.

Karlsson, G. (2001), 'Thickness measurements of lacey carbon films', *Journal of microscopy* **203**(3), 326–328.

Katoh, K., Rozewicki, J. and Yamada, K. D. (2017), 'Mafft online service: multiple sequence alignment, interactive sequence choice and visualization', *Briefings in bioinformatics* .

Katpally, U., Fu, T.-M., Freed, D. C., Casimiro, D. R. and Smith, T. J. (2009), 'Antibodies to the buried n terminus of rhinovirus vp4 exhibit cross-serotypic neutralization', *Journal of virology* **83**(14), 7040–7048.

Kawahama, S. (2007), 'Rapid photolytic release of adenosine 5'-triphosphate from a protected analogue: utilization by the na: K pump of human red blood cell ghosts', *Tanpakushitsu kakusan koso. Protein, nucleic acid, enzyme* **52**(13 Suppl), 1772.

Kemmerling, S., Arnold, S. A., Bircher, B. A. et al. (2013), 'Single-cell lysis for visual analysis by electron microscopy', *Journal of structural biology* **183**(3), 467–473.

Kidmose, R. T., Juhl, J., Nissen, P. et al. (2019), 'Namdinator–automatic molecular dynamics flexible fitting of structural models into cryo-em and crystallography experimental maps', *IUCrJ* **6**(4).

Kim, D.-S., Son, K.-Y., Koo, K.-M. et al. (2016), 'Porcine sapelovirus uses α2, 3-linked sialic acid on gd1a ganglioside as a receptor', *Journal of virology* **90**(8), 4067–4077.

Kim, S., Smith, T. J., Chapman, M. S. et al. (1989), 'Crystal structure of human rhinovirus serotype 1a (hrv1a)', *Journal of molecular biology* **210**(1), 91–111.

Kim, Y., Lovell, S., Tiew, K.-C. et al. (2012), 'Broad-spectrum antivirals against 3c or 3c-like proteases of picornaviruses, noroviruses, and coronaviruses', *Journal of virology* **86**(21), 11754–11762.

King, A. M. Q., Burman, A. J., Audonnet, J.-C. F. and Lombard, M. F. A. (2009), 'Vaccine against foot-and-mouth disease'. US Patent 7,531,182.

Kitson, J. D., McCAHON, D. and Belsham, G. J. (1990), 'Sequence analysis of monoclonal antibody resistant mutants of type o foot and mouth disease virus: evidence for the involvement of the three surface exposed capsid proteins in four antigenic sites', *Virology* **179**(1), 26–34.

Klotz, E., Doyle, R., Gross, E. and Mattson, B. (2011), 'The equilibrium constant for bromothymol blue: a general chemistry laboratory experiment using spectroscopy', *Journal of Chemical Education* **88**(5), 637–639.

Knight-Jones, T. and Rushton, J. (2013), 'The economic impacts of foot and mouth disease–what are they, how big are they and where do they occur?', *Preventive veterinary medicine* **112**(3-4), 161–173.

Knowles, N. and Barnett, I. (1985), 'A serological classification of bovine enteroviruses', *Archives of virology* **83**(3-4), 141–155.

Koiwai, A., Suzuki, F., Matsuzaki, T. and Kawashima, N. (1983), 'The fatty acid composition of seeds and leaves of nicotiana species', *Phytochemistry* **22**(6), 1409–1412.

Kontziampasis, D., Klebl, D. P., Iadanza, M. G. et al. (2019), 'A cryo-em grid preparation device for time-resolved structural studies', *bioRxiv* p. 563254.

Kotecha, A., Seago, J., Scott, K. et al. (2015), 'Structure-based energetics of protein interfaces guides foot-and-mouth disease virus vaccine design', *Nature structural & molecular biology* **22**(10), 788.

Kotecha, A., Wang, Q., Dong, X. et al. (2017), 'Rules of engagement between αvβ6 integrin and foot-and-mouth disease virus', *Nature communications* **8**, 15408.

Kotecha, A., Zhang, F., Juleff, N. et al. (2016), 'Application of the thermofluor pastry technique for improving foot-and-mouth disease virus vaccine formulation', *The Journal of general virology* **97**(7), 1557.

Kouiavskaia, D. V., Dragunsky, E. M., Liu, H.-M. et al. (2011), 'Immunological and pathogenic properties of poliovirus variants selected for resistance to antiviral drug v-073', *Antiviral therapy* **16**(7), 999.

Krissinel, E. and Henrick, K. (2007), 'Inference of macromolecular assemblies from crystalline state', *Journal of molecular biology* **372**(3), 774–797.

Kroeger, T., Frieg, B., Zhang, T. et al. (2017), 'Edta aggregates induce sypro orange-based fluorescence in thermal shift assay', *PloS one* **12**(5), e0177024.

Kubota, M., Takeuchi, K., Watanabe, S. et al. (2016), 'Trisaccharide containing α2, 3-linked sialic acid is a receptor for mumps virus', *Proceedings of the National Academy of Sciences* **113**(41), 11579–11584.

Kwaks, T. H. and Otte, A. P. (2006), 'Employing epigenetics to augment the expression of therapeutic proteins in mammalian cells', *Trends in biotechnology* **24**(3), 137–142.

Lacroix, C., Querol-Audí, J., Roche, M. et al. (2014), 'A novel benzonitrile analogue inhibits rhinovirus replication', *Journal of Antimicrobial Chemotherapy* **69**(10), 2723–2732.

Landsteiner, K. (1909), 'Ubertragung der poliomyelitis acuta auf affen', *Z Immunitatsforsch* **2**, 377–390.

Laskowski, R. A. and Swindells, M. B. (2011), 'Ligplot+: multiple ligand–protein interaction diagrams for drug discovery'.

Layne, E. (1957), '[73] spectrophotometric and turbidimetric methods for measuring proteins'.

Le Bouvier, G. (1959), 'The d□ c change in poliovirus particles', *British journal of experimental pathology* **40**(6), 605.

Le Bouvier, G. et al. (1955), 'The modification of poliovirus antigens by heat and ultraviolet light.', *Lancet* pp. 1013–16.

Lea, S., Abu-Ghazaleh, R., Blakemore, W. et al. (1995), 'Structural comparison of two strains of foot-and-mouth disease virus subtype o1 and a laboratory antigenic variant, g67', *Structure (London, England : 1993)* **3**(6), 571–580. LR: 20061115; JID: 101087697; 0 (Antibodies, Monoclonal); 0 (Antigens, Viral); ppublish.

Lea, S., Hernéndez, J., Blakemore, W. et al. (1994), 'The structure and antigenicity of a type c foot-and-mouth disease virus', *Structure* **2**(2), 123–139.

Lee, H., Shingler, K. L., Organtini, L. J. et al. (2016), 'The novel asymmetric entry intermediate of a picornavirus captured with nanodiscs', *Science advances* **2**(8), e1501929.

Lee, J.-H., von Wagner, M., Roth, W. K. et al. (1998), 'Effect of ribavirin on virus load and quasispecies distribution in patients infected with hepatitis c virus', *Journal of hepatology* **29**(1), 29–35.

Lee, S., Roh, S. H., Lee, J. et al. (2019), 'Cryo-em structures of the hsp104 protein disaggregase captured in the atp conformation', *Cell reports* **26**(1), 29–36.

Levy, H. C., Bostina, M., Filman, D. J. and Hogle, J. M. (2010), 'Catching a virus in the act of rna release: a novel poliovirus uncoating intermediate characterized by cryo-electron microscopy', *Journal of virology* **84**(9), 4426–4441.

Lewis, J. K., Bothner, B., Smith, T. J. and Siuzdak, G. (1998), 'Antiviral agent blocks breathing of the common cold virus', *Proceedings of the National Academy of Sciences* **95**(12), 6774–6778.

Ley, V., Higgins, J. and Fayer, R. (2002), 'Bovine enteroviruses as indicators of fecal contamination', *Appl. Environ. Microbiol.* **68**(7), 3455–3461.

Li, C., Wang, J. C.-Y., Taylor, M. W. and Zlotnick, A. (2012a), 'In vitro assembly of an empty picornavirus capsid follows a dodecahedral path', *Journal of virology* **86**(23), 13062–13069.

Li, F., Browning, G. F., Studdert, M. J. and Crabb, B. S. (1996), 'Equine rhinovirus 1 is more closely related to foot-and-mouth disease virus than to other picornaviruses', *Proceedings of the National Academy of Sciences* **93**(3), 990–995.

Li, H.-Y., Han, J.-F., Qin, C.-F. and Chen, R. (2013), 'Virus-like particles for enterovirus 71 produced from saccharomyces cerevisiae potently elicits protective immune responses in mice', *Vaccine* **31**(32), 3281–3287.

Li, Q., Yafal, A. G., Lee, Y., Hogle, J. and Chow, M. (1994), 'Poliovirus neutralization by antibodies to internal epitopes of vp4 and vp1 results from reversible exposure of these sequences at physiological temperature.', *Journal of virology* **68**(6), 3965–3970.

Li, Z., Yi, Y., Yin, X., Zhang, Z. and Liu, J. (2008), 'Expression of foot-and-mouth disease virus capsid proteins in silkworm-baculovirus expression system and its utilization as a subunit vaccine', *PloS one* **3**(5), e2273.

Li, Z., Yi, Y., Yin, X. et al. (2012b), 'Development of a foot-and-mouth disease virus serotype a empty capsid subunit vaccine using silkworm (bombyx mori) pupae', *PloS one* **7**(8), e43849.

Li, Z., Yin, X., Yi, Y. et al. (2011), 'Fmd subunit vaccine produced using a silkworm–baculovirus expression system: Protective efficacy against two type asia1 isolates in cattle', *Veterinary microbiology* **149**(1-2), 99–103.

Lin, J., Lee, L. Y., Roivainen, M. et al. (2012), 'Structure of the fab-labeled "breathing" state of native poliovirus', *Journal of virology* **86**(10), 5959–5962.

Lipton, H. L., Kumar, A. M., Hertzler, S. and Reddi, H. V. (2006), 'Differential usage of carbohydrate co-receptors influences cellular tropism of theiler's murine encephalomyelitis virus infection of the central nervous system', *Glycoconjugate journal* **23**(1-2), 39–49.

Liu, B., Guthridge, S., Li, S. Q. et al. (2012c), 'The end of the australia antigen? an ecological study of the impact of universal newborn hepatitis b vaccination two decades on', *Vaccine* **30**(50), 7309–7314.

Liu, H.-M., Roberts, J. A., Moore, D. et al. (2012d), 'Characterization of poliovirus variants selected for resistance to the antiviral compound v-073', *Antimicrobial agents and chemotherapy* **56**(11), 5568–5574.

Liu, H., Ma, Y., Wang, H. and Liu, Q. (2012a), 'Quantitative evaluation of the effect of the hepatitis b vaccine based on the hbsag-and anti-hbs-positive rates in the chinese population over the last 33 years', *Vaccine* **30**(23), 3483–3487.

Liu, Y., Sheng, J., Baggen, J. et al. (2015), 'Sialic acid-dependent cell entry of human enterovirus d68', *Nature communications* **6**, 8865.

Liu, Y., Sheng, J., van Vliet, A. L. et al. (2018), 'Molecular basis for the acid-initiated uncoating of human enterovirus d68', *Proceedings of the National Academy of Sciences* **115**(52), E12209–E12217.

Logan, D. (1991), Investigation of the Three Dimensional Structure of Viruses, PhD thesis, University of Oxford.

Logan, D., Abu-Ghazaleh, R., Blakemore, W. et al. (1993), 'Structure of a major immunogenic site on foot-and-mouth disease virus', *Nature* **362**(6420), 566–568.

Lonberg-Holm, K., Gosser, L. B. and Kauer, J. (1975), 'Early alteration of poliovirus in infected cells and its specific inhibition', *Journal of General Virology* **27**(3), 329–342.

Lu, Z., Bao, H., Cao, Y. et al. (2008), 'Protection of guinea pigs and swine by a recombinant adenovirus expressing o serotype of foot-and-mouth disease virus whole capsid and 3c protease', *Vaccine* **26**, G48–G53.

Lyu, K., Ding, J., Han, J.-F. et al. (2014), 'Human enterovirus 71 uncoating captured at atomic resolution', *Journal of virology* **88**(6), 3114–3126.

Lyu, K., He, Y.-L., Li, H.-Y. and Chen, R. (2015), 'Crystal structures of yeast-produced enterovirus 71 and enterovirus 71/coxsackievirus a16 chimeric virus-like particles provide the structural basis for novel vaccine design against hand-foot-and-mouth disease', *Journal of virology* **89**(12), 6196–6208.

Ma, H.-C., Liu, Y., Wang, C. et al. (2014), 'An interaction between glutathione and the capsid is required for the morphogenesis of c-cluster enteroviruses', *PLoS pathogens* **10**(4), e1004052.

Madden, T. L., Tatusov, R. L. and Zhang, J. (1996), [9] applications of network blast server, *in* 'Methods in enzymology', Vol. 266, Elsevier, pp. 131–141.

Malik, N., Kotecha, A., Gold, S. et al. (2017), 'Structures of foot and mouth disease virus pentamers', *PLoS Pathogens* .

Marquardt, O., Rahman, M. M. and Freiberg, B. (2000), 'Genetic and antigenic variance of foot-and-mouth disease virus type asia1', *Archives of virology* **145**(1), 149–157.

Marsian, J., Fox, H., Bahar, M. W. et al. (2017), 'Plant-made polio type 3 stabilized vlps—a candidate synthetic polio vaccine', *Nature communications* **8**(1), 245.

Martikainen, M., Salorinne, K., Lahtinen, T. et al. (2015), 'Hydrophobic pocket targeting probes for enteroviruses', *Nanoscale* **7**(41), 17457–17467.

Martín-Acebes, M. A., Rincón, V., Armas-Portela, R., Mateu, M. G. and Sobrino, F. (2010), 'A single amino acid substitution in the capsid of foot-and-mouth disease virus can increase acid lability and confer resistance to acid-dependent uncoating inhibition', *Journal of virology* **84**(6), 2902–2912.

Martín-Acebes, M. A., Vázquez-Calvo, Á., Rincón, V., Mateu, M. G. and Sobrino, F. (2011), 'A single amino acid substitution in the capsid of foot-and-mouth disease virus can increase acid resistance', *Journal of virology* **85**(6), 2733–2740.

Mason, P. W., Baxt, B., Brown, F. et al. (1993), 'Antibody-complexed foot-and-mouth disease virus, but not poliovirus, can infect normally insusceptible cells via the fc receptor', *Virology* **192**(2), 568–577.

Mason, P. W., Rieder, E. and Baxt, B. (1994), 'Rgd sequence of foot-and-mouth disease virus is essential for infecting cells via the natural receptor but can be bypassed by an antibody-dependent enhancement pathway.', *Proceedings of the National Academy of Sciences* **91**(5), 1932–1936.

Mateu, M., Martinez, M., Capucci, L. et al. (1990), 'A single amino acid substitution affects multiple overlapping epitopes in the major antigenic site of foot-and-mouth disease virus of serotype c', *Journal of General Virology* **71**(3), 629–637.

Mattion, N., König, G., Seki, C. et al. (2004), 'Reintroduction of foot-and-mouth disease in argentina: characterisation of the isolates and development of tools for the control and eradication of the disease', *Vaccine* **22**(31-32), 4149–4162.

Mayer, M. M., Rapp, H. J., Roizman, B. et al. (1957), 'The purification of poliomyelitis virus as studied by complement fixation', *The Journal of Immunology* **78**(6), 435–455.

McCarthy, F. M., Smith, G. and Mattick, J. (1999), 'Molecular characterisation of australian bovine enteroviruses', *Veterinary microbiology* **68**(1-2), 71–81.

McKenna, R., Xia, D., Willingmann, P. et al. (1992), 'Atomic structure of single-stranded dna bacteriophage φx174 and its functional implications', *Nature* **355**(6356), 137.

McLeish, N. J., Williams, Ç. H., Kaloudas, D., Roivainen, M. M. and Stanway, G. (2012), 'Symmetry-related clustering of positive charges is a common mechanism for heparan sulfate binding in enteroviruses', *Journal of virology* **86**(20), 11163–11170.

McMullan, G., Faruqi, A. and Henderson, R. (2016), Direct electron detectors, *in* 'Methods in enzymology', Vol. 579, Elsevier, pp. 1–17.

Merilahti, P., Koskinen, S., Heikkilä, O., Karelehto, E. and Susi, P. (2012), 'Endocytosis of integrin-binding human picornaviruses', *Advances in virology* **2012**.

Miller, S. T., Hogle, J. M. and Filman, D. J. (2001), 'Ab initio phasing of high-symmetry macromolecular complexes: successful phasing of authentic poliovirus data to 3.0 å resolution', *Journal of molecular biology* **307**(2), 499–512.

Minor, P. D., Ferguson, M., Evans, D. M., Almond, J. W. and Icenogle, J. P. (1986), 'Antigenic structure of polioviruses of serotypes 1, 2 and 3', *Journal of General Virology* **67**(7), 1283–1291.

Minor, W. and Otwinowski, Z. (1997), 'Hkl2000 (denzo-smn) software package', *Processing of X-ray Diffraction Data Collected in Oscillation Mode. In Methods in Enzymology, Macromolecular Crystallography* .

Moll, T. and Davis, A. (1959), 'Isolation and characterization of cytopathogenic enteroviruses from cattle with respiratory disease.', *American journal of veterinary research* **20**, 27–32.

Moore, A., Mercer, J., Dutina, G. et al. (1997), 'Effects of temperature shift on cell cycle, apoptosis and nucleotide pools in cho cell batch cultues', *Cytotechnology* **23**(1-3), 47–54.

Moraes, M., Mayr, G., Mason, P. and Grubman, M. (2002), 'Early protection against homologous challenge after a single dose of replication-defective human adenovirus type 5 expressing capsid proteins of foot-and-mouth disease virus (fmdv) strain a24', *Vaccine* **20**(11-12), 1631–1639.

Muckelbauer, J. K., Kremer, M., Minor, I. et al. (1995), 'The structure of coxsackievirus b3 at 3.5 å resolution', *Structure* **3**(7), 653–667.

Mullapudi, E., Nováček, J., Pálková, L. et al. (2016), 'Structure and genome release mechanism of the human cardiovirus saffold virus 3', *Journal of virology* **90**(17), 7628–7639.

Mullis, K., Faloona, F., Scharf, S. et al. (1986), Specific enzymatic amplification of dna in vitro: the polymerase chain reaction, *in* 'Cold Spring Harbor symposia on quantitative biology', Vol. 51, Cold Spring Harbor Laboratory Press, pp. 263–273.

Murai, T., Nakagawa, Y., Maeda, H. and Terada, K. (2001), 'Altered regulation of cell cycle machinery involved in interleukin-1-induced g1 and g2 phase growth arrest of a375s2 human melanoma cells', *Journal of Biological Chemistry* **276**(9), 6797–6806.

Nasr, M. L., Baptista, D., Strauss, M. et al. (2017), 'Covalently circularized nanodiscs for studying membrane proteins and viral entry', *Nature methods* **14**(1), 49.

Neff, S., Sá-Carvalho, D., Rieder, E. et al. (1998), 'Foot-and-mouth disease virus virulent for cattle utilizes the integrin $\alpha v\beta 3$ as its receptor', *Journal of virology* **72**(5), 3587–3594.

Nelsen-Salz, B., Eggers, H. J. and Zimmermann, H. (1999), 'Integrin $\alpha v\beta 3$ (vitronectin receptor) is a candidate receptor for the virulent echovirus 9 strain barty', *Journal of General Virology* **80**(9), 2311–2313.

Newman, J., Rowlands, D. and Brown, F. (1973), 'A physico-chemical sub-grouping of the mammalian picornaviruses', *Journal of General Virology* **18**(2), 171–180.

Newman, J., Rowlands, D., Brown, F. et al. (1977), 'Physicochemical characterization of two serologically unrelated equine rhinoviruses', *Intervirology* **8**(3), 145–154.

Nilsson, E. C., Jamshidi, F., Johansson, S. M., Oberste, M. S. and Arnberg, N. (2008), 'Sialic acid is a cellular receptor for coxsackievirus a24 variant, an emerging virus with pandemic potential', *Journal of virology* **82**(6), 3061–3068.

Nokhbeh, M. R., Hazra, S., Alexander, D. A. et al. (2005), 'Enterovirus 70 binds to different glycoconjugates containing α2, 3-linked sialic acid on different cell lines', *Journal of virology* **79**(11), 7087–7094.

O'donnell, V., Pacheco, J., Gregg, D. and Baxt, B. (2009), 'Analysis of foot-and-mouth disease virus integrin receptor expression in tissues from naive and infected cattle', *Journal of comparative pathology* **141**(2), 98–112.

Ohi, M., Li, Y., Cheng, Y. and Walz, T. (2004), 'Negative staining and image classification—powerful tools in modern electron microscopy', *Biological procedures online* **6**(1), 23.

Oliveira, A. C., Ishimaru, D., Gonçalves, R. B. et al. (1999), 'Low temperature and pressure stability of picornaviruses: implications for virus uncoating', *Biophysical Journal* **76**(3), 1270–1279.

Oliveira, M. A., Zhao, R., Lee, W.-M. et al. (1993), 'The structure of human rhinovirus 16', *Structure* **1**(1), 51–68.

Organtini, L. J., Makhov, A. M., Conway, J. F., Hafenstein, S. and Carson, S. D. (2014), 'Kinetic and structural analysis of coxsackievirus b3 receptor interactions and formation of the a-particle', *Journal of virology* **88**(10), 5755–5765.

Ostberg, C. O., Zhu, P., Wight, T. N. and Qwarnstrom, E. E. (1995), 'Fibronectin attachment is permissive for il-1 mediated gene regulation', *FEBS letters* **367**(1), 93–97.

Paavonen, J., Jenkins, D., Bosch, F. X. et al. (2007), 'Efficacy of a prophylactic adjuvanted bivalent l1 virus-like-particle vaccine against infection with human papillomavirus types 16 and 18 in young women: an interim analysis of a phase iii double-blind, randomised controlled trial', *The Lancet* **369**(9580), 2161–2170.

Pacheco, J. M., Brum, M. C., Moraes, M. P., Golde, W. T. and Grubman, M. J. (2005), 'Rapid protection of cattle from direct challenge with foot-and-mouth disease virus (fmdv) by a single inoculation with an adenovirus-vectored fmdv subunit vaccine', *Virology* **337**(2), 205–209.

Palmenberg, A., Neubauer, D. and Skern, T. (2010), Genome organization and encoded proteins, *in* 'The picornaviruses', American Society of Microbiology, pp. 3–17.

Panjwani, A., Strauss, M., Gold, S. et al. (2014), 'Capsid protein vp4 of human rhinovirus induces membrane permeability by the formation of a size-selective multimeric pore', *PLoS pathogens* **10**(8), e1004294.

Pantoliano, M. W., Bone, R. F., Rhind, A. W. and Salemme, F. R. (2000), 'Microplate thermal shift assay apparatus for ligand development and multi-variable protein chemistry optimization'. US Patent 6,036,920.

Park, K., Song, J., Baek, K. et al. (2011a), 'Genetic diversity of a korean echovirus 5 isolate and response of the strain to five antiviral drugs', *Virology journal* **8**(1), 79.

Park, K., Yeo, S., Baek, K. et al. (2011b), 'Molecular characterization and antiviral activity test of common drugs against echovirus 18 isolated in korea', *Virology journal* **8**(1), 516.

Parmar, M., Rawson, S., Scarff, C. A. et al. (2018), 'Using a smalp platform to determine a sub-nm single particle cryo-em membrane protein structure', *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1860**(2), 378–383.

Parry, N., Barnett, P. V., Ouldridge, E., Rowlands, D. and Brown, F. (1989), 'Neutralizing epitopes of type o foot-and-mouth disease virus. ii. mapping three conformational sites with synthetic peptide reagents', *Journal of general virology* **70**(6), 1493–1503.

Pasch, A., Küpper, J.-H., Wolde, A., Kandolf, R. and Selinka, H.-C. (1999), 'Comparative analysis of virus–host cell interactions of haemagglutinating and non-haemagglutinating strains of coxsackievirus b3', *Journal of general virology* **80**(12), 3153–3158.

Pascual, E., Mata, C. P., Carrascosa, J. L. and Castón, J. R. (2017), 'Assembly/disassembly of a complex icosahedral virus to incorporate heterologous nucleic acids', *Journal of Physics: Condensed Matter* **29**(49), 494001.

Paton, D., Valarcher, J., Bergmann, I. et al. (2005), 'Selection of foot and mouth disease vaccine strains-a review', *Revue Scientifique Et Technique-Office International Des Epizooties* **24**(3), 981.

Paule, S., Aljofan, M., Simon, C., Rombauts, L. J. and Nie, G. (2012), 'Cleavage of endometrial α-integrins into their functional forms is mediated by proprotein convertase 5/6', *Human reproduction* **27**(9), 2766–2774.

Pécheur, E.-I., Borisevich, V., Halfmann, P. et al. (2016), 'The synthetic antiviral drug arbidol inhibits globally prevalent pathogenic viruses', *Journal of virology* **90**(6), 3086–3092.

Pellinen, T. and Ivaska, J. (2006), 'Integrin traffic', *Journal of cell science* **119**(18), 3723–3731.

Pereira, H. (1981), 'Foot-and-mouth disease', *Virus diseases of food animals* **2**, 333–363.

Perlmutter, J. D. and Hagan, M. F. (2015), 'Mechanisms of virus assembly', *Annual review of physical chemistry* **66**, 217–239.

Pettersen, E. F., Goddard, T. D., Huang, C. C. et al. (2004), 'Ucsf chimera—a visualization system for exploratory research and analysis', *Journal of computational chemistry* **25**(13), 1605–1612.

Pevear, D. C., Fancher, M., Felock, P. et al. (1989), 'Conformational change in the floor of the human rhinovirus canyon blocks adsorption to hela cell receptors.', *Journal of virology* **63**(5), 2002–2007.

Pevear, D. C., Tull, T. M., Seipel, M. E. and Groarke, J. M. (1999), 'Activity of pleconaril against enteroviruses', *Antimicrobial agents and chemotherapy* **43**(9), 2109–2115.

Pietiäinen, V. M., Marjomäki, V., Heino, J. and Hyypiä, T. (2005), 'Viral entry, lipid rafts and caveosomes', *Annals of medicine* **37**(6), 394–403.

Ping, Z., XIONG, W., RODGERS, G. and QWARNSTROM, E. E. (1998), 'Regulation of interleukin 1 signalling through integrin binding and actin reorganization: disparate effects on nf-κb and stress kinase pathways', *Biochemical Journal* **330**(2), 975–981.

Plevka, P., Perera, R., Cardosa, J., Kuhn, R. J. and Rossmann, M. G. (2012), 'Crystal structure of human enterovirus 71', *Science* **336**(6086), 1274–1274.

Plevka, P., Perera, R., Yap, M. L. et al. (2013), 'Structure of human enterovirus 71 in complex with a capsid-binding inhibitor', *Proceedings of the National Academy of Sciences* **110**(14), 5463–5467.

Plummer, G. and Kerry, J. (1962), 'Studies on an equine respiratory virus', *Vet. Rec* **74**, 967–970.

Porta, C., Kotecha, A., Burman, A. et al. (2013a), 'Rational engineering of recombinant picornavirus capsids to produce safe, protective vaccine antigen', *PLoS pathogens* **9**(3), e1003255.

Powell, R. M., Ward, T., Evans, D. J. and Almond, J. W. (1997), 'Interaction between echovirus 7 and its receptor, decay-accelerating factor (cd55): evidence for a secondary cellular factor in a-particle formation.', *Journal of virology* **71**(12), 9306–9312.

Pulli, T., Lankinen, H., Roivainen, M. and Hyypiä, T. (1998), 'Antigenic sites of coxsackievirus a9', *Virology* **240**(2), 202–212.

Punjani, A., Rubinstein, J. L., Fleet, D. J. and Brubaker, M. A. (2017), 'cryosparc: algorithms for rapid unsupervised cryo-em structure determination', *Nature methods* **14**(3), 290.

Pytela, R., Pierschbacher, M. D. and Ruoslahti, E. (1985), 'A 125/115-kda cell surface receptor specific for vitronectin interacts with the arginine-glycine-aspartic acid adhesion sequence derived from fibronectin', *Proceedings of the National Academy of Sciences* **82**(17), 5766–5770.

Querol-Audí, J., Konecsni, T., Pous, J. et al. (2009), 'Minor group human rhinovirus–receptor interactions: Geometry of multimodular attachment and basis of recognition', *FEBS letters* **583**(1), 235–240.

Quispe, J., Damiano, J., Mick, S. E. et al. (2007), 'An improved holey carbon film for cryo-electron microscopy', *Microscopy and Microanalysis* **13**(5), 365–371.

Racaniello, V. R. (2007), 'Picornaviridae, the virus and their replication', *Fields virology* pp. 795–838.

Razinkov, I., Dandey, V. P., Wei, H. et al. (2016), 'A new method for vitrifying samples for cryoem', *Journal of structural biology* **195**(2), 190–198.

Reeve, R., Blignaut, B., Esterhuysen, J. J. et al. (2010), 'Sequence-based prediction for vaccine strain selection and identification of antigenic variability in foot-and-mouth disease virus', *PLoS computational biology* **6**(12), e1001027.

Reeve, R., Borley, D. W., Maree, F. F. et al. (2016), 'Tracking the antigenic evolution of foot-and-mouth disease virus', *PloS one* **11**(7), e0159360.

Ren, J., Wang, X., Hu, Z. et al. (2013), 'Picornavirus uncoating intermediate captured in atomic detail', *Nature communications* **4**, 1929.

Ren, J., Wang, X., Zhu, L. et al. (2015), 'Structures of coxsackievirus a16 capsids with native antigenicity: implications for particle expansion, receptor binding, and immunogenicity', *Journal of virology* **89**(20), 10500–10511.

Rhinow, D. and Kühlbrandt, W. (2008), 'Electron cryo-microscopy of biological specimens on conductive titanium–silicon metal glass films', *Ultramicroscopy* **108**(7), 698–705.

Robinson, W., Cordeiro, M., Abdel-Malek, S. et al. (1996), 'Dicaffeoylquinic acid inhibitors of human immunodeficiency virus integrase: inhibition of the core catalytic domain of human immunodeficiency virus integrase.', *Molecular pharmacology* **50**(4), 846–855.

Robson, K. J., Harris, T. and Brown, F. (1977), 'An assessment by competition hybridization of the sequence homology between the rnas of the seven serotypes of fmdv', *Journal of General Virology* **37**(2), 271–276.

Roedig, P., Ginn, H. M., Pakendorf, T. et al. (2017), 'High-speed fixed-target serial virus crystallography', *Nature methods* **14**(8), 805.

Rohou, A. and Grigorieff, N. (2015), 'Ctffind4: Fast and accurate defocus estimation from electron micrographs', *Journal of structural biology* **192**(2), 216–221.

Roivainen, M., Piirainen, L., Ryä, T., Närvänen, A. and Hovi, T. (1993), 'An immunodominant n-terminal region of vp1 protein of poliovirion that is buried in crystal structure can be exposed in solution', *Virology* **195**(2), 762–765.

Rombaut, B., Brioen, P. and Boeyé, A. (1990), 'Disoxaril stabilization and immunogenicity of poliovirus procapsids', *Journal of general virology* **71**(5), 1081–1086.

Rombaut, B. and Jore, J. (1997), 'Immunogenic, non-infectious polio subviral particles synthesized in saccharomyces cerevisiae.', *Journal of general virology* **78**(8), 1829–1832.

Rossmann, M. G., Arnold, E., Erickson, J. W. et al. (1985), 'Structure of a human common cold virus and functional relationship to other picornaviruses', *Nature* **317**(6033), 145.

Rossmann, M. G. and Johnson, J. E. (1989), 'Icosahedral rna virus structure', *Annual review of biochemistry* **58**(1), 533–569.

Rowlands, D., Sangar, D. and Brown, F. (1971), 'Buoyant density of picornaviruses in caesium salts', *Journal of General Virology* **13**(1), 141–152.

Rowlands, D., Sangar, D. and Brown, F. (1975), 'A comparative chemical and serological study of the full and empty particles of foot-and mouth disease virus', *Journal of General Virology* **26**(3), 227–238.

Royston, L., Essaidi-Laziosi, M., Pérez-Rodríguez, F. J. et al. (2018), 'Viral chimeras decrypt the role of enterovirus capsid proteins in viral tropism, acid sensitivity and optimal growth temperature', *PLoS pathogens* **14**(4), e1006962.

Rozov, A., Khusainov, I., El Omari, K. et al. (2019), 'Importance of potassium ions for ribosome structure and function revealed by long-wavelength x-ray diffraction', *Nature communications* **10**(1), 2519.

Rubinstein, J. L., Guo, H., Ripstein, Z. A. et al. (2019), 'Shake-it-off: A simple ultrasonic cryo-em specimen preparation device', *bioRxiv* p. 632125.

Russo, C. J. and Passmore, L. A. (2014), 'Ultrastable gold substrates for electron cryomicroscopy', *Science* **346**(6215), 1377–1380.

Sa-Carvalho, D., Rieder, E., Baxt, B. et al. (1997), 'Tissue culture adaptation of foot-and-mouth disease virus selects viruses that bind to heparin and are attenuated in cattle.', *Journal of virology* **71**(7), 5115–5123.

Sabin, A. B. (1957), 'Present status of attenuated live virus poliomyelitis vaccine', *Bulletin of the New York Academy of Medicine* **33**(1), 17.

Saiki, R. K., Scharf, S., Faloona, F. et al. (1985), 'Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia', *Science* **230**(4732), 1350–1354.

Salem, R., El-Kholy, A. A. and Ibrahim, M. (2019), 'Eight novel single chain antibody fragments recognising vp2 of foot-and-mouth disease virus serotypes a, o, and sat 2', *Virology* **533**, 145–154.

Salk, J. E., Krech, U., Youngner, J. et al. (1954), 'Formaldehyde treatment and safety testing of experimental poliomyelitis vaccines', *American Journal of Public Health and the Nations Health* **44**(5), 563–570.

Sangar, D., Newton, S., Rowlands, D. and Clarke, B. (1987), 'All foot and mouth disease virus serotypes initiate protein synthesis at two separate augs.', *Nucleic acids research* **15**(8), 3305.

Sankar, S. and Porter, A. (1992), 'Point mutations which drastically affect the polymerization activity of encephalomyocarditis virus rna-dependent rna polymerase correspond to the active site of escherichia coli dna polymerase i.', *Journal of Biological Chemistry* **267**(14), 10168–10176.

Sastry, S. K. and Burridge, K. (2000), 'Focal adhesions: a nexus for intracellular signaling and cytoskeletal dynamics', *Experimental cell research* **261**(1), 25–36.

Satheshkumar, P., Lokesh, G., Sangita, V. et al. (2004), 'Role of metal ion-mediated interactions in the assembly and stability of sesbania mosaic virus t= 3 and t= 1 capsids', *Journal of molecular biology* **342**(3), 1001–1014.

Savithri, H. and Erickson, J. W. (1983), 'The self-assembly of the cowpea strain of southern bean mosaic virus: formation of t= 1 and t= 3 nucleoprotein particles', *Virology* **126**(1), 328–335.

Saxton, W. and Baumeister, W. (1982), 'The correlation averaging of a regularly arranged bacterial cell envelope protein', *Journal of microscopy* **127**(2), 127–138.

Scheres, S. H. (2012), 'Relion: implementation of a bayesian approach to cryo-em structure determination', *Journal of structural biology* **180**(3), 519–530.

Scheres, S. H. and Chen, S. (2012), 'Prevention of overfitting in cryo-em structure determination', *Nature methods* **9**(9), 853.

Schmidli, C., Albiez, S., Rima, L. et al. (2019), 'Microfluidic protein isolation and sample preparation for high resolution cryo-em', *bioRxiv* p. 556068.

Schwarz, K., Meijerink, E., Speiser, D. E. et al. (2005), 'Efficient homologous prime-boost strategies for t cell vaccination based on virus-like particles', *European journal of immunology* **35**(3), 816–821.

Schwerdt, C. E. and Fogh, J. (1957), 'The ratio of physical particles per infectious unit observed for poliomyelitis viruses', *Virology* **4**(1), 41–52.

Scott, K. A., Kotecha, A., Seago, J. et al. (2017), 'Sat2 foot-and-mouth disease virus structurally modified for increased thermostability', *Journal of virology* **91**(10), e02312–16.

Seitsonen, J., Susi, P., Heikkilä, O. et al. (2010), 'Interaction of αvβ3 and αvβ6 integrins with human parechovirus 1', *Journal of virology* **84**(17), 8509–8519.

Shafren, D. R., Dorahy, D. J., Ingham, R. A., Burns, G. F. and Barry, R. D. (1997), 'Coxsackievirus a21 binds to decay-accelerating factor but requires intercellular adhesion molecule 1 for cell entry.', *Journal of virology* **71**(6), 4736–4743.

Shah, P. N., Filman, D. J., Karunatilaka, K. S. et al. (2020), 'Cryo-em structures reveal two distinct conformational states in a picornavirus cell entry intermediate.', *bioRxiv* .

Shakeel, S., Seitsonen, J. J., Kajander, T. et al. (2013), 'Structural and functional analysis of coxsackievirus a9 integrin αvβ6 binding and uncoating', *Journal of virology* **87**(7), 3943–3951.

Sheldrick, G. M. (2008), 'A short history of shelx', *Acta Crystallographica Section A: Foundations of Crystallography* **64**(1), 112–122.

Sherry, B., Mosser, A., Colonno, R. and Rueckert, R. (1986), 'Use of monoclonal antibodies to identify four neutralization immunogens on a common cold picornavirus, human rhinovirus 14.', *Journal of virology* **57**(1), 246–257.

Shiman, R. and Draper, D. E. (2000), 'Stabilization of rna tertiary structure by monovalent cations', *Journal of molecular biology* **302**(1), 79–91.

Shingler, K. L., Yoder, J. L., Carnegie, M. S. et al. (2013), 'The enterovirus 71 a-particle forms a gateway to allow genome release: a cryoem study of picornavirus uncoating', *PLoS pathogens* **9**(3), e1003240.

Sievers, F., Wilm, A., Dineen, D. et al. (2011), 'Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega', *Molecular systems biology* **7**(1).

Sinkala, Y., Simuunza, M., Pfeiffer, D. et al. (2014), 'Challenges and economic implications in the control of foot and mouth disease in sub-saharan africa: Lessons from the zambian experience', *Veterinary medicine international* **2014**.

Skern, T. (2010), '100 years poliovirus: from discovery to eradication. a meeting report', *Archives of virology* **155**(9), 1371–1381.

Smith, A. D. and Dawson, H. (2006), 'Glutathione is required for efficient production of infectious picornavirus virions', *Virology* **353**(2), 258–267.

Smith, J. W. and Cheresh, D. A. (1990), 'Integrin (alpha v beta 3)-ligand interaction. identification of a heterodimeric rgd binding site on the vitronectin receptor.', *Journal of Biological Chemistry* **265**(4), 2168–2172.

Smyth, M., Hoey, E., Trudgett, A., Martin, S. and Brown, F. (1990), 'Chemically synthesized peptides elicit neutralizing antibody to bovine enterovirus', *Journal of general virology* **71**(1), 231–234.

Smyth, M. and Martin, J. H. (2001), 'Structural, biochemical and electrostatic basis of serotype specificity in bovine enteroviruses', *Archives of virology* **146**(2), 347–355.

Smyth, M. and Martin, J. H. (2002), 'Picornavirus uncoating', *Molecular Pathology* **55**(4), 214.

Smyth, M., Pettitt, T., Symonds, A. and Martin, J. (2003), 'Identification of the pocket factors in a picornavirus', *Archives of virology* **148**(6), 1225–1233.

Smyth, M., Tate, J., Hoey, E. et al. (1995), 'Implications for viral uncoating from the structure of bovine enterovirus', *Nature structural biology* **2**(3), 224.

Smyth, M., Trudgett, A., Hoey, E., Martin, S. and Brown, F. (1992), 'Characterization of neutralizing antibodies to bovine enterovirus elicited by synthetic peptides', *Archives of virology* **126**(1-4), 21–33.

Song, B., Lenhart, J., Flegler, V. J. et al. (2019), 'Capabilities of the falcon iii detector for single-particle structure determination', *Ultramicroscopy* **203**, 145–154.

Speir, J. A., Bothner, B., Qu, C. et al. (2006), 'Enhanced local symmetry interactions globally stabilize a mutant virus capsid that maintains infectivity and capsid dynamics', *Journal of virology* **80**(7), 3582–3591.

Speir, J. A., Munshi, S., Wang, G., Baker, T. S. and Johnson, J. E. (1995), 'Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by x-ray crystallography and cryo-electron microscopy', *Structure* **3**(1), 63–78.

Stevenson, R. A., Huang, J.-a., Studdert, M. J. and Hartley, C. A. (2004), 'Sialic acid acts as a receptor for equine rhinitis a virus binding and infection', *Journal of general virology* **85**(9), 2535–2543.

Stoner, G. D., Williams, B., Zniazeff, A. and Shimkin, M. B. (1973), 'Effect of neuraminidase pretreatment on the susceptibility of normal and transformed mammalian cells to bovine enterovirus 261', *Nature* **245**(5424), 319.

Strauss, M., Levy, H. C., Bostina, M., Filman, D. J. and Hogle, J. M. (2013), 'Rna transfer from poliovirus 135s particles across membranes is mediated by long umbilical connectors', *Journal of virology* **87**(7), 3903–3914.

Strauss, M., Schotte, L., Thys, B., Filman, D. J. and Hogle, J. M. (2016), 'Five of five vhhs neutralizing poliovirus bind the receptor-binding site', *Journal of virology* **90**(7), 3496–3505.

Stuart, D. I., Levine, M., Muirhead, H. and Stammers, D. K. (1979), 'Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 å', *Journal of molecular biology* **134**(1), 109–142.

Subramanian, B. M., Madhanmohan, M., Sriraman, R. et al. (2012), 'Development of foot-and-mouth disease virus (fmdv) serotype o virus-like-particles (vlps) vaccine and evaluation of its potency', *Antiviral research* **96**(3), 288–295.

Takagi, J., Petre, B. M., Walz, T. and Springer, T. A. (2002), 'Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling', *Cell* **110**(5), 599–611.

Tamkun, J. W., DeSimone, D. W., Fonda, D. et al. (1986), 'Structure of integrin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin', *Cell* **46**(2), 271–282.

Tan, C. W., Lai, J. K. F., Sam, I.-C. and Chan, Y. F. (2014), 'Recent developments in antiviral agents against enterovirus 71 infection', *Journal of biomedical science* **21**(1), 14.

Tan, C. W., Poh, C. L., Sam, I.-C. and Chan, Y. F. (2013), 'Enterovirus 71 uses cell surface heparan sulfate glycosaminoglycan as an attachment receptor', *Journal of virology* **87**(1), 611–620.

Tavakkol, A. and Burness, A. T. H. (1990), 'Evidence for a direct role for sialic acid in the attachment of encephalomyocarditis virus to human erythrocytes', *Biochemistry* **29**(47), 10684–10690.

Tegunov, D. and Cramer, P. (2018), 'Real-time cryo-em data pre-processing with warp', *BioRxiv* p. 338558.

Teterina, N. L., Gorbalenya, A. E., Egger, D. et al. (2006), 'Testing the modularity of the n-terminal amphipathic helix conserved in picornavirus 2c proteins and hepatitis c ns5a protein', *Virology* **344**(2), 453–467.

Thibaut, H. J., Leyssen, P., Puerstinger, G. et al. (2011), 'Towards the design of combination therapy for the treatment of enterovirus infections', *Antiviral research* **90**(3), 213–217.

Thibaut, H. J., van der Linden, L., Jiang, P. et al. (2014), 'Binding of glutathione to enterovirus capsids is essential for virion morphogenesis', *PLoS pathogens* **10**(4), e1004039.

Thomas, A., Woortmeijer, R., Puijk, W. and Barteling, S. (1988), 'Antigenic sites on foot-and-mouth disease virus type a10.', *Journal of virology* **62**(8), 2782–2789.

Thorn, A. and Sheldrick, G. M. (2011), 'Anode: anomalous and heavy-atom density calculation', *Journal of applied crystallography* **44**(6), 1285–1287.

Tsang, S. K., Danthi, P., Chow, M. and Hogle, J. M. (2000), 'Stabilization of poliovirus by capsid-binding antiviral drugs is due to entropic effects', *Journal of molecular biology* **296**(2), 335–340.

Tsang, S. K., McDermott, B. M., Racaniello, V. R. and Hogle, J. M. (2001), 'Kinetic analysis of the effect of poliovirus receptor on viral uncoating: the receptor as a catalyst', *Journal of virology* **75**(11), 4984–4989.

Tsunoda, I., Libbey, J. E. and Fujinami, R. S. (2009), 'Theiler's murine encephalomyelitis virus attachment to the gastrointestinal tract is associated with sialic acid binding', *Journal of neurovirology* **15**(1), 81–89.

Tuller, G., Nemec, T., Hrastnik, C. and Daum, G. (1999), 'Lipid composition of subcellular membranes of an fy1679-derived haploid yeast wild-type strain grown on different carbon sources', *Yeast* **15**(14), 1555–1564.

Tuthill, T. J., Bubeck, D., Rowlands, D. J. and Hogle, J. M. (2006), 'Characterization of early steps in the poliovirus infection process: receptor-decorated liposomes induce conversion of the virus to membrane-anchored entry-intermediate particles', *Journal of virology* **80**(1), 172–180.

Tuthill, T. J., Groppelli, E., Hogle, J. M. and Rowlands, D. J. (2010), Picornaviruses, *in* 'Cell Entry by Non-Enveloped Viruses', Springer, pp. 43–89.

Tuthill, T. J., Harlos, K., Walter, T. S. et al. (2009), 'Equine rhinitis a virus and its low ph empty particle: clues towards an aphthovirus entry mechanism?', *PLoS pathogens* **5**(10), e1000620.

Urakawa, T., Ferguson, M., Minor, P. D. et al. (1989), 'Synthesis of immunogenic, but non-infectious, poliovirus particles in insect cells by a baculovirus expression vector', *Journal of general virology* **70**(6), 1453–1463.

van Vlijmen, H. W., Curry, S., Schaefer, M. and Karplus, M. (1998), 'Titration calculations of foot-and-mouth disease virus capsids and their stabilities as a function of ph', *Journal of molecular biology* **275**(2), 295–308.

Verdaguer, N., Blaas, D. and Fita, I. (2000), 'Structure of human rhinovirus serotype 2 (hrv2)', *Journal of molecular biology* **300**(5), 1179–1194.

Verdaguer, N., Fita, I., Reithmayer, M., Moser, R. and Blaas, D. (2004), 'X-ray structure of a minor group human rhinovirus bound to a fragment of its cellular receptor protein', *Nature structural & molecular biology* **11**(5), 429.

Viljoen, B., Kock, J. and Lategan, P. (1986), 'Long-chain fatty acid composition of selected genera of yeasts belonging to the endomycetales', *Antonie van leeuwenhoek* **52**(1), 45–51.

Wachsman, M., Castilla, V. and Coto, C. (1998), 'Inhibition of foot and mouth disease virus (fmdv) uncoating by a plant-derived peptide isolated from melia azedarach l leaves', *Archives of virology* **143**(3), 581–590.

Walker, J. W., Reid, G. P., McCray, J. A. and Trentham, D. R. (1988), 'Photolabile 1-(2-nitrophenyl) ethyl phosphate esters of adenine nucleotide analogs. synthesis and mechanism of photolysis', *Journal of the American Chemical Society* **110**(21), 7170–7177.

Wallace, A. C., Laskowski, R. A. and Thornton, J. M. (1995), 'Ligplot: a program to generate schematic diagrams of protein-ligand interactions', *Protein engineering, design and selection* **8**(2), 127–134.

Walter, T., Diprose, J., Brown, J. et al. (2003), 'A procedure for setting up high-throughput nanolitre crystallization experiments. i. protocol design and validation', *Journal of Applied Crystallography* **36**(2), 308–314.

Walter, T. S., Diprose, J. M., Mayo, C. J. et al. (2005), 'A procedure for setting up high-throughput nanolitre crystallization experiments. crystallization workflow for initial screening, automated storage, imaging and optimization', *Acta Crystallographica Section D: Biological Crystallography* **61**(6), 651–657.

Walter, T. S., Ren, J., Tuthill, T. J. et al. (2012), 'A plate-based high-throughput assay for virus stability and vaccine formulation', *Journal of virological methods* **185**(1), 166–170.

Wang, X., Peng, W., Ren, J. et al. (2012), 'A sensor-adaptor mechanism for enterovirus uncoating from structures of ev71', *Nature structural & molecular biology* **19**(4), 424.

Wang, X., Xiao, X., Zhao, M. et al. (2016), 'Ev71 virus-like particles produced by co-expression of capsid proteins in yeast cells elicit humoral protective response against ev71 lethal challenge', *BMC research notes* **9**(1), 42.

Warburg, O. and Christian, W. (1941), 'Isolierung und kristallisation des gärungsferments enolase', *Naturwissenschaften* **29**(39), 589–590.

Warner, S., Hartley, C. A., Stevenson, R. A. et al. (2001), 'Evidence that equine rhinitis a virus vp1 is a target of neutralizing antibodies and participates directly in receptor binding', *Journal of virology* **75**(19), 9274–9281.

Waterman, D. G., Winter, G., Parkhurst, J. M. et al. (2013), 'The dials framework for integration software', *CCP4 Newsletter on Protein Crystallography* **49**, 16–19.

Watson, M. L. (1956), 'Carbon films and specimen stability', *The Journal of biophysical and biochemical cytology* **2**(4), 31.

Wei, H., Dandey, V. P., Zhang, Z. et al. (2018), 'Optimizing "self-wicking" nanowire grids', *Journal of structural biology* **202**(2), 170–174.

Wien, M. W., Curry, S., Filman, D. J. and Hogle, J. M. (1997), 'Structural studies of poliovirus mutants that overcome receptor defects', *Nature structural biology* **4**(8), 666.

Wilson, K. and Walker, J. (2010), *Principles and techniques of biochemistry and molecular biology*, Cambridge university press.

Winn, M. D., Ballard, C. C., Cowtan, K. D. et al. (2011), 'Overview of the ccp4 suite and current developments', **67**(4), 235–242.

Winter, G. (2010), 'xia2: an expert system for macromolecular crystallography data reduction.', *J. Appl. Cryst.* **43**, 186–190.

Witz, J. and Brown, F. (2001), 'Structural dynamics, an intrinsic property of viral capsids', *Archives of virology* **146**(12), 2263–2274.

Woo, P. C., Lau, S. K., Choi, G. K. et al. (2012), 'Complete genome sequence of a novel picornavirus, canine picornavirus, discovered in dogs'.

Wu, G. and Zhu, J. (2012), 'Nmr studies of alkali metal ions in organic and biological solids.', *Progress in nuclear magnetic resonance spectroscopy* **61**, 1.

Xiao, C., Bator-Kelly, C. M., Rieder, E. et al. (2005a), 'The crystal structure of coxsackievirus a21 and its interaction with icam-1', *Structure* **13**(7), 1019–1033.

Xiao, C., Chipman, P. R., Battisti, A. J. et al. (2005b), 'Cryo-electron microscopy of the giant mimivirus', *Journal of molecular biology* **353**(3), 493–496.

Xiao, C., Kuznetsov, Y. G., Sun, S. et al. (2009), 'Structural studies of the giant mimivirus', *PLoS biology* **7**(4), e1000092.

Xie, C., Zhu, J., Chen, X. et al. (2010), 'Structure of an integrin with an αi domain, complement receptor type 4', *The EMBO journal* **29**(3), 666–679.

Xiong, J.-P., Stehle, T., Diefenbach, B. et al. (2001), 'Crystal structure of the extracellular segment of integrin $\alpha v\beta 3$', *Science* **294**(5541), 339–345.

Xu, Y., Ma, S., Huang, Y. et al. (2019), 'Virus-like particle vaccines for poliovirus types 1, 2, and 3 with enhanced thermostability expressed in insect cells', *Vaccine* **37**(17), 2340–2347.

Yang, B., Chuang, H. and Yang, K. D. (2009), 'Sialylated glycans as receptor and inhibitor of enterovirus 71 infection to dld-1 intestinal cells', *Virology journal* **6**(1), 141.

Yang, Z., Gao, F., Wang, X. et al. (2019), 'Development and characterization of an enterovirus 71 (ev71) virus-like particles (vlps) vaccine produced in pichia pastoris', *Human vaccines & immunotherapeutics* pp. 1–9.

Yano, T., Premashthira, S., Dejyong, T., Tangtrongsup, S. and Salman, M. D. (2018), 'The effectiveness of a foot and mouth disease outbreak control programme in thailand 2008–2015: Case studies and lessons learned', *Veterinary sciences* **5**(4), 101.

Yoshioka, C., Carragher, B. and Potter, C. S. (2010), 'Cryomesh™: A new substrate for cryo-electron microscopy', *Microscopy and Microanalysis* **16**(1), 43–53.

Young, G. and Kukura, P. (2019), 'Interferometric scattering microscopy', *Annual review of physical chemistry* **70**, 301–322.

Yu, Y., Zhu, C., Wang, S. et al. (2013), 'Homosecoiridoid alkaloids with amino acid units from the flower buds of lonicera japonica', *Journal of natural products* **76**(12), 2226–2233.

Zambon, M. and Martin, J. (2018), 'Polio eradication: next steps and future challenges', *Eurosurveillance* **23**(47).

Zell, R. (2018), 'Picornaviridae—the ever-growing vírus family', *Archives of virology* **163**(2), 299–317.

Zell, R., Delwart, E., Gorbalenya, A. et al. (2017), 'Ictv virus taxonomy profile: Picornaviridae', *The Journal of general virology* **98**(10), 2421.

Zell, R., Krumbholz, A., Dauber, M., Hoey, E. and Wutzler, P. (2006), 'Molecular-based reclassification of the bovine enteroviruses', *Journal of general virology* **87**(2), 375–385.

Zhang, A., Smith, J. and Burgess, G. (1990), 'Blocking enzyme-linked immunosorbent assay for detection of antibodies against bovine enterovirus', *Veterinary microbiology* **21**(3), 275–281.

Zhang, C., Liu, Q., Ku, Z. et al. (2016), 'Coxsackievirus a16-like particles produced in pichia pastoris elicit high-titer neutralizing antibodies and confer protection against lethal viral challenge in mice', *Antiviral research* **129**, 47–51.

Zhang, G., Zhou, F., Gu, B. et al. (2012), 'In vitro and in vivo evaluation of ribavirin and pleconaril antiviral activity against enterovirus 71 infection', *Archives of virology* **157**(4), 669–679.

Zhang, Y., Simpson, A. A., Ledford, R. M. et al. (2004), 'Structural and virological studies of the stages of virus replication that are affected by antirhinovirus compounds', *Journal of virology* **78**(20), 11061–11069.

Zhao, H., Li, H.-Y., Han, J.-F. et al. (2013), 'Virus-like particles produced in saccharomyces cerevisiae elicit protective immunity against coxsackievirus a16 in mice', *Applied microbiology and biotechnology* **97**(24), 10445–10452.

Zhao, R., Hadfield, A. T., Kremer, M. J. and Rossmann, M. G. (1997), 'Cations in human rhinoviruses', *Virology* **227**(1), 13–23.

Zhao, R., Pevear, D. C., Kremer, M. J. et al. (1996), 'Human rhinovirus 3 at 3.0 å resolution', *Structure* **4**(10), 1205–1220.

Zheng, M., Jin, N., Zhang, H. et al. (2006), 'Construction and immunogenicity of a recombinant fowlpox virus containing the capsid and 3c protease coding regions of foot-and-mouth disease virus', *Journal of virological methods* **136**(1-2), 230–237.

Zheng, S. Q., Palovcak, E., Armache, J.-P. et al. (2017), 'Motioncor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy', *Nature methods* **14**(4), 331.

Zhou, D., Zhao, Y., Kotecha, A. et al. (2019), 'Unexpected mode of engagement between enterovirus 71 and its receptor scarb2', *Nature microbiology* **4**(3), 414.

Zhou, L., Luo, Y., Wu, Y., Tsao, J. and Luo, M. (2000), 'Sialylation of the host receptor may modulate entry of demyelinating persistent theiler's virus', *Journal of virology* **74**(3), 1477–1485.

Zhu, L., Wang, X., Ren, J. et al. (2015), 'Structure of ljungan virus provides insight into genome packaging of this picornavirus', *Nature communications* **6**, 8316.

Zhu, P., Xiong, W., Rodgers, G. and Qwarnstrom, E. E. (1998), 'Regulation of interleukin 1 signalling through integrin binding and actin reorganization: disparate effects on nf-kappab and stress kinase pathways.', *Biochemical Journal* **330**(Pt 2), 975.

Zhu, R., Xu, L., Zheng, Q. et al. (2018), 'Discovery and structural characterization of a therapeutic antibody against coxsackievirus a10', *Science Advances* **4**(9).

Zivanov, J., Nakane, T., Forsberg, B. O. et al. (2018), 'New tools for automated high-resolution cryo-em structure determination in relion-3', *Elife* **7**, e42166.

Zocher, G., Mistry, N., Frank, M. et al. (2014), 'A sialic acid binding site in a human picornavirus', *PLoS pathogens* **10**(10), e1004401.

Chapter 10

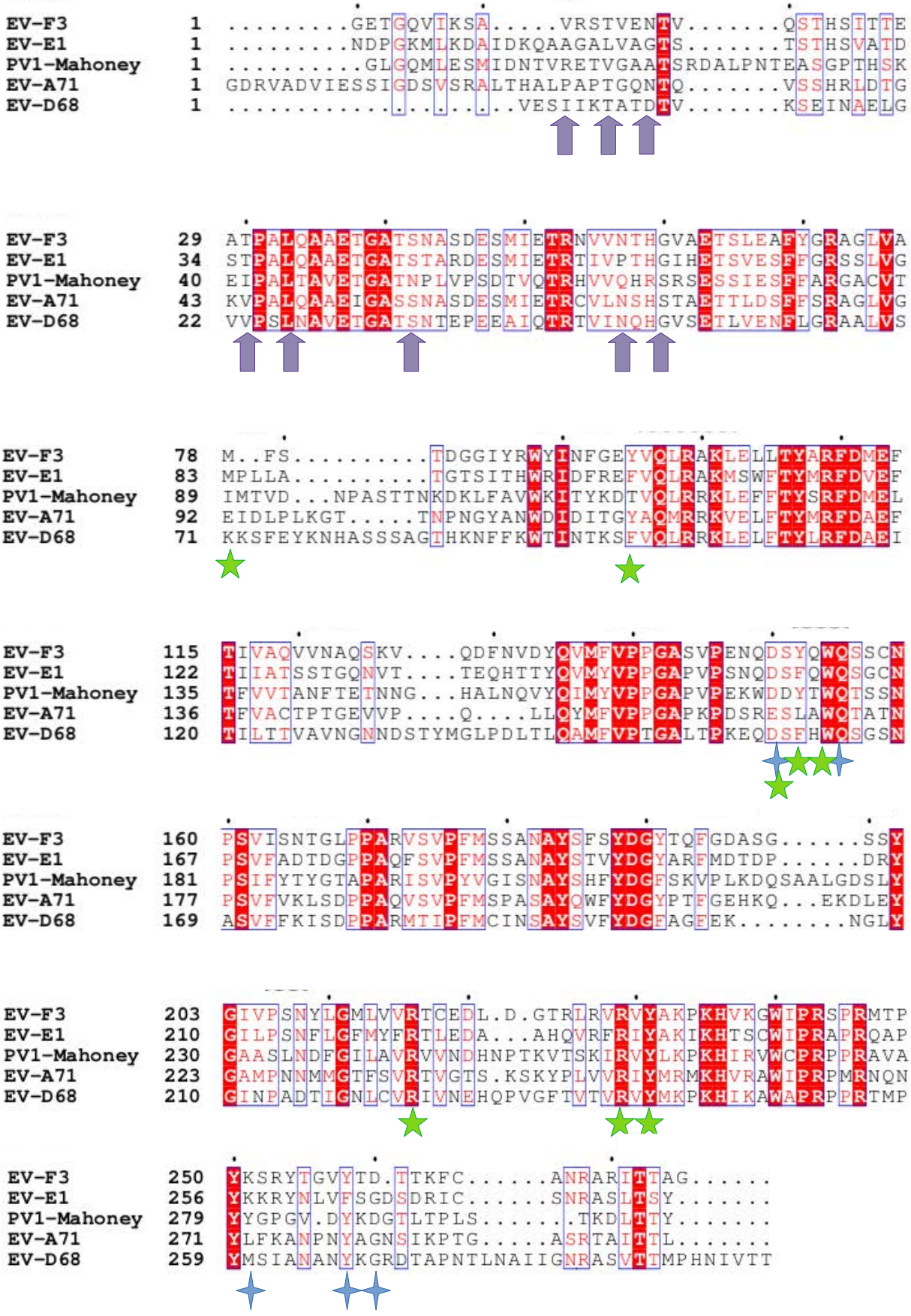
Appendix

10.1 Description of contents

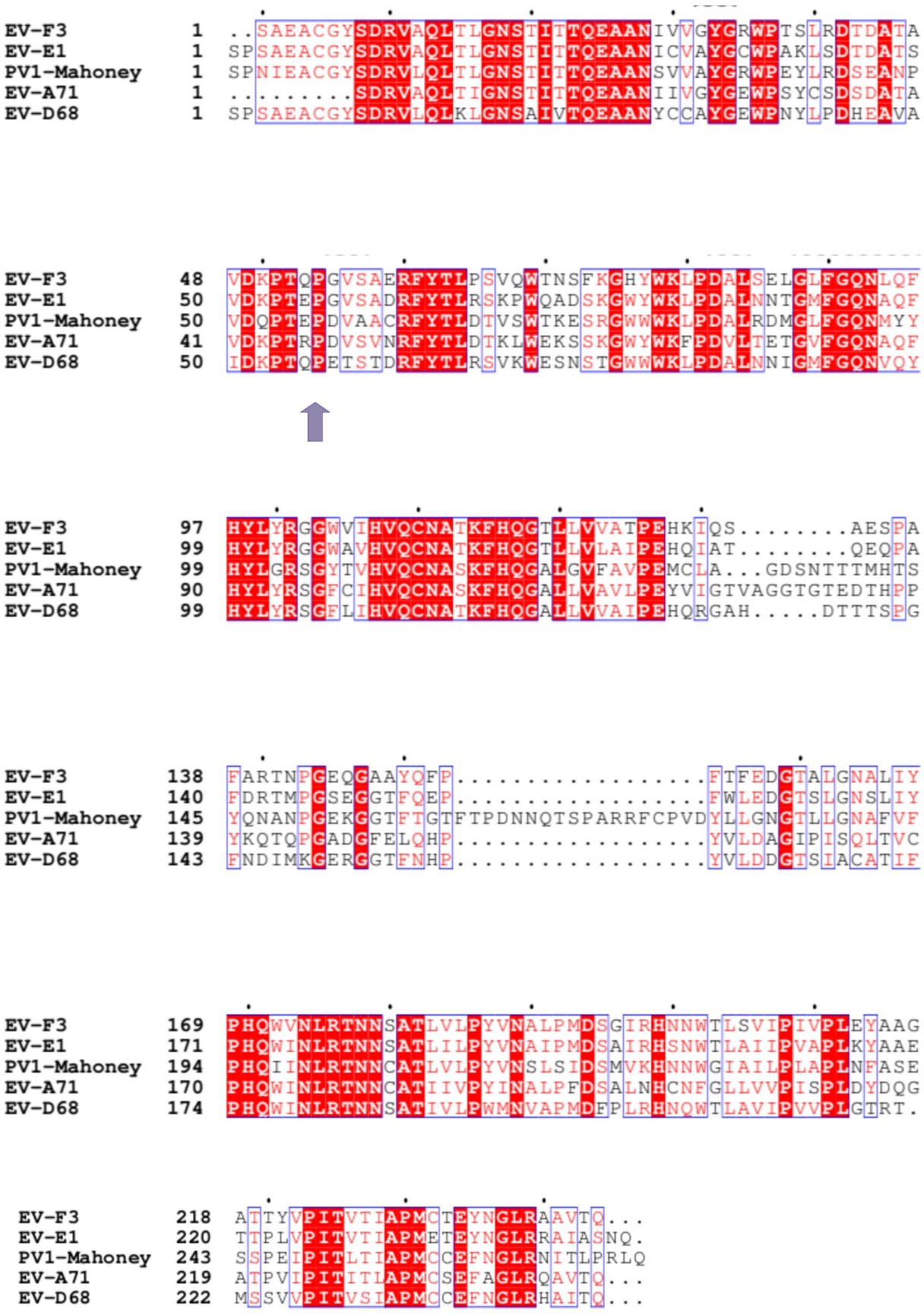
- Sequence for EV-F3 for Chapters 3 and 4. Residues that appear to contact modelled K^+ ions are indicated by purple arrows. Those that were found to contact GSH with green stars, and those that contacted our modelled SI by blue three-pointed stars.
- Sequence alignment for FMDV VLP constructs and selected wild-type FMDV serotypes for Chapters 5 and 7.
- Sequence alignment for PV VLP constructs. Pertinent for Chapter 6.
- Copies of published papers.

EV-F3 Sequence alignment

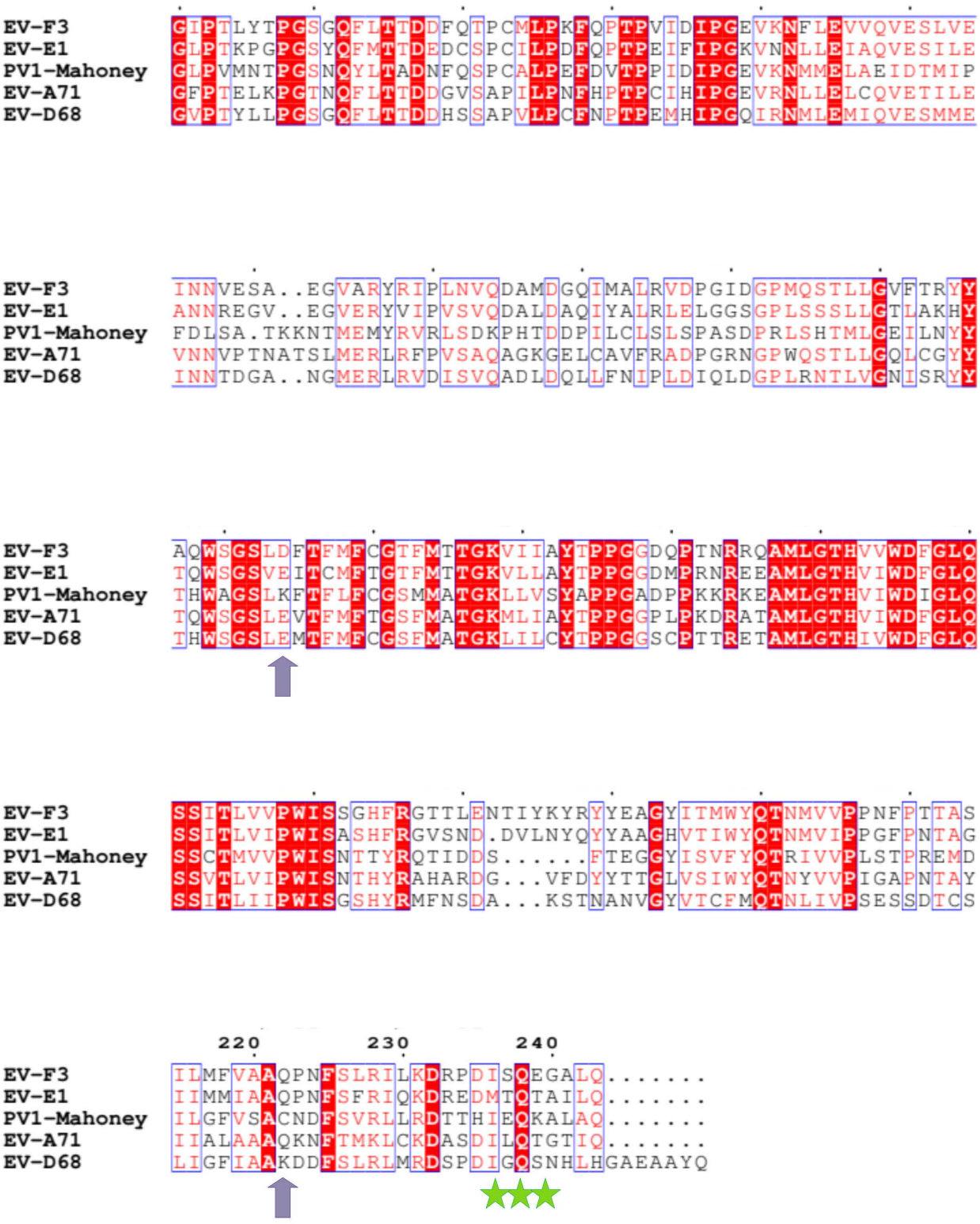
VP1



VP2

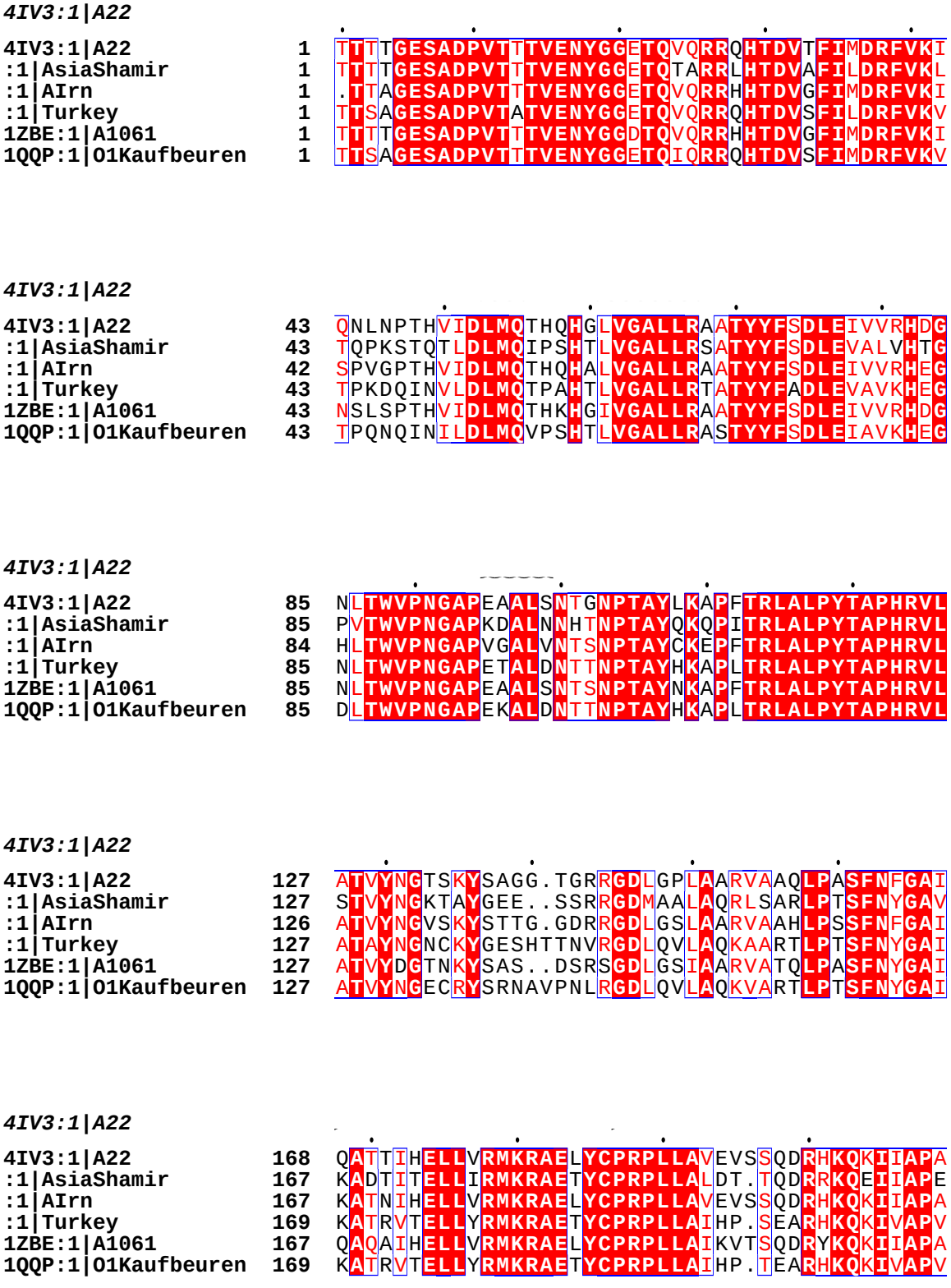

EV-F3 1 ..SAEACGYSDRVAQLTLGNSTITTQEAANIVVGYGRWPTSLRDTDATA
EV-E1 1 SPSAEACGYSDRVAQLTLGNSTITTQEAANICVAYGCWPAKLSDTDATS
PV1-Mahoney 1 SPNIEACGYSDRVLQLTLGNSTITTQEAANSVVAYGRWPEYLRDSEANP
EV-A71 1 ..SDRVAQLTIGNSTITTQEAANIIVGYGEWPSYCSDSDATA
EV-D68 1 SPSAEACGYSDRVLQLKLGNSAIVTQEAANYCCAYGEWPNYLPDHEAVA
EV-F3 48 VDKPTQPGVSAERFYTLPSVQWTNSFKGHYWKLPDALSELGLFGQNLQF
EV-E1 50 VDKPTEPGVSADRFYTLRSKPWQADSKGWYWKLPDALNNTGMFGQNAQF
PV1-Mahoney 50 VDQPTEPDVAACRFYTLDTVSWTKESRGWWWKLPDALRDMGLFGQNMYY
EV-A71 41 VDKPTRPDVSVNRFYTLDTKLWEKSSKGWYWKFPDVLTETGVFGQNAQF
EV-D68 50 IDKPTQPETSTDRFYTLRSVKWESNSTGWWWKLPDALNNIGMFGQNVQY
EV-F3 97 HYLYRGGWVIHVQCNATKFHQGTLLVVATPEHKIQS.....AESPA
EV-E1 99 HYLYRGGWAVHVQCNATKFHQGTLLVLAIPEHQIAT.....QEQPA
PV1-Mahoney 99 HYLGRSGYTVHVQCNASKFHQGALGVFAVPEMCLA...GDSNTTTMHTS
EV-A71 90 HYLYRSGFCIHVQCNASKFHQGALLVAVLPEYVIGTVAGGTGTEDTHPP
EV-D68 99 HYLYRSGFLIHVQCNATKFHQGALLVVAIPEHQRGAH.....DTTTSPG
EV-F3 138 FARTNPGEQGAAYQFP.....FTFEDGTALGNALIY
EV-E1 140 FDRTMPGSEGGTFQEP.....FWLEDGTSLGNSLIY
PV1-Mahoney 145 YQNANPGEKGGTFTGTFTPDNNQTSPARRFCPVDYLLGNGTLLGNAFVF
EV-A71 139 YKQTQPGADGFELQHP.....YVLDAGIPISQLTVC
EV-D68 143 FNDIMKGERGGTFNHP.....YVLDDGTSIACATIF
EV-F3 169 PHQWVNLRTNNSATLVLPYVNALPMDSGIRHNNWTLSVIPIVPLEYAAG
EV-E1 171 PHQWINLRTNNSATLILPYVNAIPMDSAIRHSNWTLAIIPVAPLKYAAE
PV1-Mahoney 194 PHQIINLRTNNCATLVLPYVNSLSIDSMVKHNNWGIAILPLAPLNFASE
EV-A71 170 PHQWINLRTNNCATIIVPYINALPFDSALNHCNFGLLVVPISPLDYDQG
EV-D68 174 PHQWINLRTNNSATIVLPWMNVAPMDFPLRHNQWTLAVIPVVPLGTRT.
EV-F3 218 ATTYVPITVTIAPMCTEYNGLRAAVTQ...
EV-E1 220 TTPLVPITVTIAPMETEYNGLRRAIASNQ.
PV1-Mahoney 243 SSPEIPITLTIAPMCCEFNGLRNITLPRLQ
EV-A71 219 ATPVIPITITLAPMCSEFAGLRQAVTQ...
EV-D68 222 MSSVVPITVSIAPMCCEFNGLRHAITQ...

VP3



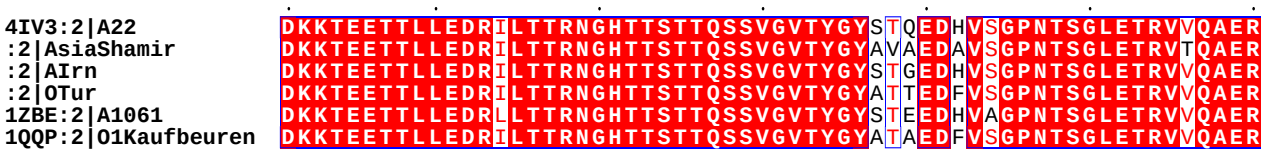
FMDV Sequence alignment

VP1



VP2

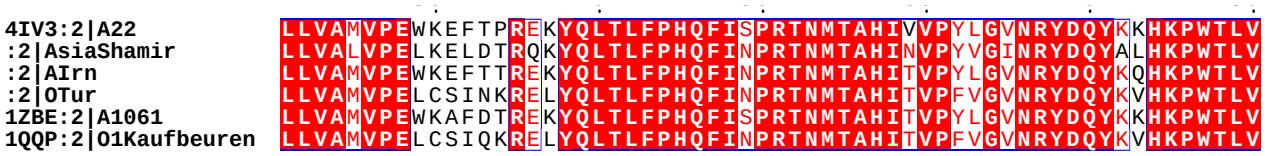
4IV3:2|A22



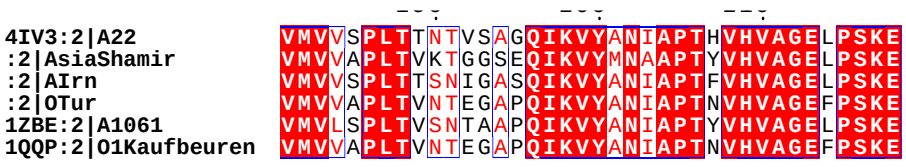
4IV3:2|A22



4IV3:2|A22



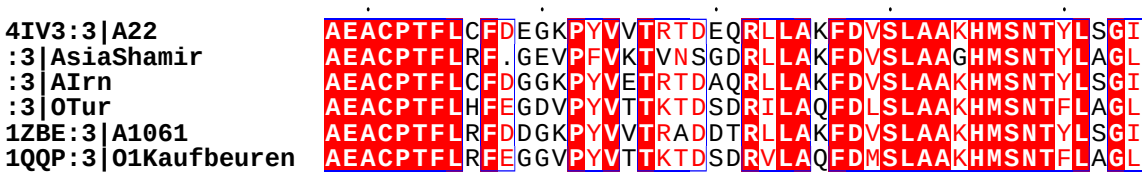
4IV3:2|A22



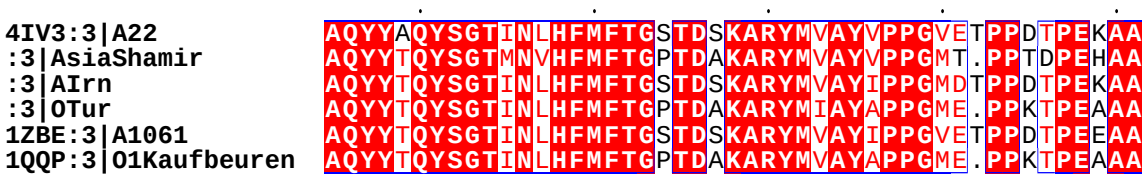
VP3

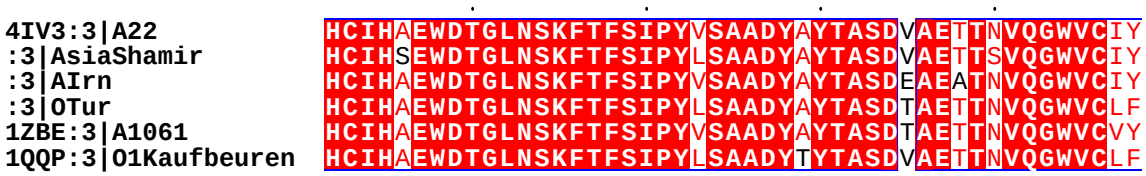
4IV3:3|A22

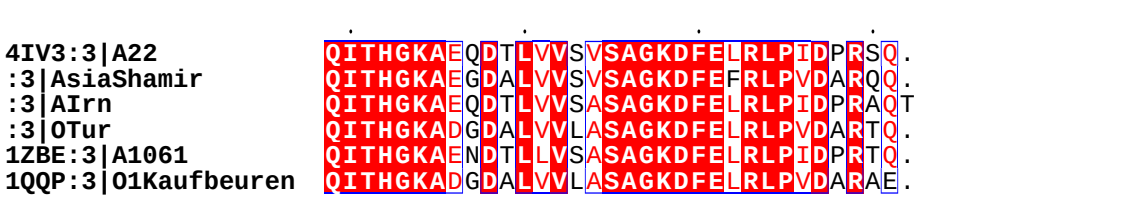






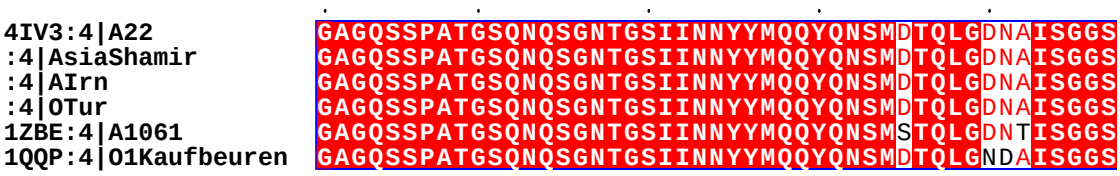






VP4

4IV3:4|A22

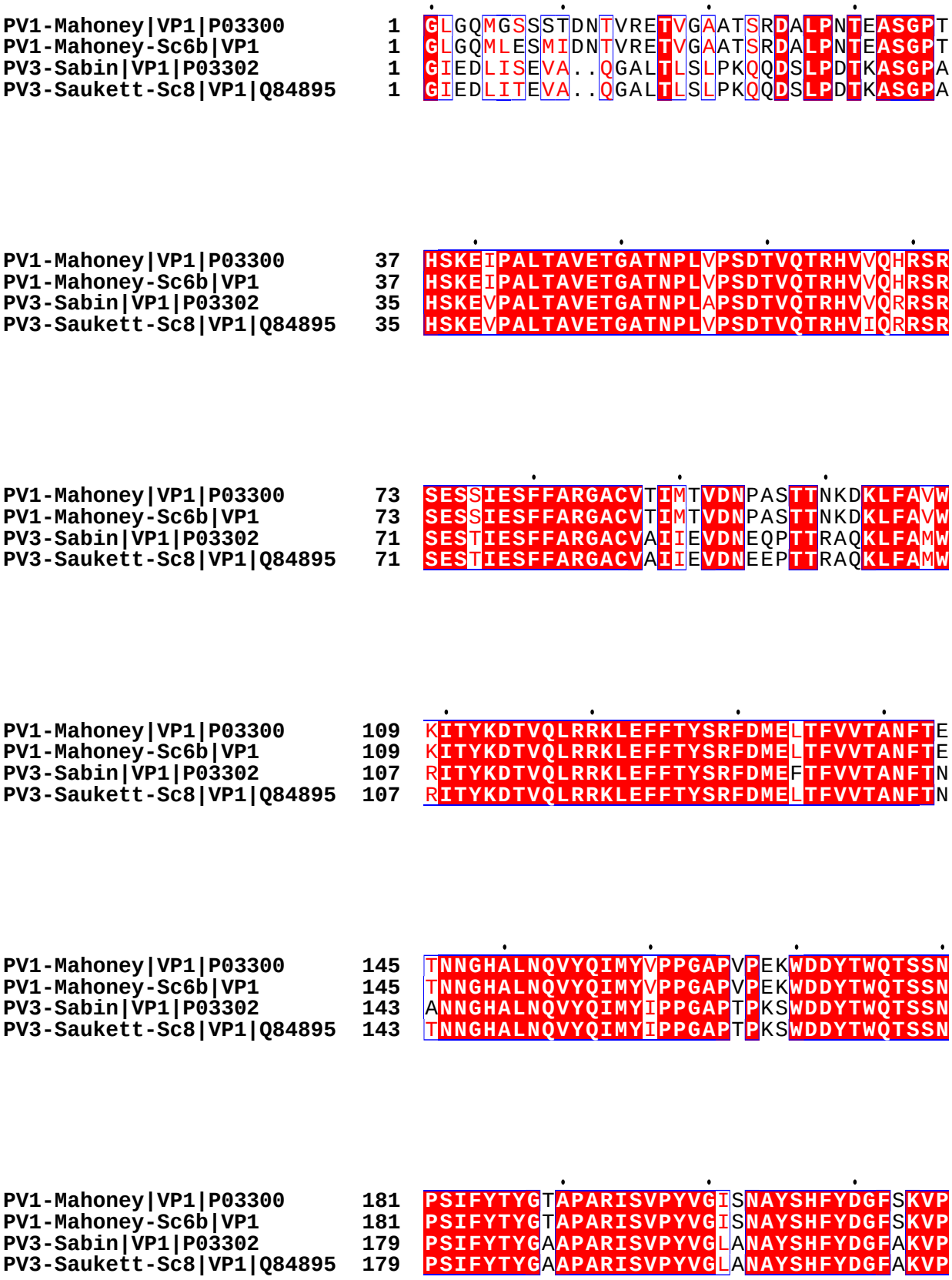


4IV3:4|A22

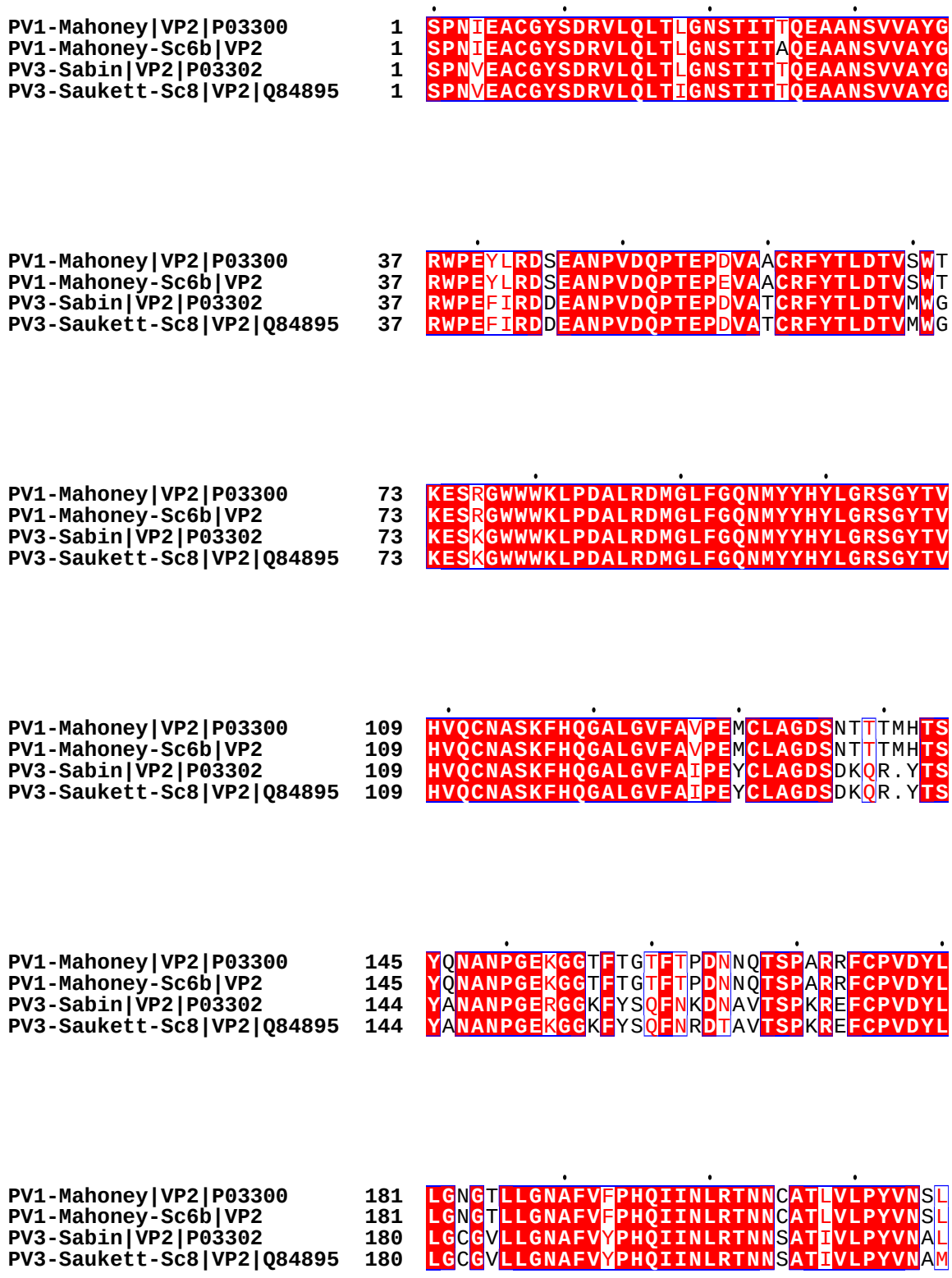


PV Sequence alignment

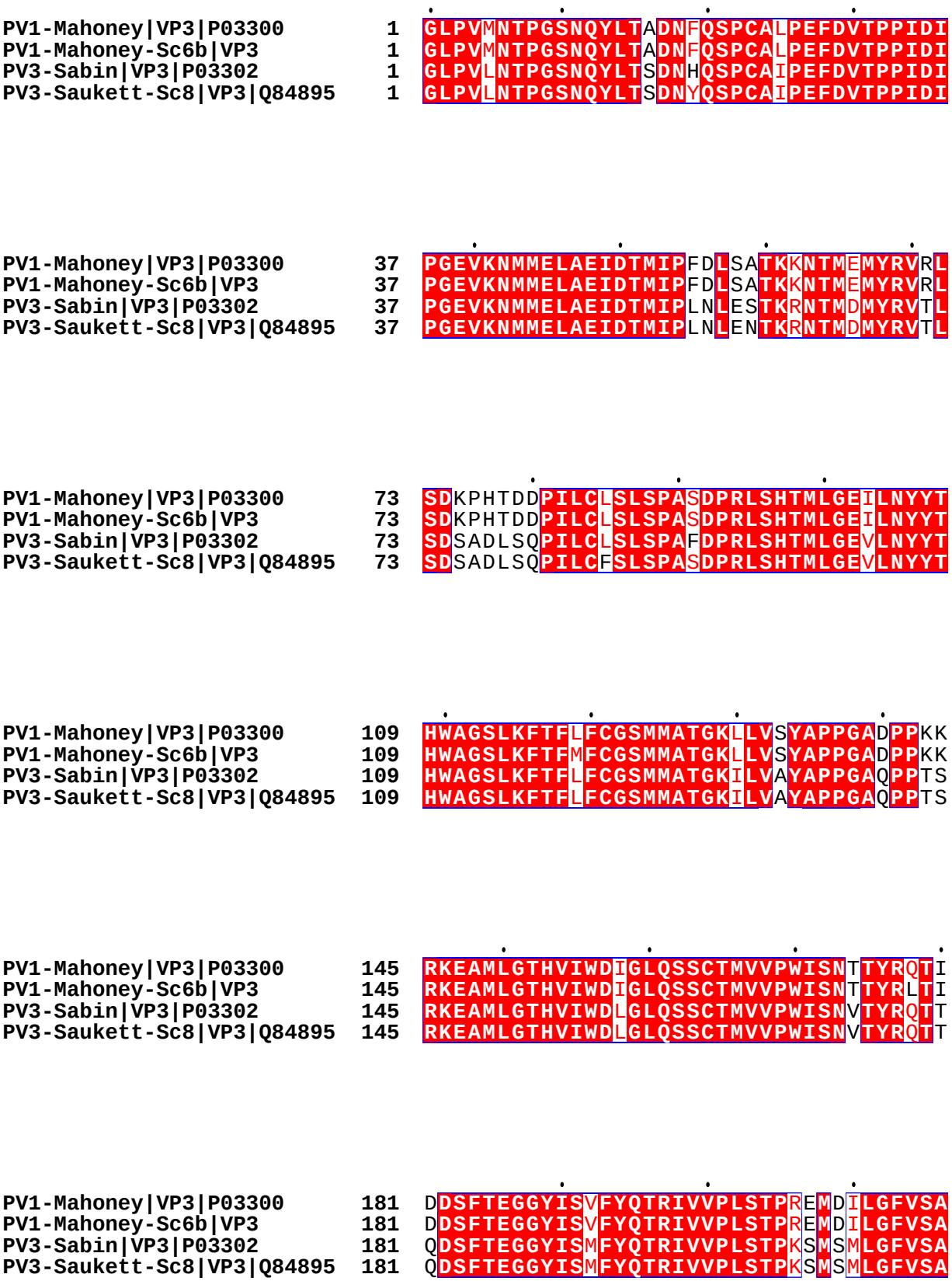
VP1



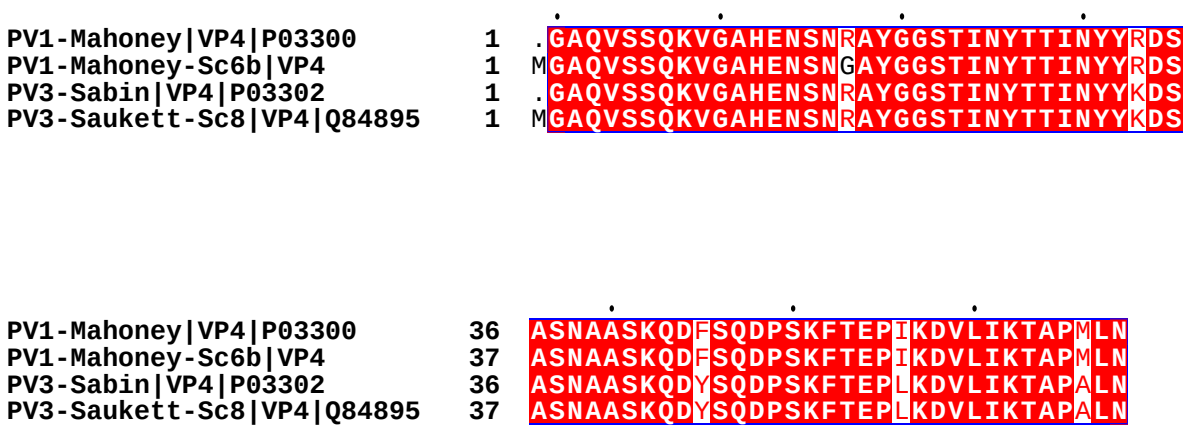
VP2


PV1-Mahoney|VP2|P03300 1 SPNIEACGYSDRVLQLTLGNSTITTQEAANSVVAYG
PV1-Mahoney-Sc6b|VP2 1 SPNIEACGYSDRVLQLTLGNSTITAQEAANSVVAYG
PV3-Sabin|VP2|P03302 1 SPNVEACGYSDRVLQLTLGNSTITTQEAANSVVAYG
PV3-Saukett-Sc8|VP2|Q84895 1 SPNVEACGYSDRVLQLTIGNSTITTQEAANSVVAYG
PV1-Mahoney|VP2|P03300 37 RWPEYLRDSEANPVDQPTEPDVAACRFYTLDTVSWT
PV1-Mahoney-Sc6b|VP2 37 RWPEYLRDSEANPVDQPTEPEVAACRFYTLDTVSWT
PV3-Sabin|VP2|P03302 37 RWPEFIRDDEANPVDQPTEPDVATCRFYTLDTVMWG
PV3-Saukett-Sc8|VP2|Q84895 37 RWPEFIRDDEANPVDQPTEPDVATCRFYTLDTVMWG
PV1-Mahoney|VP2|P03300 73 KESRGWWWKLPDALRDMGLFGQNMYYHYLGRSGYTV
PV1-Mahoney-Sc6b|VP2 73 KESRGWWWKLPDALRDMGLFGQNMYYHYLGRSGYTV
PV3-Sabin|VP2|P03302 73 KESKGWWWKLPDALRDMGLFGQNMYYHYLGRSGYTV
PV3-Saukett-Sc8|VP2|Q84895 73 KESKGWWWKLPDALRDMGLFGQNMYYHYLGRSGYTV
PV1-Mahoney|VP2|P03300 109 HVQCNASKFHQGALGVFAVPEMCLAGDSNTTTMHTS
PV1-Mahoney-Sc6b|VP2 109 HVQCNASKFHQGALGVFAVPEMCLAGDSNTTTMHTS
PV3-Sabin|VP2|P03302 109 HVQCNASKFHQGALGVFAIPEYCLAGDSDKQR.YTS
PV3-Saukett-Sc8|VP2|Q84895 109 HVQCNASKFHQGALGVFAIPEYCLAGDSDKQR.YTS
PV1-Mahoney|VP2|P03300 145 YQNANPGEKGGTFTGTFTPDNNQTSPARRFCPVDYL
PV1-Mahoney-Sc6b|VP2 145 YQNANPGEKGGTFTGTFTPDNNQTSPARRFCPVDYL
PV3-Sabin|VP2|P03302 144 YANANPGERGGKFYSQFNKDNAVTSPKREFCPVDYL
PV3-Saukett-Sc8|VP2|Q84895 144 YANANPGEKGGKFYSQFNRDTAVTSPKREFCPVDYL
PV1-Mahoney|VP2|P03300 181 LGNGTLLGNAFVFPHQIINLRTNNCATLVLPYVNSL
PV1-Mahoney-Sc6b|VP2 181 LGNGTLLGNAFVFPHQIINLRTNNCATLVLPYVNSL
PV3-Sabin|VP2|P03302 180 LGCGVLLGNAFVYPHQIINLRTNNSATIVLPYVNAL
PV3-Saukett-Sc8|VP2|Q84895 180 LGCGVLLGNAFVYPHQIINLRTNNSATIVLPYVNAM

VP3


PV1-Mahoney|VP3|P03300 1 GLPVMNTPGSNQYLTADNFQSPCALPEFDVTPPIDI
PV1-Mahoney-Sc6b|VP3 1 GLPVMNTPGSNQYLTADNFQSPCALPEFDVTPPIDI
PV3-Sabin|VP3|P03302 1 GLPVLNTPGSNQYLTSDNHQSPCAIPEFDVTPPIDI
PV3-Saukett-Sc8|VP3|Q84895 1 GLPVLNTPGSNQYLTSDNYQSPCAIPEFDVTPPIDI
PV1-Mahoney|VP3|P03300 37 PGEVKNMMELAEIDTMIPFDLSATKKNTMEMYRVRL
PV1-Mahoney-Sc6b|VP3 37 PGEVKNMMELAEIDTMIPFDLSATKKNTMEMYRVRL
PV3-Sabin|VP3|P03302 37 PGEVKNMMELAEIDTMIPLNLESTKRNTMDMYRVTL
PV3-Saukett-Sc8|VP3|Q84895 37 PGEVKNMMELAEIDTMIPLNLENTKRNTMDMYRVTL
PV1-Mahoney|VP3|P03300 73 SDKPHTDDPILCLSLSPASDPRLSHTMLGEILNYYT
PV1-Mahoney-Sc6b|VP3 73 SDKPHTDDPILCLSLSPASDPRLSHTMLGEILNYYT
PV3-Sabin|VP3|P03302 73 SDSADLSQPILCLSLSPAFDPRLSHTMLGEVLNYYT
PV3-Saukett-Sc8|VP3|Q84895 73 SDSADLSQPILCFSLSPASDPRLSHTMLGEVLNYYT
PV1-Mahoney|VP3|P03300 109 HWAGSLKFTFLFCGSMMATGKLLVSYAPPGADPPKK
PV1-Mahoney-Sc6b|VP3 109 HWAGSLKFTFMFCGSMMATGKLLVSYAPPGADPPKK
PV3-Sabin|VP3|P03302 109 HWAGSLKFTFLFCGSMMATGKILVAYAPPGAQPPTS
PV3-Saukett-Sc8|VP3|Q84895 109 HWAGSLKFTFLFCGSMMATGKILVAYAPPGAQPPTS
PV1-Mahoney|VP3|P03300 145 RKEAMLGTHVIWDIGLQSSCTMVVPWISNTTYRQTI
PV1-Mahoney-Sc6b|VP3 145 RKEAMLGTHVIWDIGLQSSCTMVVPWISNTTYRLTI
PV3-Sabin|VP3|P03302 145 RKEAMLGTHVIWDLGLQSSCTMVVPWISNVTYRQTT
PV3-Saukett-Sc8|VP3|Q84895 145 RKEAMLGTHVIWDLGLQSSCTMVVPWISNVTYRQTT
PV1-Mahoney|VP3|P03300 181 DDSFTEGGYISVFYQTRIVVPLSTPREMDILGFVSA
PV1-Mahoney-Sc6b|VP3 181 DDSFTEGGYISVFYQTRIVVPLSTPREMDILGFVSA
PV3-Sabin|VP3|P03302 181 QDSFTEGGYISMFYQTRIVVPLSTPKSMSMLGFVSA
PV3-Saukett-Sc8|VP3|Q84895 181 QDSFTEGGYISMFYQTRIVVPLSTPKSMSMLGFVSA

VP4


PV1-Mahoney|VP4|P03300 1 .GAQVSSQKVGAHENSNRAYGGSTINYTTINYYRDS
PV1-Mahoney-Sc6b|VP4 1 MGAQVSSQKVGAHENSNGAYGGSTINYTTINYYRDS
PV3-Sabin|VP4|P03302 1 .GAQVSSQKVGAHENSNRAYGGSTINYTTINYYKDS
PV3-Saukett-Sc8|VP4|Q84895 1 MGAQVSSQKVGAHENSNRAYGGSTINYTTINYYKDS
PV1-Mahoney|VP4|P03300 36 ASNAASKQDFSQDPSKFTEPIKDVLIKTAPMLN
PV1-Mahoney-Sc6b|VP4 37 ASNAASKQDFSQDPSKFTEPIKDVLIKTAPMLN
PV3-Sabin|VP4|P03302 36 ASNAASKQDYSQDPSKFTEPLKDVLIKTAPALN
PV3-Saukett-Sc8|VP4|Q84895 37 ASNAASKQDYSQDPSKFTEPLKDVLIKTAPALN

SCIENTIFIC REPORTS

OPEN

Towards *in cellulo* virus crystallography

Helen M. E. Duyvesteyn[1,2], Helen M. Ginn[1,2], Maija K. Pietilä[3,5], Armin Wagner[2], Johan Hattne[4,6], Jonathan M. Grimes[1,2], Elina Hirvonen[3], Gwyndaf Evans[2], Marie-Laure Parsy[1], Nicholas K. Sauter[4], Aaron S. Brewster[4], Juha T. Huiskonen[1,3,7], David I. Stuart[1,2], Geoff Sutton[1] & Dennis H. Bamford[3]

Received: 23 October 2017
Accepted: 30 January 2018
Published online: 28 February 2018

Viruses are a significant threat to both human health and the economy, and there is an urgent need for novel anti-viral drugs and vaccines. High-resolution viral structures inform our understanding of the virosphere, and inspire novel therapies. Here we present a method of obtaining such structural information that avoids potentially disruptive handling, by collecting diffraction data from intact infected cells. We identify a suitable combination of cell type and virus to accumulate particles in the cells, establish a suitable time point where most cells contain virus condensates and use electron microscopy to demonstrate that these are ordered crystalline arrays of empty capsids. We then use an X-ray free electron laser to provide extremely bright illumination of sub-micron intracellular condensates of bacteriophage phiX174 inside living *Escherichia coli* at room temperature. We have been able to collect low resolution diffraction data. Despite the limited resolution and completeness of these initial data, due to a far from optimal experimental setup, we have used novel methodology to determine a putative space group, unit cell dimensions, particle packing and likely maturation state of the particles.

Whilst viral complexity and size often limits the growth of crystals suitable for classical crystallographic characterisation[1], recent advances in electron cryo-microscopy (cryo-EM) have opened another route to high resolution structure determination[2–4]. Alternatively, X-ray free electron lasers (XFELs) offer less stringent crystal size requirements; using femtosecond-duration pulses that are over a billion times brighter than synchrotron radiation[5]. This has recently allowed the determination of a high-resolution structure from virus microcrystals[6].

Although crystals are conventionally grown *in vitro*, the production of protein crystals *in vivo* may be achieved via numerous biological processes[7]. The relatively small size of *in vivo* crystals means that useful diffraction data can be collected only at microfocus synchrotron beamlines or XFELs[8–10]. Collecting diffraction data *in cellulo* minimises potential mechanical damage of fragile crystals and so has the potential for improving the quality of diffraction, however the signal-to-noise of the diffraction data will be adversely affected by the presence of extraneous cellular material[11,12]. Thus, for large unit cells and small crystals the intensities of the Bragg peaks will be dramatically reduced (for instance, similar sized crystals of small picornaviruses will have average intensities hundreds of times less than for lysozyme), whereas the background noise increases in proportion to the amount of extraneous material illuminated. Therefore, very careful experimental design will be needed to obtain useful measured diffraction intensities.

Most bacteriophages terminate an infection cycle by host cell lysis, which limits the accumulation of virus particles within the cell. However, there are some mutant cell lines available which are resistant to lysis by certain bacteriophages. For instance, mutation of the host gene *slyD* (sensitivity to lysis), which encodes peptidyl-prolyl cis-trans isomerase, can block lysis by phiX174, precluding viral escape[13,14]. We have therefore chosen to

[1]Division of Structural Biology, University of Oxford, The Henry Wellcome Building for Genomic Medicine Headington, Oxford, UK. [2]Diamond Light Source, Harwell Science and Innovation Campus, Didcot, UK. [3]Molecular and Integrative Biosciences Research Program, Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki, Finland. [4]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, USA. [5]Present address: Department of Microbiology, Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland. [6]Present address: Howard Hughes Medical Institute, Janelia Research Campus, Ashburn, VA, USA. [7]Laboratory of Structural Biology, Helsinki Institute of Life Science, University of Helsinki, Helsinki, Finland. Helen M. E. Duyvesteyn, Helen M. Ginn and Maija K. Pietilä contributed equally to this work. Correspondence and requests for materials should be addressed to D.I.S. (email: dave@strubi.ox.ac.uk) or D.H.B. (email: dennis.bamford@helsinki.fi)

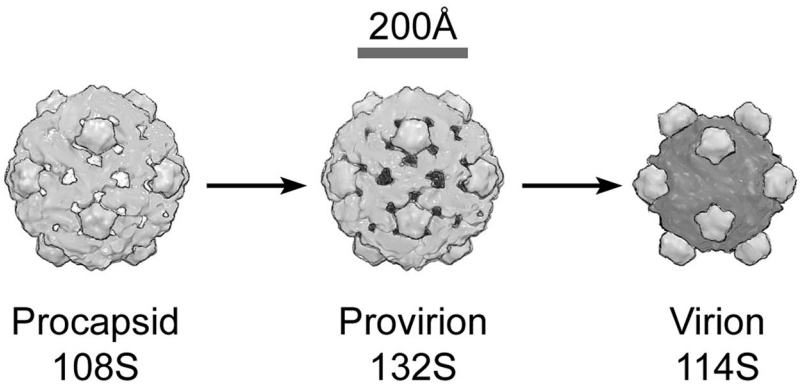


Figure 1. Schematic of the stages in the assembly and maturation of phiX174, a T = 1 bacteriophage. The procapsid (pdb 1cd3[16]) contains proteins B, D, F, G & H and develops into the provirion[20] by losing at least some of the B proteins and gaining protein J and DNA. Subsequent formation of the mature virion (pdb 2bpa[18]) occurs through the loss of D proteins and the remaining B proteins.

investigate the structure of phiX174 within infected *E. coli slyD* cells as a proof of principle for what could be developed into a more general vehicle.

PhiX174′s lifecycle encompasses three distinct states (Fig. 1), all of which are icosahedral and have been studied, *in vitro*, by crystallography[15–19] and electron microscopy[20,21]. The first stage is assemblage into a procapsid (108 S) form, which, upon loss of the B scaffolding protein, and gain of a J protein, concomitant with packaging of the ssDNA genome, forms the 132 S intermediate provirion, which is of a similar size[16,17,22–24]. Maturation to the virion (114 S) occurs with loss of the D scaffolding protein, and a significant collapse of the structure[22,24]. We report analysis of cells infected with both wild-type (wt) virus and *AmbJ*$^-$ mutant virus. The latter are unable to package DNA, since they do not possess the packaging protein, J. Using electron microscopy (EM) and X-ray diffraction, we demonstrate that small crystalline arrays form within cells. The experimental setup-up available, whilst unable to fully explore the potential of the method, nonetheless provided useful information, demonstrating the arrest of PhiX174 maturation at the procapsid state.

Results

phiX174 forms crystalline condensates in most cells 4.5 h post infection. Thin-section EM of *Escherichia coli* C990 *slyD* infected with both wt and *AmbJ*$^-$ phiX174 revealed the presence intracellular viral condensates in lysis-defective conditions[14] (see Methods), the result of particle accumulation (Fig. 2a). For example, by 1.5 h post infection, aggregates of wt virus particles were seen in some 50% of the cells. By 4.5 h post infection, these had developed so that approximately 75% of cells contained condensates of empty particles, of which roughly one third appeared to be crystalline. We therefore selected 4.5 h post infection as the time point for further analysis.

Electron Cryo-Microscopy indicates the crystal lattice is reasonably well-ordered. The crystallinity of the crystal-like arrays seen at 4.5 h post infection (for both wt and *AmbJ*$^-$ virus) was then visualised using whole-cell three-dimensional tomographic reconstructions and sub-tomogram averaging. The condensates, *circa* $200 \times 200 \times 200\,nm^3$ in volume, nearly filled the cell interior. Power spectra calculated from 2D projection images of cells confirmed that the condensates possessed crystalline order up to a resolution of ~52 Å (Fig. 2b,c). It appears that a major spacing of the lattice is around 340 Å. Sub-tomogram averaging suggested that the capsids were ordered on a lattice and were empty (Fig. 2d).

XFEL analysis of infected cells. Cells were grown and infected adjacent to the Linac Coherent Light Source (LCLS), at Stanford University. A slurry of 4.5 h post infection cells was pumped through a gas dynamic virtual nozzle (GDVN)[25] at room temperature into the X-ray beam at the LCLS CXI beamline[5,26]. Diffraction patterns were collected on a CSPAD detector[27] ~2.5 m from the interaction region of the beam with the jet to record low angle diffraction from condensates.

The majority of data were collected on wt phiX174 at a wavelength of 1.768 Å, while approximately 30% of the dataset was taken from the *ambJ*$^-$ mutant. Data for this latter sample were collected at both 1.768 and 1.306 Å wavelengths. At these wavelengths, the maximal resolutions at the detector edge were approximately 46 and 34 Å, respectively.

At the time the experiment was performed the beam and jet were larger than would have been ideal for this experiment. Ideally one would match the beam and the jet to the crystal size, whereas in practice these had diameters of ~2 and 4 μm, respectively. Nevertheless, Bragg diffraction was observed, but limited to ~50 Å spacings (Fig. 3a).

Initial XFEL Parameters. Based on calibrations from thermolysin data collected shortly before phiX174[28], the detector distance was estimated to be 2.49 m. Note that this is dependent on accurate estimates of the thermolysin unit cell[29] and wavelength calibration at LCLS, neither of which were known to a confidence higher than 1%.

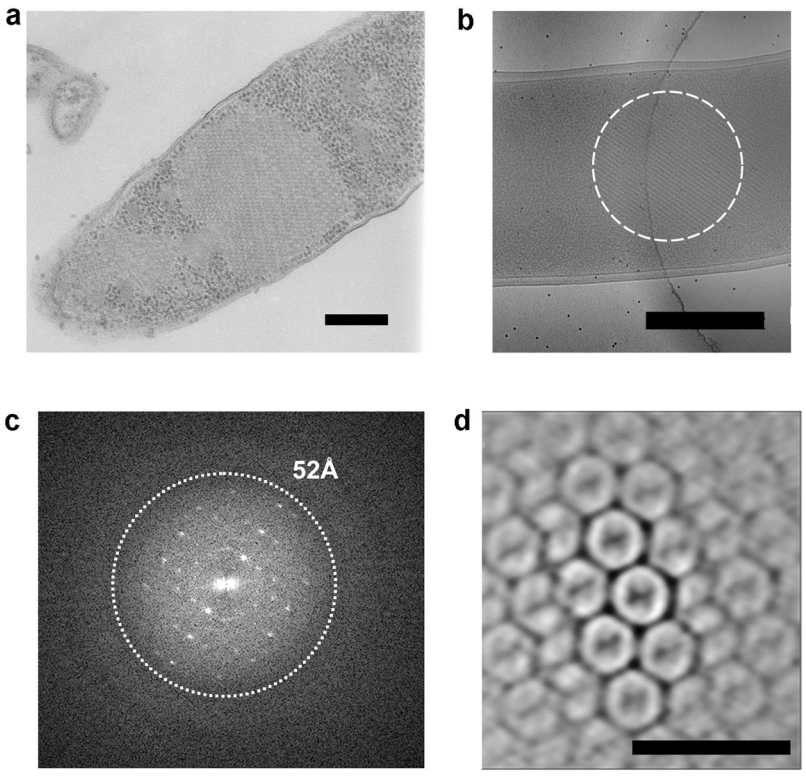


Figure 2. Electron microscopy. (**a**) Ultrathin section of a *E. coli* C990 *slyD1* cell infected with wt phiX174 at 4.5 h p.i. Scale bar shows 200 nm. (**b**) Two-dimensional projection image of similarly infected *E. coli* C990 *slyD1* cell. Scale bar shows 500 nm. (**c**) Fourier transform of the area shown in (**b**). (**d**) Section through averaged tomogram density. Scale bar shows 100 nm.

The two wavelengths of 1.768 and 1.306 Å were initially derived from LCLS log files. Wavelengths may, however, be taken on a per image basis (Supplementary Fig. 1). The shot-to-shot variation in wavelength is much less that the likely errors in distance, and reprocessing some of the data using these unique per shot energies gave no significant improvement, therefore all further analysis used the overall wavelengths from the log file.

Very low crystal-to-beam hit rates and highly intense artifacts from flaring, a likely consequence of reflection of the incident beam from the edge of the liquid jet (see Fig. 3a), precluded the use of conventional spot finding algorithms even with stringent masking. 168,456 images were initially triaged on the basis of the compressed image size. The extent of jpeg compression is related to information content in the image and we observed that the 3.1 M pixel images compressed to a wide range of sizes (0.124 to 0.421 MB for 2×2 binned images). Inspection of a small subset of images revealed that those showing diffraction were <0.3 MB. Using this criteria we were able to immediately eliminate 72% of the images (examination of a substantial subset of these revealed that no diffraction patterns had been missed), whereas the triaged images for the wt sample showed a success rate of 7.2% when screened manually. For the ~680 diffraction patterns identified the coordinates of spots were recorded by hand. Spot positions on the detector were mapped into reciprocal space and pair-wise vectors between all observed spots on each image were calculated. By generating a histogram of these inter-spot distances, a 'pseudo-powder pattern' was formed, as described in Ginn 2016[30] (shown in light blue on Fig. 3b). There was no significant difference in the diffraction between the wt and mutant particles.

Based on the apparent high symmetry indicated by the sparse powder pattern a cubic space group seemed likely (this was consistent with the condensates seen in the cryo-ET reconstructions). In addition, as will emerge in the discussion below, although we strictly need to observe symmetry in the diffraction intensities to confirm the cubic space group, a cubic space group is consistent with elements of the inherent symmetry of the virus particle and the alignment of particle and crystal symmetry elements is a common feature of most virus crystals. For simplicity, we therefore assume that this is the case here. Thus, a systematic fitting of allowed face-centred (F), body-centred (I) and primitive- (P) space groups was performed (Fig. 3b). The powder pattern facilitates the identification of centering operations, since successively more reflections are systematically absent for P, I and F lattices. For a primitive lattice, all reflections can be of non-zero intensity (shown in red in Fig. 3b). For a body centered lattice the sum of the three components of the indices must be even (shown with dark blue hatching) for non-zero intensity and for face centering all three pair-wise sums of the components must be even (shown with green hatching and a star). This produced a convincing best fit for a face-centred cubic space group, with an approximate cell edge of 500 Å (see Supplementary Fig. 2). Assuming a cubic lattice this limits the possible space groups which can accommodate a chiral virus particle to F*23*, F*432* and F4_132. The highest subgroup of icosahedral symmetry that can be accommodated in a crystal lattice is 23 symmetry. All of the putative space groups

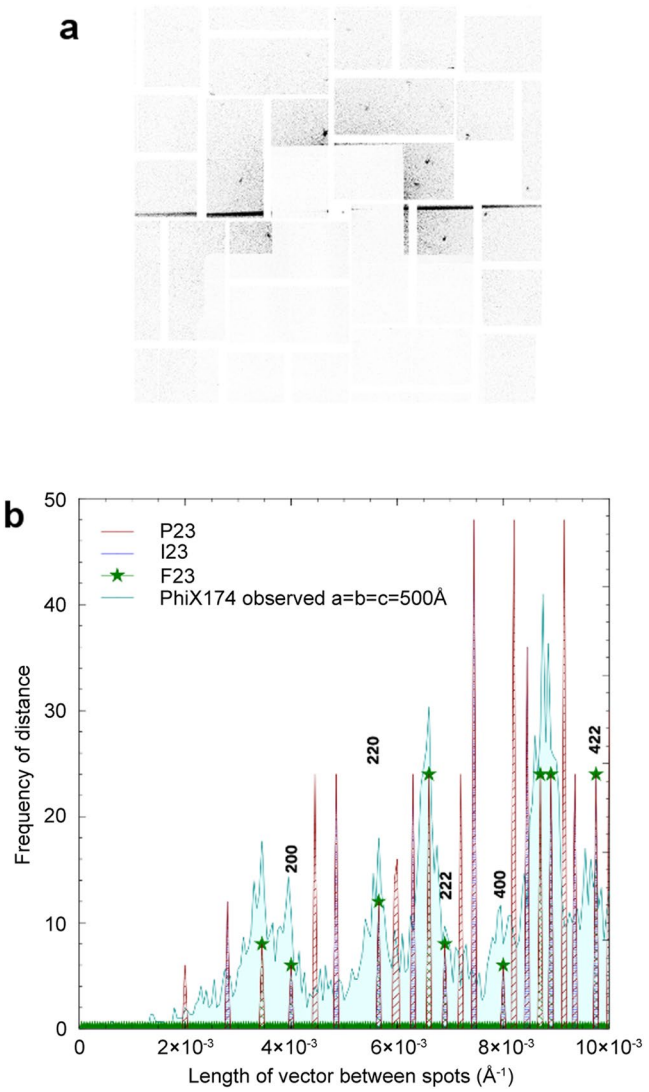


Figure 3. Diffraction data and analysis. (**a**) Example of a strong diffraction pattern from phiX174 wild-type with a clear lattice. Panel shadowing was a consequence of a silicon support used to protect the detector. The black streak was observed on many patterns, although at different angles and is presumably derived from the beam reflecting from the jet edge. (**b**) Histogram of vector distances illustrating selection process of correct cubic space group. Blue fill peaks correspond to combined data from wild type and mutant data sets. Patterned peaks show predicted frequency of vector distances for face- (F), body- (I) and primitive- (P) centred structures. The closest match to our observed data (pale blue fill, with blue outline) is the face-centred cubic space group (green stars).

contain points of 23 symmetry, F*23* has four such points whilst F*432* and F4_132 each have eight. Even with the most economical packing, only F*23* can accommodate a particle of the size of phiX174 on a point of 23 symmetry, since four, but not eight, particles can be accommodated in a cell of this size (detailed packing arguments discussed below). We therefore conclude that the most likely space group of the viral arrays grown in infected cells is F*23*, with a = b = c = ~500 Å. Note that purified viral particles and procapsids have been crystallised previously, however none of these crystals are isomorphous with those grown in infected cells (the calculated powder patterns agree poorly, Supplementary Fig. 3).

Putative Maturation State. Since the diffraction from the wt and *AmbJ*$^-$ was indistinguishable and protein J is required for genomic packaging, it seemed likely that our sample is that of a procapsid (furthermore in electron micrographs the condensates appeared to be empty particles, Fig. 2). It is still possible that the condensates represent the mature virion, but in an empty form. Spontaneous maturation in some viral procapsids has been observed *in vitro*[16,31–33], and most notably, for the homologous bacteriophage G4 (sequence identity of 40–66%), where both conformational change, and external scaffolding led to the formation of an empty 'mature virion'[34].

Knowing the space group and unit cell dimension, we can calculate the separation of the centres of neighbouring virus particles (*e.g.* at 0, 0, 0 and 0, ½, ½) as ~353 Å. From published data, the minimum/maximum

Property	Maturation State	
	Procapsid	Mature virus
PDB Code	1cd3[16]	2bpa[18]
Space Group	$I2_13$	$P2_1$
Unit cell dimensions		
a, b, c/Å;	774.00, 774.00, 774.0	305.58, 360.78, 299.4
α, ß, γ/°	90.00, 90.00, 90.00	90.00, 92.89, 90.00
Outer Radius/Å	176	171
Outer Diameter/Å	352	342
Outer Spherical Volume/Å^3	22, 836, 345.90	20, 944, 834.90

Table 1. Summary published crystal data for phiX174.

diameters of the virion and procapsid are 300/342 and 348/352 Å respectively (Table 1). This strongly suggests that the in-cell crystals contain procapsids. Although there are some experimental uncertainties, the true unit cell dimension will lie well within the range of 480 to 520 Å (which correspond to the extreme distances and wavelengths within the experimental uncertainty). Within this range, the only sensible packing for either a procapsid or mature virus is for the procapsid with a cell dimension close to 500 Å, thus in the best packing for a mature virion (480 Å cell) the closest approach of sidechains is about 20 Å (Supplementary Fig. 4).

Discussion

We have established a routine procedure to produce cells, 75% of which contain phiX174 viral condensates. We have also demonstrated that these intracellular viral condensates form ordered sub-micron crystals that are capable of diffracting X-rays to produce informative patterns at XFELs. This represents a proof of principle for the extension of XFEL crystallography to systems with orders of magnitude lower signal than previously reported, and lends promise to further exploration of generic methods for viral structural analysis to bypass traditional crystallisation and handling. This holds potential for difficult-to-crystallise viral targets.

There is a clear route to improving the signal-to-noise for the diffraction data. At the time the experiments were performed, the beam and jet were larger than ideal. Assuming that the beam could be trimmed to the maximal crystal cross-section within the X-ray beam (~$0.4 \times 0.4\,\mu m^2$ supposing maximum crystal dimensions are approximately twice those observed on average by EM), compared to the beam used here (which intersects approximately $2 \times 2\,\mu m^2$ of jet) a 25-fold improvement in signal-to-noise could be obtained. Furthermore, if the jet could be thinned from ~4 μm to the dimensions of a cell, ~1 μm, the total potential gain would be ~100-fold. Although the bacterial cells will scatter somewhat more strongly than the liquid jet, this is probably not an unreasonable estimate. Finally, focussing the entire beam onto the crystal would increase the signal ~25-fold and the signal-to-noise by a further factor of five. Although extraordinarily challenging with current liquid jet technology, since the hit-rate will decline precipitously, solid-phase supports, offering high speed and precision, may be used instead to recover much of this potential 500-fold gain[35,36]. A significant improvement in hit rate may also be achieved with the recently reported Acoustic Droplet Ejection technology[37].

If it is possible to extend the applicability of this method, and render it a generic vehicle for icosahedral, and perhaps any non-pleomorphic, virus particles, *in cellulo* crystallisation might provide new targets that may only be crystallised within intracellular conditions. It remains to be seen whether such advances provide a method competitive with electron microscopy for large complexes, but at the very least they may offer a synergistic tool.

Methods

Sample preparation. Cultures were grown in Luria broth. Lysis-defective cells of logarithmic *Escherichia coli* C990 *slyD1* cultures were infected with phix174 using a multiplicity of infection of 10 and $CaCl_2$ and $MgCl_2$ were added to final concentrations of 5 and 10 mM, respectively. After infection, cells were incubated for 5 min without aeration and then aerobically at 37 °C. For thin-section EM, cells were fixed 4.5 h post infection (p.i.) with 3% (v/v) glutaraldehyde and sections were prepared as previously described[38]. The micrographs were taken with JEOL1200EX electron microscope operating at 80 kV.

Electron cryo-microscopy and tomography. For electron cryo-microscopy and tomography, unfixed infected cells and 10 nm gold clusters were added to an electron microscopy grid 4.5 h p.i. and the grid was plunge-frozen in a liquid ethane-propane mixture. Data were collected on a Tecnai Polara electron microscope (Thermo Fischer) operating at 300 kV at liquid nitrogen temperature and equipped with a 4k × 4k CCD camera (Gatan) mounted behind an energy filter (Gatan) operating at zero-loss mode (20 eV slit). For two-dimensional imaging, a series of 40 images (total dose of 80 $e^-Å^{-2}$) was acquired at 12 μm under focus and images were computationally aligned in IMOD[39]. Images were high-pass filtered to 10000 $Å^{-1}$ to remove low frequency features arising from variations in sample thickness. For three-dimensional tomographic reconstruction of viral aggregates, a tilt series was collected from −60° to +60° in 4° increments at 12 μm under focus, with a maximum dose of 70 $e^-Å^{-2}$ in SerialEM[40] and reconstructed in IMOD using the gold clusters as fiducial markers. Template matching and sub-tomogram averaging of ordered condensate regions was performed in Jsubtomo[41,42], taking into account the missing wedge of tomographic reconstructions. Averages were low pass filtered to 80 $Å^{-1}$ for visualization.

Serial femtosecond crystallography. Diffraction data were collected at the Linac Coherent Light Source (LCLS) of SLAC National Accelerator Laboratory (Menlo Park, CA, USA) on the Coherent X-ray Imaging (CXI) beamline[26]. Intact cells were injected at room temperature using a Gas Dynamic Virtual Nozzle[25] at a flow rate of 50 µl min^{-1}. X-ray pulses were of 50 fs duration with 2.0×10^{12} photons per pulse and a beam diameter of c.a. 2 µm. In order to calculate the X-ray wavelength, electron energies were converted to photon energies, based on the undulator parameter K[43].

Femtosecond diffraction snapshots were recorded at 120 Hz on the downstream position Cornell-SLAC pixel array (CSPAD) detector[27] at a sample-detector distance estimated to be 2.49 m.

Pseudo-powder pattern generation. Approximately 2000 images with diffraction were derived from the output streams. Diffraction spots were then manually selected from these diffraction images and the corresponding pixel coordinates recorded (Ginn, unpublished software). Vectors between pairs of spots were generated using cppxfel[30], and a histogram showing the frequency of the occurrence of the lengths of these vectors was termed the "pseudo-powder pattern". This method has been described previously[44].

Crystal symmetry determination and packing analysis. Packing of the procapsid (PDB 1cd3)[16] and mature capsid (PDB 2bpa)[18] asymmetric units was visualised using PyMOL[45].

References

1. Cockburn, J., Bamford, J., Grimes, J., Bamford, D. & Stuart, D. Crystallization of the membrane-containing bacteriophage PRD1 in quartz capillaries by vapour diffusion. *Acta Crystallographica Section D: Biological Crystallography* **59**, 538–540 (2003).
2. Bai, X.-C., McMullan, G. & Scheres, S. H. How cryo-EM is revolutionizing structural biology. *Trends in Biochemical Sciences* **40**, 49–57 (2015).
3. Campbell, M. G., Veesler, D., Cheng, A., Potter, C. S. & Carragher, B. 2.8 Å resolution reconstruction of the Thermoplasma acidophilum 20S proteasome using cryo-electron microscopy. *Elife* **4**, e06380 (2015).
4. Amunts, A. *et al.* Structure of the yeast mitochondrial large ribosomal subunit. *Science* **343**, 1485–1489 (2014).
5. Emma, P. *et al.* First lasing and operation of an ångstrom-wavelength free-electron laser. *nature photonics* **4**, 641–647 (2010).
6. Roedig, P. *et al.* High-speed fixed-target serial virus crystallography. *Nat Meth* **14**, 805–810 (2017).
7. Doye, J. P. K. & Poon, W. C. K. Protein crystallization *in vivo*. *Current Opinion in Colloid & Interface Science* **11**, 40–46 (2006).
8. Smith, J. L., Fischetti, R. F. & Yamamoto, M. Micro-crystallography comes of age. *Current Opinion in Structural Biology* **22**, 602–612 (2012).
9. Sawaya, M. R. *et al.* Protein crystal structure obtained at 2.9 Å resolution from injecting bacterial cells into an X-ray free-electron laser beam. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 12769–12774 (2014).
10. Jakobi, A. J. *et al.* In cellulo serial crystallography of alcohol oxidase crystals inside yeast cells. *IUCrJ* **3**, 88–95 (2016).
11. Axford, D., Ji, X., Stuart, D. I. & Sutton, G. In cellulo structure determination of a novel cypovirus polyhedrin. *Acta Crystallographica Section D: Biological Crystallography* **70**, 1435–1441 (2014).
12. Boudes, M., Garriga, D., Fryga, A., Caradoc-Davies, T. & Coulibaly, F. A pipeline for structure determination of *in vivo*-grown crystals using in cellulo diffraction. *Acta Crystallographica. Section D, Structural Biology* **72**, 576–585 (2016).
13. Bernhardt, T. G., Roof, W. D. & Young, R. Genetic evidence that the bacteriophage φX174 lysis protein inhibits cell wall synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 4297–4302 (2000).
14. Bernhardt, T. G., Roof, W. D. & Young, R. The Escherichia coli FKBP-type PPIase SlyD is required for the stabilization of the E lysis protein of bacteriophage φX174. *Molecular microbiology* **45**, 99–108 (2002).
15. Bernal, R. A., Hafenstein, S., Esmeralda, R., Fane, B. A. & Rossmann, M. G. The φX174 protein J mediates DNA packaging and viral attachment to host cells. *Journal of molecular biology* **337**, 1109–1122 0022–2836 (2004).
16. Dokland, T. *et al.* The role of scaffolding proteins in the assembly of the small, single-stranded DNA virus ϕX174. *Journal of molecular biology* **288**, 595–608 (1999).
17. Dokland, T., McKenna, R., Ilag, L. L. & Bowman, B. R. Structure of a viral procapsid with molecular scaffolding. *Nature* **389**, 308 (1997).
18. McKenna, R. *et al.* Atomic structure of single-stranded DNA bacteriophage ϕX174 and its functional implications. *Nature* **355**, 137 (1992).
19. McKenna, R., Ilag, L. L. & Rossmann, M. G. Analysis of the Single-stranded DNA Bacteriophage ϕX174, Refined at a Resolution of 3·0 Å. *Journal of Molecular Biology* **237**, 517–543 (1994).
20. Ilag, L. L. *et al.* DNA packaging intermediates of bacteriophage ϕX174. *Structure* **3**, 353–363 (1995).
21. Freifelder, D., Kleinschmidt, A. K. & Sinsheimer, R. L. Electron microscopy of single-stranded dna: circularity of dna of bacteriophage phi-x174. *Science* **146**, 254–5 (1964).
22. Fujisawa, H. & Hayashi, M. Assembly of bacteriophage phi X174: identification of a virion capsid precursor and proposal of a model for the functions of bacteriophage gene products during morphogenesis. *J Virol* **24**, 303–13 (1977).
23. Mukai, R., Hamatake, R. K. & Hayashi, M. Isolation and identification of bacteriophage phi X174 prohead. *Proc Natl Acad Sci USA* **76**, 4877–81 (1979).
24. Aoyama, A., Hamatake, R. K. & Hayashi, M. Morphogenesis of phi X174: *in vitro* synthesis of infectious phage from purified viral components. *Proceedings of the National Academy of Sciences* **78**, 7285–7289 (1981).
25. DePonte, D. *et al.* Gas dynamic virtual nozzle for generation of microscopic droplet streams. *Journal of Physics D: Applied Physics* **41**, 195505 (2008).
26. Liang, M. *et al.* The Coherent X-ray Imaging instrument at the Linac Coherent Light Source. *Journal of Synchrotron Radiation* **22**, 514–519 (2015).
27. Hart, P. *et al.* The CSPAD megapixel x-ray camera at LCLS. in *SPIE Optical Engineering + Applications* 85040C-85040C-11 (International Society for Optics and Photonics, 2012).
28. Kern, J. *et al.* Taking snapshots of photosynthetic water oxidation using femtosecond X-ray diffraction and spectroscopy. *Nature communications* **5**, 4371 (2014).
29. Hattne, J. *et al.* Accurate macromolecular structures using minimal measurements from X-ray free-electron lasers. *Nat Meth* **11**, 545–548 (2014).
30. Ginn, H. M., Evans, G., Sauter, N. K. & Stuart, D. I. On the release of cppxfel for processing X-ray free-electron laser images. *Journal of applied crystallography* **49** (2016).
31. Dokland, T. & Murialdo, H. Structural transitions during maturation of bacteriophage lambda capsids. *Journal of molecular biology* **233**, 682–694 0022–2836 (1993).
32. Prasad, B. V. V. *et al.* Three-dimensional transformation of capsids associated with genome packaging in a bacterial virus. *Journal of molecular biology* **231**, 65–74 0022–2836 (1993).

33. Zhou, Z. H. *et al*. Identification of the sites of interaction between the scaffold and outer shell in herpes simplex virus-1 capsids by difference electron imaging. *Proceedings of the National Academy of Sciences* **95**, 2778–2783 0027–8424 (1998).
34. McKenna, R., Bowman, B. R., Ilag, L. L., Rossmann, M. G. & Fane, B. A. Atomic structure of the degraded procapsid particle of the bacteriophage G4: induced structural changes in the presence of calcium ions and functional implications. *Journal of molecular biology* **256**, 736–750 0022–2836 (1996).
35. Roedig, P. *et al*. A micro-patterned silicon chip as sample holder for macromolecular crystallography experiments with minimal background scattering. *Scientific reports* **5** (2015).
36. Sherrell, D. A. *et al*. A modular and compact portable mini-endstation for high-precision, high-speed fixed target serial crystallography at FEL and synchrotron sources. *Journal of synchrotron radiation* **22** (2015).
37. Roessler, C. G. *et al*. Acoustic Injectors for Drop-On-Demand Serial Femtosecond Crystallography. *Structure* **24**, 631–640 (2016).
38. Bamford, D. H. & Mindich, L. Electron microscopy of cells infected with nonsense mutants of bacteriophage φ6. *Virology* **107**, 222–228 (1980).
39. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *Journal of structural biology* **116**, 71–76 (1996).
40. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *Journal of structural biology* **152**, 36–51 (2005).
41. Huiskonen, J. T. *et al*. Averaging of viral envelope glycoprotein spikes from electron cryotomography reconstructions using Jsubtomo. *Journal of visualized experiments: JoVE* (2014).
42. Huiskonen, J. Template Matching, Alignment and Averaging of Sub-Tomograms in Electron Cryo-Tomography Reconstructions using Jsubtomo. *Microscopy and Microanalysis* **20**, 2138–2139 (2014).
43. Margaritondo, G. & Rebernik Ribic, P. A simplified description of X-ray free-electron lasers. *Journal of Synchrotron Radiation* **18**, 101–108 (2011).
44. Ginn, H. M. *et al*. TakeTwo: an indexing algorithm suited to still images with known crystal parameters. *Acta Crystallographica Section D: Structural Biology* **72**, 956–965 (2016).
45. Schrödinger, L. The PyMOL molecular graphics system, version 1.7. 6.6 (2015).

Acknowledgements

PhiX174 and its host were a gift of Dr Ben Fane. We thank Ted Jardetzky for providing access to laboratories and equipment, and Helin Veskiväli for excellent technical assistance. We thank not only all staff at the CXI beamline but also, and especially, members of the Schlichting group for providing and running the GDVN, these include Marc Messerschmidt, Sebastien Boutet, Garth Williams, Dan Deponte, Thomas Barends, Sabine Botha, Bruce Doak, Ilme Schlichting & Robert Shoeman. The authors acknowledge support from FEMS (to M.K.P.), from FEBS (to M.K.P.), from the Viikki Doctoral Programme in Molecular Biosciences (M.K.P.), from the Academy of Finland (130750 and 218080 to J.T.H., 256197 and 256518 to D.H.B.), from the Wellcome DPhil Program in Structural Biology (M.-L.P.) and from the Wellcome Core Award Grant (090532/Z/09/Z). D.I.S. and G.S. were supported by the Medical Research Council, grant MR/N00065X/1 and previously G1000099. H.M.G was supported by Wellcome (studentship 075491/04), H.M.E.D. was supported by Wellcome (studentship ALR00750-B500.1) and J.M.G. was supported by the Medical Research Council (MR/L017709/1). The authors acknowledge the University of Helsinki and Academy of Finland (grant 1306833) for support provided to Centre for Virus and Macromolecular Complex Production (ICVIR), part of Instruct-FI, used in this study, which is an Instruct collaboration (now Instruct-ERIC). N.K.S. acknowledges support from U.S. National Institutes of Health grants GM102520 and GM117126, as well as from Lawrence Berkeley National Laboratory under U.S. Department of Energy contract DE-AC02-05CH11231. Use of the Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515.

Author Contributions

D.I.S. secured funding. H.M.E.D., D.I.S., G.S. wrote the manuscript. H.M.E.D. performed data analysis. H.M.G. wrote software. All authors performed experiments and contributed to the paper. D.M.B. provided reagents, expertise and feedback.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-21693-3.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018

Machining protein microcrystals for structure determination by electron diffraction

Helen M. E. Duyvesteyn[a,b,1], Abhay Kotecha[a,c,1], Helen M. Ginn[a,b], Corey W. Hecksel[b], Emma V. Beale[b], Felix de Haas[c], Gwyndaf Evans[b], Peijun Zhang[a,b], Wah Chiu[d,e,2], and David I. Stuart[a,b,2]

[a]Division of Structural Biology, University of Oxford, Headington, Oxford OX3 7BN, United Kingdom; [b]Diamond Light Source, Harwell Science and Innovation Campus, Didcot OX11 0DE, United Kingdom; [c]Materials and Structural Analysis, Thermo Fisher Scientific, Achtseweg Noord 5, 5651 GG, Eindhoven, The Netherlands; [d]Department of Bioengineering, Stanford University, Stanford, CA 94305; and [e]CryoEM and Bioimaging Division, SSRL SLAC National Accelerator Laboratory, Menlo Park, CA 94025

Contributed by Wah Chiu, August 1, 2018 (sent for review June 11, 2018; reviewed by Mike Marko and Xiaodong Zou)

We demonstrate that ion-beam milling of frozen, hydrated protein crystals to thin lamella preserves the crystal lattice to near-atomic resolution. This provides a vehicle for protein structure determination, bridging the crystal size gap between the nanometer scale of conventional electron diffraction and micron scale of synchrotron microfocus beamlines. The demonstration that atomic information can be retained suggests that milling could provide such detail on sections cut from vitrified cells.

electron crystallography | FIB | protein crystal | diffraction | structure

Despite the explosive growth in cryoelectron microscopy (cryo-EM) following the "resolution revolution" (1), X-ray methods were responsible for nearly 20-fold as many macromolecular structure depositions as single-particle cryo-EM in 2017, and diffraction remains the favored method for genuine atomic resolution analysis. Crystal growth is often a bottleneck for such analyses, and it can sometimes be difficult to grow crystals of sufficient size for analysis at conventional synchrotron macromolecular crystallography (sMX) beamlines. This has led to the development of microfocus beamlines that can routinely analyze crystals down to ∼5 μm (2). However, it is unknown how often crystallographic analyses fail due to the failure to recognize smaller crystals. The next generation of synchrotron beamlines may reduce the size requirement to around 1 μm, and the potential for structure determination from submicron crystals has been demonstrated by application of X-ray free electron laser (XFEL) sources, which achieve a brightness ∼10^8-fold greater than a typical synchrotron beamline (3). XFEL sources also sometimes mitigate, by achieving "diffraction before destruction," the radiation damage problem that limits the collection of synchrotron X-ray data (4, 5).

Electron diffraction (ED) can provide high-resolution structure determination from 3D crystals (6–12). Electrons have potential advantages over X-rays for the analysis of very small crystals: Their interaction with matter is sufficiently strong for a measurable signal to be obtained from a typical laboratory instrument; the experiment is performed in a vacuum, thereby reducing background scatter; and the useful data obtained before radiation damage occurs are much greater (6, 13, 14). However, the strong interaction with matter is also a limitation for ED, since the beam is absorbed very strongly by a few hundred nanometers of crystal. Furthermore, it has often been held that kinematic theories of diffraction, upon which X-ray analysis rests, are useless for crystals of such thickness (15, 16). For protein crystals, due to the large size of the unit cell, many diffracted beams are stimulated simultaneously and, through conservation of energy, the amplitude of each is greatly reduced, mitigating the effects of multiple scattering. Although dynamical effects have been observed with thicker protein crystals (6), the utility of the kinematic theory for thin crystals is shown in our study below, where systematic absences, which would be polluted by multiple scattering, are maintained. Indeed, in recent years, the power of ED to analyze nanometer-sized crystals has been amply demonstrated (6, 12, 17); however, there remains a significant range of size, from a few hundred nanometers to a few microns, that is accessible only using XFEL radiation, which is both hard to access and extremely expensive.

Micromachining using a focused ion beam (FIB milling) has been established in physical science applications for many years (18), and, although far from routine, cryo-FIB milling is now increasingly applied in the life sciences, for instance, to create thin lamella milled from frozen, hydrated cells to reveal cellular architecture (19). Cryo-FIB milling has been shown to preserve the structure of frozen, hydrated biological samples better than mechanical cryosectioning, although how far the preservation of order extends has not been established (19).

Results

We reasoned that if cryo-FIB milling does not distort the internal structure of cut lamellae, then milled crystals would maintain their diffraction capacity and milling would provide a general vehicle to bridge the crystal size gap between sMX and ED. As a proof-of-principle experiment, we grew tetragonal microcrystals (*ca.* $10 \times 10 \times 5$ μm^3) of hen egg-white lysozyme. A

Significance

Electron cryomicroscopy and diffraction are increasingly powerful methods for the analysis of biological structures at near-atomic resolution. However, the very strong interaction of electrons with matter limits the thickness of specimens that can be analyzed at high resolution to ∼100 nm. We show here that a frozen, hydrated lysozyme crystal milled with a focused ion beam to a suitable thickness produces useful electron diffraction to ∼2-Å resolution. This approach provides a means to determine structures of protein crystals that would otherwise be too thick for electron crystallographic analysis. This demonstration also suggests that this technique will allow lamella of cryopreserved biological material prepared by milling to preserve the structural details, approaching atomic detail.

Author contributions: A.K., W.C., and D.I.S. designed research; H.M.E.D., A.K., C.W.H., E.V.B., F.d.H., and G.E. performed research; H.M.E.D., H.M.G., W.C., and D.I.S. analyzed data; and H.M.E.D., A.K., C.W.H., P.Z., W.C., and D.I.S. wrote the paper.

Reviewers: M.M., New York State Department of Health; and X.Z., Stockholm University.

The authors declare no conflict of interest.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.wwpdb.org (PDB ID code 6H3B). Mandexing was created for this article, licensed under GPLv3, and deposited on GitHub (available at https://www.github.com/helenginn/mandexing).

[1]H.M.E.D. and A.K. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: wahc@stanford.edu or david.stuart@strubi.ox.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1809978115/-/DCSupplemental.

Published online August 31, 2018.

Downloaded by guest on March 12, 2020

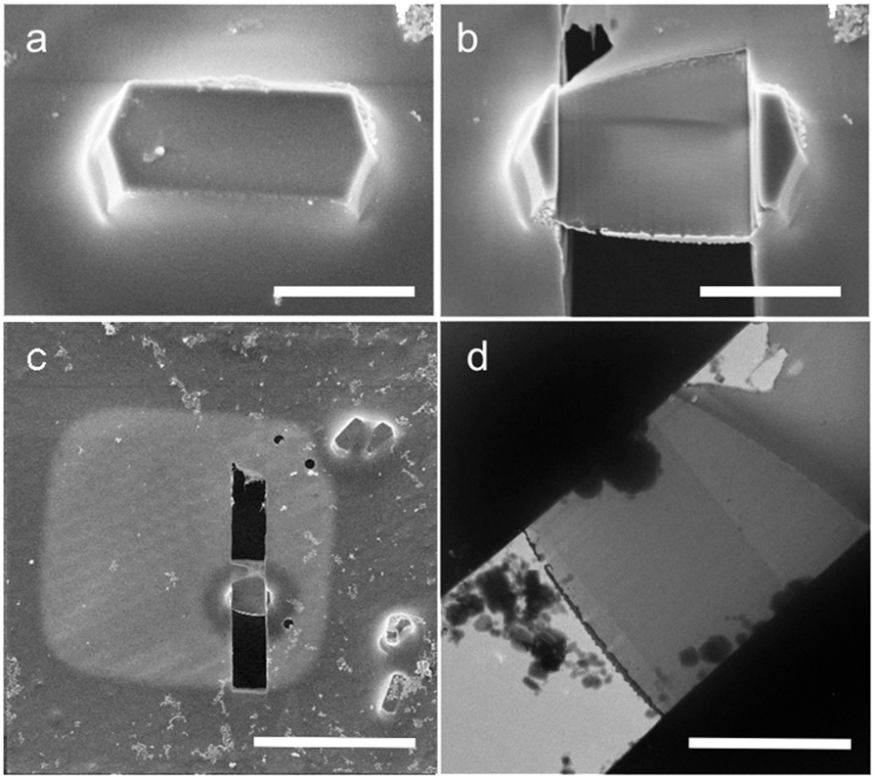


Fig. 1. Images of the frozen, hydrated lysozyme crystal. Representative SEM images are shown before (*A*) and after (*B* and *C*) milling. (*D*) Transmission EM image of the milled lamella. (Scale bars: *A* and *B*, 4 μm; *C*, 30 μm; *D*, 3 μm.)

limitation of lysozyme as a test system is that it shows significant crystal-to-crystal variation in unit cell parameters (*SI Appendix*, Fig. S1); however, it generally diffracts to a resolution of 2 Å or better. Crystals were grown using established methods and deposited on standard EM grids, which were then blotted and vitrified by plunging into liquid ethane (*Materials and Methods*). The crystals, maintained at 100 K, were coated with platinum before milling in a dual-beam instrument using a beam of gallium ions, to thicknesses of *ca.* 200 nm (*Materials and Methods* and Fig. 1*A*). The final stages of milling used a low-voltage ion beam, a strategy shown to minimize surface damage and strain in metallic crystals (20). Unlike some current protocols, our procedure did not involve subsequent sputtering with platinum (which is thought to allow improved dissipation of charge). A representative example of lamella used for data collection is shown in Fig. 1. Diffraction data were collected on a Thermo Fisher Scientific Talos Arctica microscope operating at 200 kV, with a beam size ∼2 μm in diameter as still snapshots with the stage tilted from −60 to +10° in steps of 0.5° between images, resulting in 134 individual diffraction patterns, at a total electron dose of ∼25 $e^{-}/Å^2$ (*Materials and Methods*). Data were recorded on a Ceta 16M (Thermo Fisher Scientific) detector (https://www.fei.com/accessories/ceta-16m/), with measured detective quantum efficiency at 1/2 Nyquist of 0.12. No correction was made for possible saturation of the detector leading to nonlinearity of response because the distribution of pixel values shows no evidence of saturation. Typical patterns are shown in Fig. 2, and the full set is shown in Movie S1. In the early patterns, visible diffraction spots extended to a Bragg spacing of 1.9 Å, while by pattern 110, the effective resolution had declined to ∼4 Å due to severe radiation damage, as expected since we markedly exceeded the advised maximum electron dose (7, 14, 21, 22) (Movie S1). A principal axis may be seen in pattern 118, confirming that the systematic absences are well obeyed (Fig. 2*B*). Autoindexing procedures developed for sMX and XFEL analysis were not able to determine crystal orientation. However, using the principal axis as an anchor, the orientation matrix was fully defined using a purpose-written, graphical user interface-driven program (*Materials and Methods*). Since the data were monochromatic still images, and the mosaic spread of the crystal was remarkably small (∼0.01°), the data were treated as independent diffraction patterns. Further analysis of the indexed data was therefore achieved using software originally developed for XFEL images (*Materials and Methods*). In this initial experiment, fully automated analysis was hindered by an apparent detector tilt and by drift in the direct beam position on the detector. Details of the analysis are provided in *Materials and Methods*. The unit cell dimensions of the crystal were determined as a = b = 79.9, c = 38.7 Å, which is within the range expected for lysozyme crystals under cryoconditions. Clearly, the data

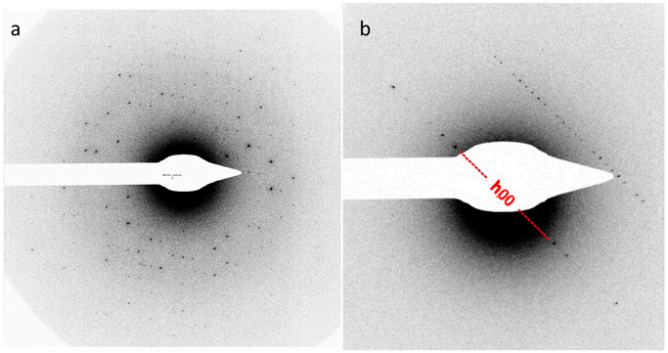


Fig. 2. ED patterns of the frozen, hydrated lysozyme crystal. The edges of the image correspond to Bragg spacings of ∼1.8 Å. The first image collected (*A*) and image 118 (*B*), with the *h,0,0* row of reflections indicated, are shown. Comparison of these reflections with those from a general *h* row (clearly seen running parallel to the *h,0,0* row) reveals the systematic absences.

Downloaded by guest on March 12, 2020

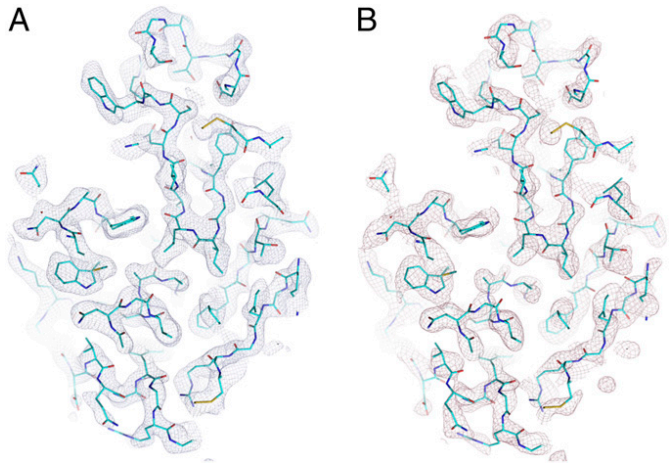


Fig. 3. Section of the lysosome crystal density map for a traditional 2Fo-Fc synthesis (*A*) and a composite omit map from Phenix (*B*). Contours are drawn at a level of 1 σ. Note the composite omit map eliminates bias from the phases (derived from a rigid body-refined molecular replacement model), and thus represents the useful information in the measured amplitudes.

collection method was suboptimal (essentially all reflections were partially recorded), and the detector was almost an order of magnitude less efficient than an ideal detector. Nevertheless, an RSplit of 28 (25)% and correlation coefficient ($CC_{1/2}$) of 85 (85)% between two random half datasets on all data to 2.0 Å (values to 4.5 Å are shown in parentheses) for a dataset of over 80% completeness at a resolution of 4.5 Å (falling to <30% at a resolution of 3 Å; *SI Appendix*, Figs. S2 and S3) was obtained from a single lamella. B-factor increases (estimated from the fall-off of the high-resolution data on successive diffraction patterns) were consistent with significant radiation damage (*SI Appendix*, Fig. S4). This dataset is isomorphous with previous data collected from lysozyme (CC = 74%), and rigid body refinement with no manual intervention, to avoid overfitting, against a typical lysozyme model [Protein Data Bank (PDB) ID code 5wr9] yielded Rwork = 29.1% and Rfree = 28.2%. The moderate quality of the data may be judged from the potential function maps shown in Fig. 3. Therefore, although this experiment was far from ideal (for instance, a complete dataset with little radiation damage could be collected if a more sensitive detector was used, and properly integrated measurements could be obtained from continuous recording during crystal rotation), the resulting data are not drastically worse in quality than data reported from unmilled lysozyme crystals (25.5%/27.8% Rwork/Rfree after full refinement) (6).

Discussion

This work is a proof of principle; better detectors, better experimental procedures, continuous rotation of the crystal, and improved data analysis methods (including consideration of dynamical scattering effects) are required to establish the full potential for the collection of diffraction data from lamellae. In addition, we have not determined the depth of any surface damage to the lamella, or the effect of beam-induced charging. We expect that, with optimization of the method, the use of multiple lamellae will enable the collection of complete, accurate, high-resolution sets of diffraction intensities, as is now routine for crystals of a few microns with microfocus sMX. Thus, the use of milled crystals will remove the size restriction on crystals for ED, bridging the size gap to synchrotron X-ray diffraction, which currently struggles with crystals less than ∼5 μm across, and potentially releasing the limited availability on XFEL beamlines for some of the structure questions. Furthermore, and perhaps more importantly, the retention of high-resolution diffraction demonstrates that cryo-FIB milling preserves atomic-level structure within the lamellae, suggesting that this method might provide a vehicle for the determination of molecular and atomic structures via electron imaging and diffraction from not only protein crystals but also cellular structures, where precision milling can isolate the region of interest in a slice thin enough to provide good signal to noise. Analysis of such lamellae, most likely via electron tomography, would provide a link from high-resolution structural biology to cell biology. Furthermore, subtomogram averaging or even in situ single-particle analysis of repeated structures might then allow the visualization at near-atomic detail within the cell.

Table 1. Data collection and analysis

Data	Instrument/parameters	Specification/measurements
Data collection	Accelerating voltage, kV	200
	Electron source	Field emission gun
	Wavelength, Å	0.0251
	Total electron dose per crystal, $e^{-}/Å^2$	25
	No. of diffraction patterns per crystal	134
	No. of crystals	1
	Total reflections to 1.9 Å	8,463
Data analysis	Space group	$P4_32_12$
	Unit cell dimensions, Å	
	a = b	79.90
	c	38.67
	α = β = γ	90°
	Resolution, Å (highest resolution bin)	10–1.9 (2.14–1.90)
	Total unique reflections	3,510
	Reflections in working set	3,337 (307)
	Reflections in free set	173 (10)
	Multiplicity	0.82
	Completeness	38.7 (10.5)
	RSplit, %	27.9 (93.2)
	Rwork/Rfree, %	29.1/28.3 (53.2/72.9)
	Overall correlation coefficient, %	81.3 (29.2)
	Ramachandran % for outliers, favored	0, 98.4

The highest resolution shell statistics are shown in parentheses.

Downloaded by guest on March 12, 2020

Materials and Methods

Crystal and Grid Preparation. A solution of 50 mg/mL lysozyme (Sigma–Aldrich) in 0.02 M sodium acetate (pH 4.6) was batch-crystallized by addition of 1 M sodium acetate (pH 3.0), 6% (wt/vol) PEG 6000, and 10% (wt/vol) NaCl in a volumetric ratio of 1:4 at 25 °C. A 4-μL aliquot of a 1:10 dilution of the batch crystallization was applied onto a glow-discharged, holey, carbon-coated copper grid (C-flat, CF-2/1-2C; Protochips) at room temperature. Grids were blotted for 3–4 s from the reverse side and plunge-frozen in liquid ethane using a manual plunger. Crystals were typically ∼10 μm in length. Frozen grids were transferred into liquid nitrogen for storage until needed.

Lamella Preparation. Plunge-frozen grids were clipped into autogrid (Thermo Fisher Scientific) rims and loaded into the cryo-FIB (Thermo Fisher Scientific) using a Quorum cryosystem (PP3010T) with a custom-designed cryostage and shuttle (Max Planck Institute of Biochemistry and Thermo Fisher Scientific) (23, 24). The cryostage and anticontaminator were cooled for a minimum of 90 min before loading samples, using a continuous nitrogen gas flow rate of 3.6 L·min^{-1} and 4 L·min^{-1} at −165 °C and −175 °C, respectively. Milling was performed as previously described (23, 25). Before milling, samples were precoated with a platinum layer ∼2 μm thick using the gas injection system (GIS). This was accomplished by tilting the stage to 7°, rotating by 180° (such that the surface of the grid faced the ion column), and lowering the sample by 5 mm (12 mm as measured from the SEM column). The GIS needle was inserted and heated to 27 °C, and samples were coated for 5 s per grid. SEM imaging was performed at 5 kV and 25 pA beam current. The lamella was cut at an angle of 10° with respect to the grid, with a final thickness of ∼0.2 μm. Milling was performed in a stepwise fashion, using a 30-kV beam and beam currents of 300 pA, 100 pA, 49 pA, and 30 pA to generate lamella with a progressive thickness of ∼5 μm, ∼2 μm, ∼1 μm, and ∼0.2 μm, respectively, measured using the ion beam image. During the final milling step, lamella thickness was monitored using the SEM with a 3-kV beam at 25 pA as described previously (25). All images were recorded using an Everhart–Thornley detector to record secondary electrons or FIB-induced secondary electrons.

Data Collection. ED data were collected from a single crystal lamella, ∼200 nm thick, over an angular range from −60 to +10° rotation around a tilt axis on the grid plane normal to the electron beam. Diffraction data were collected on an FEI Talos Arctica transmission microscope equipped with a field emission electron gun operating at an accelerating voltage of 200 kV, corresponding to a wavelength of 0.0251 Å. Images were recorded in diffraction mode as stills, separated by 0.5° tilt of the grid, on a Ceta camera with a sensor size of 4,096 × 4,096 pixels and at a physical camera length of 1.75 m, corresponding to an apparent detector distance of 2.44 m with lenses. The resolution at the edge of the detector was 1.8 Å. Individual diffraction frames were taken with exposure times of 2.5 s per diffraction pattern, with an electron dose rate of ∼0.042 $e^{-}\cdot Å^{2}\cdot s^{-1}$ (measured using a fluorescent screen) and with a 2-s wait between patterns to allow the stage to settle. The electron beam size on the lamella was ∼1.9 × 1.9 μm.

Crystal Indexing. Software providing a graphical user interface was written to allow manual indexing of the diffraction image (Mandexing). This is a lightweight piece of software, dependent on the Qt5 framework, which implements a simple detector geometry (flat detector perpendicular to the beam center at a given distance). Portable Network Graphics files can be loaded to provide a reference diffraction image. The beam center, detector distance, and wavelength parameters can be varied. The unit cell can be set, and two chosen Miller index vectors can be brought parallel to the screen. Rotation around the beam center is controlled by the mouse, and incremental rotations can be accessed using the "WASD" keys. These keys operate on the vertical and horizontal directions, unless a particular axis perpendicular to the beam direction has been fixed. In this case, the keys operate around that fixed axis and the axis perpendicular to both that axis and the beam axes.

The parameters of the lysozyme micro-ED experiment are shown in Table 1. The sample is rotated about an axis approximately vertical on the displayed image. The 118th pattern clearly displays the axial systematic absences resulting from a 2_1 screw axis, and this was therefore defined as the a* axis (Fig. 2*B*). Mandexing was used to fix this axis and manually rotate the diffraction pattern around it to find an orientation that matched the remaining nonaxial spots elsewhere on the image. There were a small number of solutions that, visually, fitted equally well. From these, the correct solution was selected as that which properly predicted spots on preceding and subsequent images. Lysozyme typically crystallizes in $P4_32_12$, but the unit cell parameters can vary significantly (*SI Appendix*, Fig. S1). The unit cell dimension for this analysis, a = b = 79.9 Å, was chosen based on the fit to the h00 axis for the assumed wavelength and detector distance. The c-axis dimension was determined as 38.7 Å, which corresponds to an a-to-c ratio within the range observed for the majority of lysozyme datasets. Strong spots matching predicted reflections were picked out manually for each image and refined using a reciprocal lattice point size of 0.003 $Å^{-1}$ to minimize the distance in reciprocal space to the Ewald sphere. These orientations were used for integration.

Data Processing. The derived orientation matrices were used to integrate and post-refine the raw intensities using cppxfel (26). Patterns were processed using individual direct beam positions determined using custom scripts (the beam position drifted on the detector by 30 pixels perpendicular and 91 pixels parallel to the axis of rotation across the entire data collection). A residual discrepancy between predicted and observed spot positions was successfully modeled by an 8° tilt in the detector (this could be an artifact of the electron optics), after which the average deviation between observed and predicted spot positions was less than one-half of a typical spot diameter. Integration was performed as described previously (26), with a foreground integration window size of 10 × 10 pixels. The surrounding four-pixel moat was ignored, and the outer window extended to 25 × 25 pixels for background estimation. Merging and postrefinement with partiality correction were performed as described previously (26, 27) with the following beam parameters: wavelength = 0.0251 Å, mosaic spread = 0.01°, fractional energy spread of the beam = 0.00075 (modeled as a super-Gaussian), and crystal size = 0.3 μm. Individual reflections with a partiality of less than 0.2 were discarded; otherwise, their intensity was divided by the partiality to approximately recover the true intensity (Table 1).

Structure Determination. Since the crystals were approximately isomorphous with those used for previously reported structures, the structure was determined simply by rigid body refinement of an extant structure (PDB ID code 5wr9) (Table 1) using phenix.refine with electron scattering factors and with manually added waters.

Composite omit maps were generated using Phenix with sequential 5% fractions omitted and Fc used where Fo was not recorded.

Code Availability. Mandexing (https://www.github.com/helenginn/mandexing) and cppxfel (https://github.com/helenginn/cppxfel) are publicly available on GitHub.

Data Availability. The partially refined model of a milled lysosome crystal has been deposited in the PDB (ID code 6H3B).

ACKNOWLEDGMENTS. We thank Robert Imlau for advice. H.M.E.D. (Studentship ALR00750-B500.1), H.M.G. (Studentship ALR00040), and A.K. (Grant 101122/Z/13/Z) were supported by Wellcome. D.I.S. was supported by the UK Medical Research Council (MRC; Grant MR/N00065X/1). P.Z. was supported by Wellcome (Grant 206422/Z/17/Z). W.C. was supported by the NIH (Grant P41GM103832). The Wellcome Centre for Human Genetics is supported by Wellcome (Grant 090532/Z/09/Z). This is an Oxford Centre Instruct-European Research Infrastructure Consortium contribution. Electron Bio-Imaging Centre (eBIC) is supported by Wellcome, the MRC, and the Biotechnology and Biological Sciences Research Council.

1. Kühlbrandt W (2014) Biochemistry. The resolution revolution. *Science* 343:1443–1444.
2. Grimes JM, et al. (2018) Where is crystallography going? *Acta Crystallogr D Struct Biol* 74:152–166.
3. Hajdu J (2002) The challenge offered by X-ray lasers. *Nature* 417:15.
4. Emma P, et al. (2010) First lasing and operation of an ångstrom-wavelength free-electron laser. *Nat Photonics* 4:641–647.
5. Neutze R, Wouts R, van der Spoel D, Weckert E, Hajdu J (2000) Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* 406:752–757.
6. Shi D, Nannenga BL, Iadanza MG, Gonen T (2013) Three-dimensional electron crystallography of protein microcrystals. *eLife* 2:e01345.
7. Nannenga BL, Shi D, Leslie AGW, Gonen T (2014) High-resolution structure determination by continuous-rotation data collection in MicroED. *Nat Methods* 11:927–930.
8. Yonekura K, Kato K, Ogasawara M, Tomita M, Toyoshima C (2015) Electron crystallography of ultrathin 3D protein crystals: Atomic model with charges. *Proc Natl Acad Sci USA* 112:3368–3373.
9. Clabbers MTB, et al. (2017) Protein structure determination by electron diffraction using a single three-dimensional nanocrystal. *Acta Crystallogr D Struct Biol* 73:738–748.
10. Xu H, et al. (2018) A rare lysozyme crystal form solved using highly redundant multiple electron diffraction datasets from micron-sized crystals. *Structure* 26:667–675.e3.
11. Rodriguez JA, et al. (2015) Structure of the toxic core of α-synuclein from invisible crystals. *Nature* 525:486–490.
12. de la Cruz MJ, et al. (2017) Atomic-resolution structures from fragmented protein crystals with the cryoEM method MicroED. *Nat Methods* 14:399–402.

Downloaded by guest on March 12, 2020

13. Henderson R (1995) The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q Rev Biophys* 28: 171–193.
14. Jeng T-W, Chiu W (1984) Quantitative assessment of radiation damage in a thin protein crystal. *J Microsc* 136:35–44.
15. Cowley JM (1968) Crystal structure determination by electron diffraction. *Prog Mater Sci* 13:267–321.
16. Glaeser RM, Downing K, DeRosier D, Chiu W, Frank J (2007) *Electron Crystallography of Biological Macromolecules* (Oxford Univ Press, New York).
17. Martynowycz MW, Gonen T (2018) From electron crystallography of 2D crystals to MicroED of 3D crystals. *Curr Opin Colloid Interface Sci* 34:9–16.
18. Young RJ (1993) Micro-machining using a focused ion beam. *Vacuum* 44:353–356.
19. Rigort A, et al. (2012) Focused ion beam micromachining of eukaryotic cells for cryoelectron tomography. *Proc Natl Acad Sci USA* 109:4449–4454.
20. Hofmann F, et al. (2018) Incidence focussed ion beam milling: A coherent X-ray diffraction study of 3D nano-scale lattice strains and crystal defects. *Acta Mater* 154: 113–123.
21. Chen D-H, Jakana J, Liu X, Schmid MF, Chiu W (2008) Achievable resolution from images of biological specimens acquired from a 4k x 4k CCD camera in a 300-kV electron cryomicroscope. *J Struct Biol* 163:45–52.
22. Bammes BE, Jakana J, Schmid MF, Chiu W (2010) Radiation damage effects at four specimen temperatures from 4 to 100 K. *J Struct Biol* 169:331–341.
23. Schaffer M, et al. (2015) Cryo-focused ion beam sample preparation for imaging vitreous cells by cryo-electron tomography. *Bio Protoc* 5:e1575.
24. Engel BD, et al. (2015) Native architecture of the Chlamydomonas chloroplast revealed by in situ cryo-electron tomography. *eLife* 4:e05889, and erratum (2015) 4: e11383.
25. Schaffer M, et al. (2017) Optimized cryo-focused ion beam sample preparation aimed at in situ structural studies of membrane proteins. *J Struct Biol* 197:73–82.
26. Ginn HM, Evans G, Sauter NK, Stuart DI (2016) On the release of cppxfel for processing X-ray free-electron laser images. *J Appl Crystallogr* 49:1065–1072.
27. Ginn HM, et al. (2015) A revised partiality model and post-refinement algorithm for X-ray free-electron laser data. *Acta Crystallogr D Biol Crystallogr* 71:1400–1410.

Downloaded by guest on March 12, 2020

COMMUNICATIONS BIOLOGY

ARTICLE

https://doi.org/10.1038/s42003-019-0722-x OPEN

Glutathione facilitates enterovirus assembly by binding at a druggable pocket

Helen M.E. Duyvesteyn [1,2], Jingshan Ren [1], Thomas S. Walter [1], Elizabeth E. Fry [1] & David I. Stuart [1,2*]

Enteroviruses cause a range of human and animal diseases, some life-threatening, but there remain no licenced anti-enterovirus drugs. However, a benzene-sulfonamide derivative and related compounds have been shown recently to block infection of a range of enteroviruses by binding the capsid at a positively-charged surface depression conserved across many enteroviruses. It has also been established that glutathione is essential for the assembly of many enteroviruses, interacting with the capsid proteins to facilitate the formation of the pentameric assembly intermediate, although the mechanism is unknown. Here we show, by high resolution structure analyses of enterovirus F3, that reduced glutathione binds to the same interprotomer pocket as the benzene-sulfonamide derivative. Bound glutathione makes strong interactions with adjacent protomers, thereby explaining the underlying biological role of this druggable binding pocket and delineating the pharmacophore for potential antivirals.

[1] Division of Structural Biology, University of Oxford, The Henry Wellcome Building for Genomic Medicine, Headington, Oxford OX3 7BN, UK. [2] Diamond Light Source, Harwell Science and Innovation Campus, Didcot OX11 0DE, UK. *email: dave@strubi.ox.ac.uk

The genus *Enterovirus* (EV) is the most populous genus of the *Picornaviridae* family of small, single-stranded RNA viruses. Enteroviruses are responsible for a range of animal and human diseases. These non-enveloped viruses are constructed of an icosahedral protein capsid composed of 60 copies each of viral proteins VP1–4. VP1–3 are similar to each other (comprising a β-barrel with extended surface loops and termini) and form the capsid surface. Whilst a handful of enterovirus vaccines are available, including for poliovirus (PV) and for EV-A71 (frequently responsible for serious outcomes of hand, foot and mouth disease)[1–4], there are as yet no licenced anti-enterovirus drugs, despite evidence for efficacy in clinical trials and reports of extremely high potency for compounds targeting a pocket internal to capsid protein VP1 (termed here the *P* site)[5,6]. However, a cavity on the capsid surface at an inter-protomer interface, where we previously reported extra electron density[7], has recently been identified as potentially druggable, with micromolar binding for a benzene-sulfonamide derivative[8]. This compound, and a number of related compounds, are inhibitors of Coxsackie B viruses and a range of other enteroviruses including rhinoviruses, and have been shown to act synergistically with *P* site binders[8].

Cells contain glutathione in both reduced (GSH) and oxidised (GSSG) forms, whose balance is crucial for maintaining cellular redox potential[9]. Disruption of the GSH/GSSG balance is associated with several enterovirus infections, including EV-A71[10], CV-A16 and CV-B3[11–13]. Furthermore, depletion of GSH, using L-butathione sulfoximate (BSO)[14,15] and TP219[16], blocks the assembly of protomeric units into pentamers[14,16,17]. GSH-independence can be conferred by mutations at the protomer interface, consistent with GSH binding to protomers driving the formation of pentameric units critical to assembly[16,17]. However, there are no published structural data for GSH binding to an enterovirus capsid.

Here we study EV-F3, a relatively benign bovine enterovirus which is our model system of choice[7,18–22]. We firstly show that EV-F3 is dependent on GSH and that GSH stabilises the capsid. We then determine high-resolution structures of complexes of GSH and its initial breakdown product Cys-Gly (CG) bound to EV-F3 and by a competition binding study demonstrate that GSH binds more strongly. We find that the binding site of both compounds is identical to that for a similar molecule that remains naturally attached to the virus[7], and to the binding site observed for the benzene-sulfonamide derivative in complex with Coxsackievirus B3[8], identifying the biological role for this binding site.

Results

EV-F3 is dependent on and stabilised by GSH. Infection of cells both with and without BSO treatment shows that inhibition of glutathione synthesis reduces EV-F3 growth by 3.5 log, demonstrating glutathione dependency (Supplementary Table 1). Since GSH binding has been shown to stabilise PV[17,23], we next used the PaSTRy assay[24] to demonstrate that GSH also shows a modest dose-dependent stabilisation of EV-F3 (Methods, Supplementary Fig. 1).

GSH and CG bind at an inter-protomer surface pocket on EV-F3. To determine if GSH binds to the enterovirus capsid, data to 1.8 Å resolution were collected from crystals of EV-F3 soaked with GSH by automated X-ray cryo-crystallography (Methods and Table 1), yielding high-quality electron density and a reliable model for surface-bound GSH (Fig. 1, Supplementary Fig. 2 and Table 1). GSH binds in a cavity between protomers within the pentameric building block (Fig. 1a–d), consistent with GSH's role in protomer assembly[14,16]. We term this the interface (*I*) site. There is also an additional weak binding site nestled between the VP1 and VP3 C-termini at a non-conserved site, but this is probably not biologically significant (Fig. 1a).

Table 1 Data collection and refinement statistics.

	Experiment 1: EV-F3 GSH	Experiment 2: EV-F3 CG	Experiment 3: EV-F3 GSH/CG
Data collection			
Number of crystals	2	1	2
Space group	I222	I222	I222
Cell dimensions			
a, *b*, *c* (Å)	342.8, 348.2, 351.4	342.7, 348.3, 351.6	344.0, 349.4, 352.7
α, *β*, *γ* (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	20.0–1.80 (1.83–1.80)	20.0–1.67 (1.70–1.67)	110.0–2.17 (2.20–2.17)
R_{merge}	0.394 (—)	0.216 (—)	0.374 (0.798)
R_{pim}	0.111 (—)	0.089 (—)	0.187 (0.787)
$I/\sigma I$	6.2 (0.5)	7.2 (0.7)	5.7(3.6)
Completeness (%)	99.9 (99.9)	99.7 (95.4)	82.6 (76.9)
Redundancy	13.4 (7.8)	6.9 (5.2)	3.7(1.4)
Refinement			
Resolution (Å)	20.0–1.80	20.0–1.67	20.0–2.17
No. reflections	179,939/ 95,446	2,243,186/ 117,927	862,708/ 45,615
R_{work}/R_{free}	0.205, 0.208	0.190/0.191	0.200/0.203
No. atoms			
Protein	6291	6291	6288
Ligand/ion	86	79	94
Water	717	742	612
B-factors			
Protein	22.2	22.4	18.5
Ligand/ion	44.4	37.6	50.9
Water	31.1	34.9	33.8
R.m.s. deviations			
Bond lengths (Å)	0.005	0.005	0.005
Bond angles (°)	1.4	1.4	1.4

In un-soaked EV-F3 virions, the *I* site is partially occupied by a ligand similar to GSH (Fig. 2a) and to clarify the nature of the binding we performed two further structure determinations, firstly with CG (the GSH breakdown product) and secondly a competition soak with equal concentrations of both GSH and CG (Methods and Fig. 2). In summary, these experiments demonstrated that (i) GSH can replace the natural bound factor, (ii) that Cys-Gly (CG), the initial breakdown product of GSH[9], also binds at the *I* site and (iii) that GSH binds more strongly than CG (Fig. 2, Methods and Table 1).

The *I* site comprises six segments of polypeptide chain contributed by two adjacent protomers, A and B. Three of these are from VP1 (protomer A), residues 75–79, 95–98 and 227–231 respectively, a fourth is from the C-terminus of VP3 (protomer A), and the fifth and sixth are from residues 150–155 and 216 of VP1 of protomer B (Fig. 1d). Numerous hydrophobic interactions, hydrogen bonds and salt bridges stabilise GSH binding (Fig. 3) and several interacting residues are conserved across enteroviruses (Fig. 4a). Notable are the salt bridges formed between the carboxy terminus of GSH and the side chains of VP1 R216 (protomer B) and VP1 R229 (protomer A), which form the core of a strong positive patch within the *I* site (Figs. 1d and 3). The CG portion of GSH anchors the two adjacent protomers via not only these salt bridges, but also by additional H-bonding interactions with Y231 of protomer A and the carbonyl oxygen of W154 of protomer B via a bridging water molecule. The main

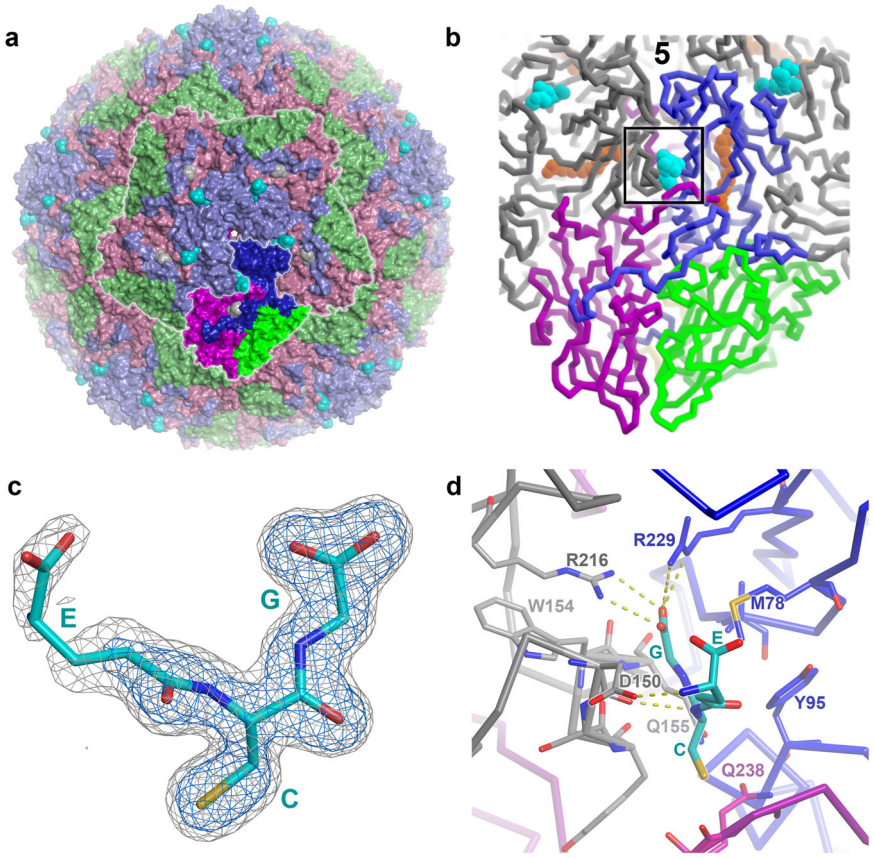


Fig. 1 Virus–glutathione interactions. a EV-F3 capsid surface coloured by protein: VP1 blue, VP2 green, VP3 magenta. GSH at Site *I* is in cyan, and a weaker binding site is shown in white. A pentameric unit and a protomeric unit are both outlined in white and the protomer is coloured bold. **b** Cartoon of protomeric unit (with adjacent protomers shown in grey) with druggable sites: *P* orange, *I* cyan. Protein chains in protomeric unit A coloured as in (**a**), with VP4 in yellow (barely visible on the inside of the capsid). The position of the fivefold axis is marked "5". **c** 15-fold NCS-averaged electron density for GSH, contoured at two levels, the blue level being 2.5 times higher. **d** GSH at site *I*. Protomer A is coloured as in (**b**), protomer B is in grey. The region shown in panel (**d**) is boxed on panel (**b**).

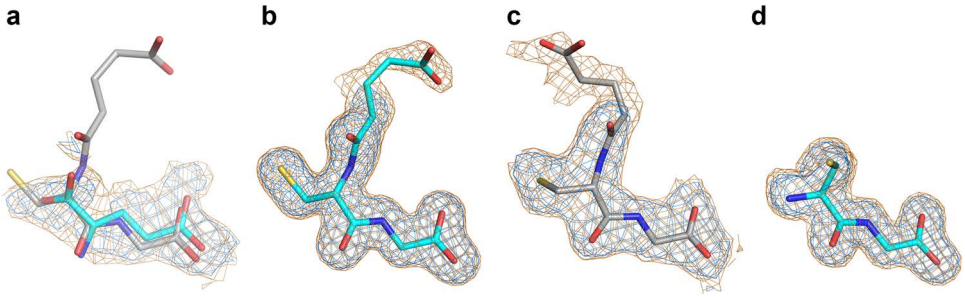


Fig. 2 Electron density for GSH environments at the *I* site. a XFEL EV-F3 (PDB ID, 5OSN). **b** GSH soak (experiment 1). **c** competition soak (experiment 3). **d** CG soak (experiment 2). All maps are NCS-averaged and the view is the same for all panels. Two contour levels are shown, with blue 2.5 times higher than orange.

chain amide and carbonyl groups of the C residue of GSH form additional H-bonds with the side chain of the carbonyl oxygen of D150, and the side chain of Q155, respectively (both protomer B). Although the carboxylate group of the E residue of GSH makes no direct contacts with the *I* site, other portions of the residue contact D150 (protomer B) and, indirectly, R227 (protomer A) at the inter-protomer interface (Figs. 1d and 3). Given that for GSH the strongest interactions are formed by the C-terminus of the ECG tripeptide, the binding of CG is not unexpected, the preference for GSH, however, confirms that additional binding energy is contributed by residue E, although its side chain can assume multiple conformations (Fig. 2b, c).

The GSH binding site is conserved and is occupied in several virus structures. The patch of positive charge within the *I* site is conserved across a panel of enteroviruses (note strict conservation in VP1 of R216 and R229 and the presence of either K or R at 233 in Fig. 4a) and although the width of the pocket varies, it remains sufficient to accommodate glutathione (if the flexible Glu side chain is allowed to rearrange, Fig. 4b). In several previously determined enterovirus structures, *I* site density has been observed and variously interpreted as water (EV-A71, 3VBH)[25], chlorine (CV-A16, 5C4W)[26], glycerol (HRV14, 4PDW)[27] and glutamic acid (EV-F3, 5OSN)[7]. These previous interpretations sometimes fail to fully explain the observed density (Supplementary Fig. 3),

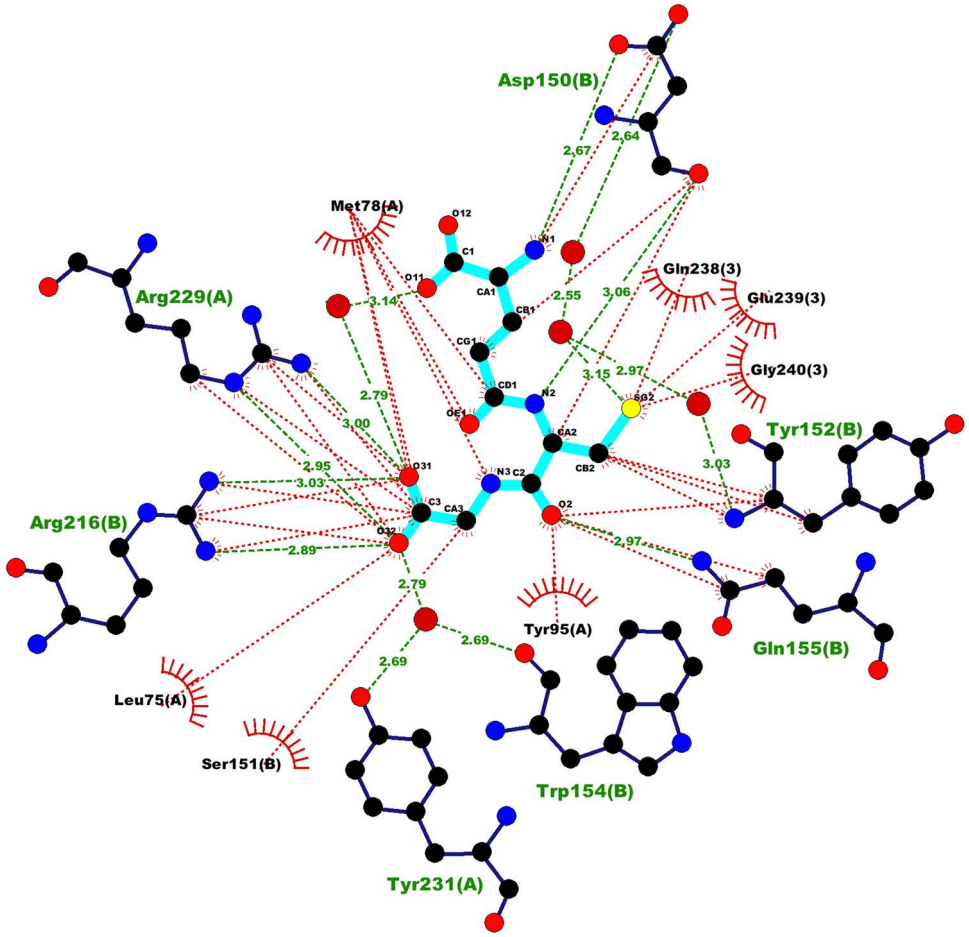


Fig. 3 Details of interactions of GSH at the *I* site. GSH is shown in cyan. Red arcs denote hydrophobic interactions. The residue numbers for the interacting protein residues are shown with the chain label in parentheses (**A** is VP1 of protomer A, **B** is VP1 of protomer B and **3** is VP3 of protomer A). Drawn with Ligplot+[41].

although these densities also do not closely resemble GSH and their identities remain unknown.

The GSH binding pocket also binds enterovirus inhibitors. Very recently, 'compound-17', a benzene sulphonamide derivative, was identified as a CV-B3 inhibitor (Fig. 5a)[8]. Elaborations of this compound were identified with activity against a range of enteroviruses[8,28], and electron cryo-microscopy revealed the approximate binding site at 4 Å resolution. Comparison of binding sites shows that GSH and compound-17 bind in the same pocket (Fig. 5), although the conformation and interactions of compound-17 are sub-optimal relative to GSH; as expected, given the much lower resolution of the analysis, there are common elements in the binding (Fig. 5).

Discussion

Pan-enterovirus conservation of key residues within the *I* site suggests that it may have potential as a broad-spectrum druggable site. While such conservation encompasses enteroviruses that have been identified as GSH-independent, e.g. EV-A71, it is plausible that GSH can still exert a stabilising effect on the capsid, blocking the transition to the altered state (known as A-particle) that is thought to be a precursor to uncoating and infection[25,29]. In line with this, the antiviral activity of compound-17 is not impeded by increased intracellular glutathione levels, consistent with inhibition of a stage of the infection process which occurs before virus assembly in the cytoplasm, such as uncoating in the endosome. Compounds that increase the stability of the virus capsid may also be useful as vaccine-stabilising additives[23]. Indeed it has been found that oral poliovirus vaccine is stabilised by GSH[23], and thus developing a more potent binder may be useful for the formulation of future virus-like-particle-based vaccines, including, but not limited to, PV[30].

In summary, our structures explain how GSH facilitates initial enteroviral assembly into pentamers. The structural conservation of the *I* site suggests that this protomer glue is widely used, with natural molecules remaining associated at this site for many purified enteroviruses. Together with the recent studies on benzene sulphonamide derivatives[8,28], it is clear that site *I* presents a druggable pocket, and our high-resolution results provide a firm basis for structure-led drug discovery.

Methods

EV-F3 production and purification. BHK-21 cells were grown to 80% confluency before infection with EV-F3 virus stocks in Glasgow modified Eagle's medium (G-MEM BHK-21, Gibco), supplemented with 2% fetal horse serum (Thermo Fisher Scientific) and minimal essential medium non-essential amino acids (MEM NEAA, Gibco). Cells were incubated at 37 °C for 3 days before harvesting. Virus-infected media was first clarified, before overnight PEG-precipitation of virus particles. PEG-precipitated virus was then re-suspended in 50 mM HEPES pH 8.0, 200 mM NaCl, 0.5% NP40 buffer, and particles pelleted on a 30% sucrose cushion. The resulting pellet was re-suspended in minimal 50 mM HEPES pH 8.0, 200 mM NaCl, applied to a 15–45% w/v sucrose gradient and subjected to ultracentrifugation at 4 °C for 3 h, 105,000 × *g*. A clear band at approximately 30% sucrose concentration was observed, extracted and the presence of EV-F3 was confirmed by SDS-Page gel electrophoresis.

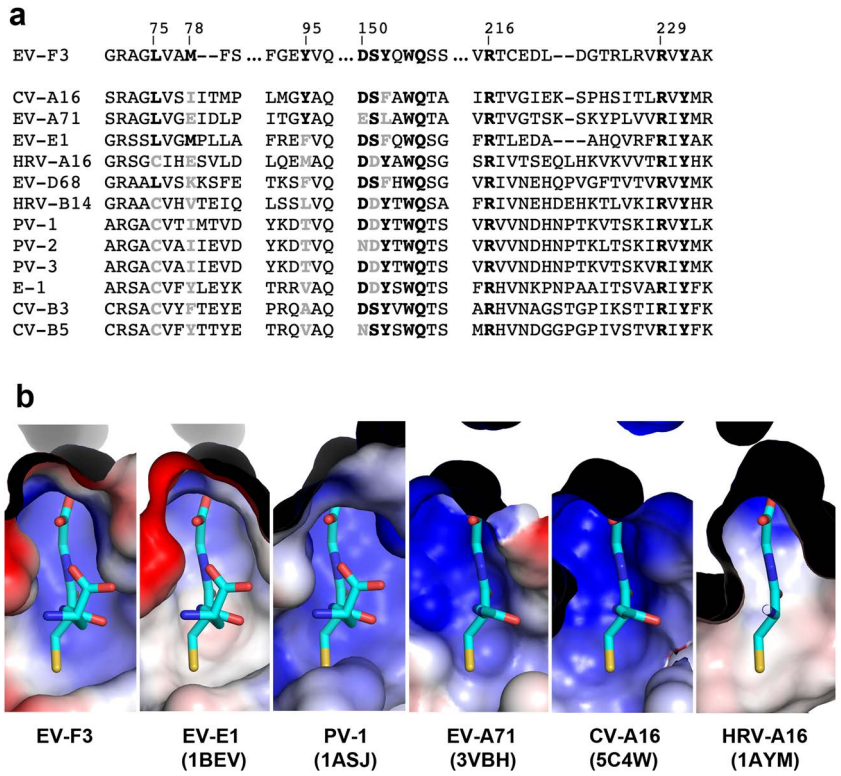


Fig. 4 Site *I* is conserved for major enteroviruses. a Sequence alignment of VP1 for several major enteroviruses. Conserved GSH-interacting residues are in bold. **b** Electrostatic charge (±5 kT e^{-1}) mapped onto the viral surface and coloured blue (positive) and red (negative), for selected enteroviruses, as labelled.

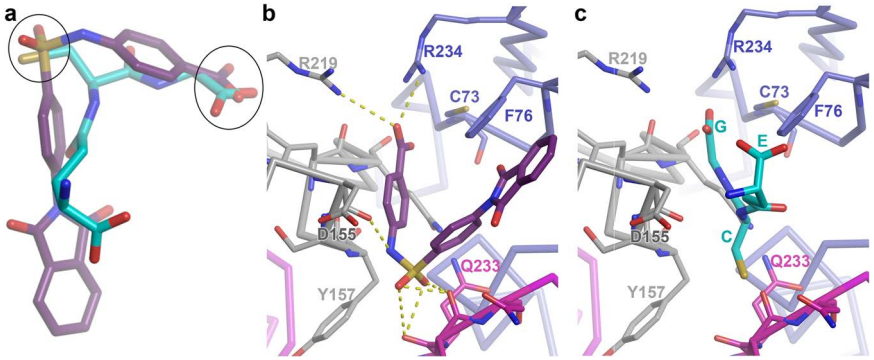


Fig. 5 Compound-17 forms sub-optimal interactions when mapped onto site *I*. a GSH (cyan) superposed on compound-17 (purple) by matching surrounding protein structure[8]. The superpositions of a carboxylic acid group and a sulphur between the two are marked with ellipses. **b** Compound-17 (purple) bound to CVB3 (PDB: 6GZV)[8]. Protomer A VP1 is blue and VP3 magenta. Promoter B is in grey. **c** GSH positioned by superposition of the surrounding protein into CB-V3.

PaSTRy. Solutions of EV-F3 and additive at either 100× or 1000× molar excesses (corresponding to 0.1 mM or 1 mM final concentration) were aliquoted into a 96-well PCR plate containing SYTO9 dye (Thermo Fisher Scientific) in triplicate with 50 mM HEPES, pH 8.0, 200 mM NaCl buffer, before centrifugation for 3 min at 4 °C. The assay was performed by heating samples in a Mx3005p qPCR machine (Agilent Technologies, USA) from 25 to 97 °C in 1 °C min^{-1} increments for 30 s in an expanding saw-tooth profile, a refinement of a previously described protocol[24]. Fluorescence changes were monitored at 25 °C with excitation and emission wavelengths of 492 and 517 nm, respectively.

Infectivity assays. Assays were performed with BSR-T7 cells maintained in G-MEM BHK-21, supplemented with 5% fetal bovine serum (FBS, Life Technology) and MEM NEAA (Gibco) seeded into 96-well plates. For BSO treatment, cells were GSH starved for 48 h before washing and addition of EV-F3 and further 0.4 mM BSO (Sigma). Control plates were also seeded 48 h prior to infection. One hundred microlitres of tenfold serial dilutions of EV-F3 in serum-free G-MEM was added to washed, confluent cells. For BSO-treatment conditions, additional BSO was added to each of the EV-F3 dilutions. CPE could clearly be seen after 12 h (compared with negative controls). Eight replicates of each dilution were plated and were monitored over 2 days at 37 °C and 5% CO_2. Wells were visually scored for virus growth at 17 h post infection (Supplementary Table 1). GSH depletion by BSO eventually disrupts cell integrity and thus it was not possible to maintain assays for more than 2 days.

Crystallisation. Crystals were grown using previously reported protocols[7,31,32]. Briefly, sitting drop crystallisations were set up in 96-well trays using a Cartesian robot with a reservoir solution of 1.5 M ammonium sulfate and 0.1 M Tris at pH 8.5, each crystallisation drop containing 100 nL each of protein and reservoir solutions.

Crystals of approximately 50–80 µm diameter were soaked overnight with either (1) GSH; (2) CG; or (3) GSH/CG at 10 mM (thus for experiment 3 the total concentration of ligand was 20 mM) and then cryo-protected using sequential addition of glycerol from 10 to 25% (v/v). The compound to be soaked was maintained during the cryo-protection process.

X-ray data collection. Diffraction data were collected mainly using automated X-ray centring and data collection at 100 K at I03 beamline, Diamond Light Source (Didcot, UK), with an X-ray wavelength of 0.976 Å. Diffraction images of 0.03° or 0.05° rotation were recorded on either a Pilatus3 6M, or Eiger2 X 16M detector (see Table 1) using a beam size of 80(H) × 20(V) μm^2. A 0.0075 s or 0.01 s exposure time per 0.03 or 0.05° image with 100% beam transmission, under cryogenic conditions, enabled the collection of up to 180° of data before radiation damage was observed. Data from a single crystal were used for experiment (2) whilst data from two crystals were merged for experiments (1) and (3) (Table 1).

X-ray data processing. Data were indexed and integrated using either the xia2-dials pipeline (experiment (3))[33,34], or XDS[35], before scaling and merging with AIMLESS (CCP4i)[36,37]. Structures were initially determined by rigid body fitting of the EV-F3 XFEL structure (PDB: 5OSN)[7] using increasing resolution cut-offs in CNS[38]. Some unit cell shrinkage was observed relative to the room-temperature XFEL structure, consistent with these data being collected under cryogenic conditions[7]. Cyclic positional and B-factor refinement with strict NCS constraints was then performed with CNS, and models were improved with COOT[39,40]. For experiments 1, 2 and 3, respectively, the percentages of residues in the favoured region of the Ramachandran plot are 94.5%, 95.1% and 94.9%. If the allowed regions are included, these percentages rise to 98.6%, 99.2% and 99.2%, respectively. Ramachandran and rotamer outliers for experiment 2 are shown in Supplementary Fig. 2, confirming our modelling. Other statistics are provided in Table 1. All structural figures were constructed using PyMol[35].

Statistics and reproducibility. Infectivity assays were performed with eight repeats. PaSTRy assays were performed in triplicate, with mean values and standard deviations shown in Supplementary Fig. 1. Associated melting temperatures were determined as the temperature half way between the bottom and top of each Boltzmann sigmoidal fit in GraphPad Prism.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Coordinates and structure factors for EV-F3 under conditions (1), (2) and (3) are deposited in the PDB with accession codes 6T4C, 6T40 and 6T48, respectively.

Received: 9 August 2019; Accepted: 2 December 2019;
Published online: 03 January 2020

References

1. Patel, M., Zipursky, S., Orenstein, W., Garon, J. & Zaffran, M. Polio endgame: the global introduction of inactivated polio vaccine. *Expert Rev. Vaccines* **14**, 749–762 (2015).
2. Mao, Q., Wang, Y., Bian, L., Xu, M. & Liang, Z. EV-A71 vaccine licensure: a first step for multivalent enterovirus vaccine to control HFMD and other severe diseases. *Emerg. Microbes Infect.* **5**, 1–7 (2016).
3. Yang, T. et al. A comparative study of multiple clinical enterovirus 71 isolates and evaluation of cross protection of inactivated vaccine strain FY-23 KB in vitro. *Virol. J.* **14**, 206 (2017).
4. Zhu, F. et al. Efficacy, safety, and immunogenicity of an enterovirus 71 vaccine in China. *N. Engl. J. Med.* **370**, 818–828 (2014).
5. Abzug, M. J. The enteroviruses: problems in need of treatments. *J. Infect.* **68**, S108–S114 (2014).
6. De Colibus, L. et al. More-powerful virus inhibitors from structure-based analysis of HEV71 capsid-binding molecules. *Nat. Struct. Mol. Biol.* **21**, 282–288 (2014).
7. Roedig, P. et al. High-speed fixed-target serial virus crystallography. *Nat. Methods* **14**, 805–810 (2017).
8. Abdelnabi, R. et al. A novel druggable interprotomer pocket in the capsid of rhino-and enteroviruses. *PLoS Biol.* **17**, e3000281 (2019).
9. Meister, A. & Anderson, M. E. Glutathione. *Annu. Rev. Biochem.* **52**, 711–760 (1983).
10. Cao, Z. et al. Isochlorogenic acid C prevents enterovirus 71 infection via modulating redox homeostasis of glutathione. *Sci. Rep.* **7**, 16278 (2017).
11. Beck, M. A., Williams-Toone, D. & Levander, O. A. Coxsackievirus B3-resistant mice become susceptible in Se/vitamin E deficiency. *Free Radic. Biol. Med.* **34**, 1263–1270 (2003).
12. Beck, M. A., Levander, O. A. & Handy, J. Selenium deficiency and viral infection. *J. Nutr.* **133**, 1463S–1467S (2003).
13. Beck, M. A., Handy, J. & Levander, O. A. Host nutritional status: the neglected virulence factor. *Trends Microbiol.* **12**, 417–423 (2004).
14. Smith, A. D. & Dawson, H. Glutathione is required for efficient production of infectious picornavirus virions. *Virology* **353**, 258–267 (2006).
15. Mikami, T., Satoh, N., Hatayama, I. & Nakane, A. Buthionine sulfoximine inhibits cytopathic effect and apoptosis induced by infection with human echovirus 9. *Arch. Virol.* **149**, 1117–1128 (2004).
16. Thibaut, H. J. et al. Binding of glutathione to enterovirus capsids is essential for virion morphogenesis. *PLoS Pathog.* **10**, e1004039 (2014).
17. Ma, H.-C. et al. An interaction between glutathione and the capsid is required for the morphogenesis of C-cluster enteroviruses. *PLoS Pathog.* **10**, e1004052 (2014).
18. Kalter, S. S. Animal "orphan" enteroviruses. *Bull. World Health Organ.* **22**, 319–337 (1960).
19. Owen, R. L. et al. Exploiting fast detectors to enter a new dimension in room-temperature crystallography. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **70**, 1248–1256 (2014).
20. Owen, R. L. et al. Outrunning free radicals in room-temperature macromolecular crystallography. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **68**, 810–818 (2012).
21. Axford, D. et al. In situ macromolecular crystallography using microbeams. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **68**, 592–600 (2012).
22. Burkhardt, A. et al. Structure determination from a single high-pressure-frozen virus crystal. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **69**, 308–312 (2013).
23. Abdelnabi, R., Delang, L. & Neyts, J. Glutathione is a highly efficient thermostabilizer of poliovirus Sabin strains. *Vaccine* **35**, 1370–1372 (2017).
24. Walter, T. S. et al. A plate-based high-throughput assay for virus stability and vaccine formulation. *J. Virol. Methods* **185**, 166–170 (2012).
25. Wang, X. et al. A sensor-adaptor mechanism for enterovirus uncoating from structures of EV71. *Nat. Struct. Mol. Biol.* **19**, 424–429 (2012).
26. Ren, J. et al. Structures of coxsackievirus A16 capsids with native antigenicity: implications for particle expansion, receptor binding, and immunogenicity. *J. Virol.* **89**, 10500–10511 (2015).
27. Lacroix, C. et al. A novel benzonitrile analogue inhibits rhinovirus replication. *J. Antimicrobial Chemother.* **69**, 2723–2732 (2014).
28. Ma, Y. et al. New class of early-stage enterovirus inhibitors with a novel mechanism of action. *Antivir. Res.* **147**, 67–74 (2017).
29. Curry, S., Chow, M. & Hogle, J. M. The poliovirus 135S particle is infectious. *J. Virol.* **70**, 7125–7131 (1996).
30. Marsian, J. et al. Plant-made polio type 3 stabilized VLPs—a candidate synthetic polio vaccine. *Nat. Commun.* **8**, 2041–1723 (2017).
31. Walter, T. S. et al. A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **61**, 651–657 (2005).
32. Brown, J. et al. A procedure for setting up high-throughput nanolitre crystallization experiments. II. Crystallization results. *J. Appl. Crystallogr.* **36**, 315–318 (2003).
33. Winter, G., Lobley, C. M. & Prince, S. M. Decision making in xia2. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **69**, 1260–1273 (2013).
34. Winter, G. et al. DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. D* **74**, 85–97 (2018).
35. Kabsch, W. Xds. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **66**, 125–132 (2010).
36. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **69**, 1204–1214 (2013).
37. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **67**, 235–242 (2011).
38. Brunger, A. T. et al. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **54**, 905–921 (1998).
39. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **60**, 2126–2132 (2004).
40. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **66**, 486–501 (2010).
41. Laskowski, R. A. & Swindells, M. B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model* **51**, 2778–2786 (2011).

Acknowledgements

We thank Diamond Light Source for beamtime (proposal mx19946), and the staff of beamline I03 for assistance. H.M.E.D. is supported by Wellcome (ALR00750), D.I.S., E.E.F. and T.S.W. by the UK Medical Research Council (MR/N00065X/1), and J.R. by Wellcome (101122/Z/13/Z).

Author contributions

Experiments were jointly conceived, H.M.E.D. and T.S.W. performed experiments. H.M.E.D., J.R., D.I.S. and E.E.F. collected data and H.M.E.D. and J.R. processed the data. H.M.E.D., J.R. and D.I.S. analysed the results and H.M.E.D., D.I.S., J.R. and E.E.F. wrote the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s42003-019-0722-x.

Correspondence and requests for materials should be addressed to D.I.S.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020

analytical chemistry

Article

Cite This: Anal. Chem. 2018, 90, 7325−7331

pubs.acs.org/ac

Signature of Antibody Domain Exchange by Native Mass Spectrometry and Collision-Induced Unfolding

Yasunori Watanabe,[†,§,⊥] Snezana Vasiljevic,[†] Joel D. Allen,[§] Gemma E. Seabright,[†,§] Helen M. E. Duyvesteyn,[⊥,#] Katie J. Doores,[∥] Max Crispin,*[,§] and Weston B. Struwe*[,†,‡]

[†]Oxford Glycobiology Institute, Department of Biochemistry, and [‡]Chemistry Research Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3QU, United Kingdom

[§]Biological Sciences & The Institute for Life Sciences, University of Southampton, Southampton SO17 1BJ, United Kingdom

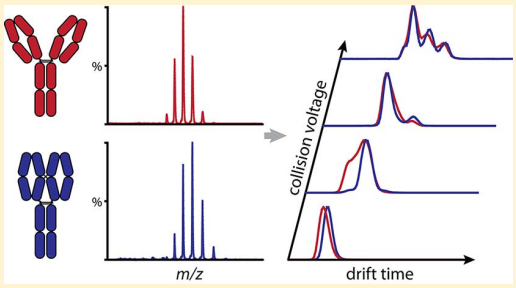
[∥]Department of Infectious Diseases, King's College London, London SE1 9RT, United Kingdom

[⊥]Division of Structural Biology, Wellcome Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, United Kingdom

[#]Diamond Light Source, Harwell Science and Innovation Campus, Didcot OX11 0DE, United Kingdom

Supporting Information

ABSTRACT: The development of domain-exchanged antibodies offers a route to high-affinity targeting to clustered multivalent epitopes, such as those associated with viral infections and many cancers. One strategy to generate these antibodies is to introduce mutations into target antibodies to drive domain exchange using the only known naturally occurring domain-exchanged anti-HIV (anti-human immunodeficiency virus) IgG1 antibody, 2G12, as a template. Here, we show that domain exchange can be sensitively monitored by ion-mobility mass spectrometry and gas-phase collision-induced unfolding. Using native 2G12 and a mutated form that disrupts domain exchange such that it has a canonical IgG1 architecture (2G12 I19R), we show that the two forms can be readily distinguished by their unfolding profiles. Importantly, the same signature of domain exchange is observed for both intact antibody and isolated Fab fragments. The development of a mass spectrometric method to detect antibody domain exchange will enable rapid screening and selection of candidate antibodies engineered to exhibit this and other unusual quaternary antibody architectures.



Antibodies (Ab) have emerged as the leading format for biotherapeutics because of their fine target specificity and their ability to recruit immune effector functions.[1] Most immunoglobulin-based therapeutics are based on the IgG format whereby target recognition is mediated through the variable domains of the fragment antigen binding (Fab) and the fragment crystallizable (Fc) domain mediates immune effector functions via complement activation and/or binding to cellular Fcγ receptors.[2−4]

The Fc activity can be fine-tuned or silenced by both protein and glycan engineering.[5−8] In this way, therapeutic antibodies can be designed that either have minimal immunological effect or can exhibit elevated immune recruitment for maximal cell killing. In contrast, much less attention has been paid to exploring enhancing the target selectivity beyond that achieved by a natural Fab architecture. One approach to circumvent this limitation has been the development of bispecific antibodies (bsAb) where target selectivity is enhanced by Fabs of differing specificity.[9−11] A further underexplored architecture is the use of domain-exchanged IgG whereby the Variable domains from the two heavy chains (V_H and $V_{H'}$) are exchanged, leading to a compact dual Fab.[12] This architecture was discovered during the epitope mapping of 2G12, an antibody that is capable of neutralizing a broad range of human immunodeficiency virus (HIV) isolates (Figure 1A).[13] This Ab is unusual in that the epitope consists of the dense cluster of N-linked glycans on the outer domain of the HIV attachment glycoprotein, gp120.[13,14] 2G12 specifically recognizes terminal Manα1−2Man-linked moieties predominantly found on $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$.[15−17] Domain exchange is critical for effective binding as the target glycans for the canonical Fab binding sites are only 20 Å apart.[15] In addition, domain exchange may generate a new glycan binding surface at the $V_H/V_{H'}$ interface, which further enhances binding.[12,18]

Domain exchange of 2G12 has been verified by X-ray crystallography of the isolated $(Fab)_2$ domain and also size-exclusion chromatography, analytical ultracentrifugation, and electron microscopy.[12,18] A single-point mutation, I19R, completely results in the undomain-exchanged, classical IgG format, and X-ray crystallography of the resulting isolated Fab

Received: February 3, 2018
Accepted: May 14, 2018
Published: May 14, 2018

Downloaded via UNIV OF OXFORD on September 18, 2019 at 02:45:22 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

© 2018 American Chemical Society

DOI: 10.1021/acs.analchem.8b00573
Anal. Chem. 2018, 90, 7325−7331

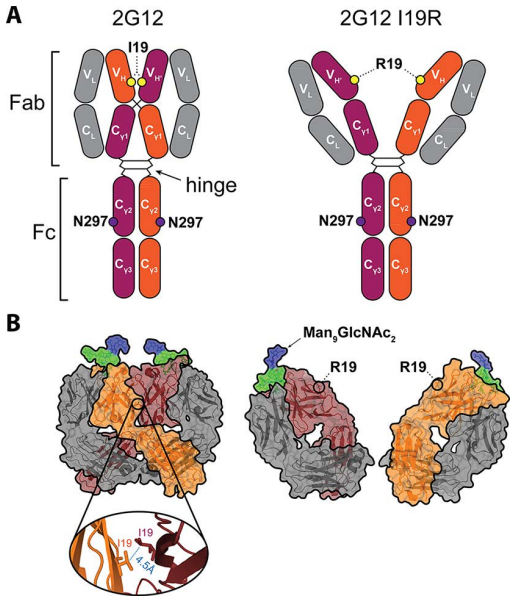


Figure 1. (A) Cartoon structures of domain-exchanged 2G12 and 2G12 I19R antibodies showing the location of Fab, hinge, and Fc regions (domains are labeled with white text). Light chains are gray and heavy chains are colored dark red and orange. The location of amino acid I19/R19 on both heavy chains is labeled yellow and the conserved Asn297 N-linked glycosylation site is purple. (B) Crystal structures of 2G12 (PDB: 1OM3) and 2G12 I19R (PDB: 3OAU) Fab domains with a single $Man_9GlcNAc_2$ glycan bound to each antigen binding site. Mannose, green. GlcNAc, blue.

was shown to exhibit identical V_H/V_L structure compared to the parental domain-exchanged form[19] (Figure 1B). It has been suggested that domain-exchange architecture on antibodies could be used to target multivalent ligands in tumors and pathogens by providing binding specificities that go beyond the classical Fab paradigm. However, identifying domain-exchanged antibodies by aforementioned techniques is low throughput and requires considerable amounts of protein.

Using native ion-mobility mass spectrometry (IM-MS), which separates ions not only by their mass-to-charge ratio but also by their three-dimensional shape, and collision-induced unfolding (CIU), which activates selected ions prior to IM separation, we sought to explore whether domain-exchanged antibody architecture of 2G12 could be distinguished. Compared to other MS-based approaches, native MS maintains proteins in their intrinsically folded state and is an emerging method for biotherapeutic analysis.[20−23] IM-MS has proven useful to study thermal stability,[24] dynamics,[25,26] interactions,[27] and biosimilarity[28] of intact monoclonal antibodies. CIU experiments utilize both collision activation to induce protein unfolding and IM separation which is based on the rotationally averaged collision cross section (CCS) of the ion.[29−31] CIU of protein complexes is an efficient tool to detect various protein features, including discriminating IgG glycosylation and disulfide bonding.[32] Additionally, native IM-MS has proven capable of characterizing antibody drug conjugates, specifically assigning drug-to-antibody ratio of trastuxumab emtansine.[33]

With antibodies at the forefront of biotherapeutics, a high-throughput analytical technique for the assessment of their structural properties is desirable.[34] Here we continue the development of IM-MS and CIU as an analytical technique, and show that gas-phase methods offer a route to rapid and sensitive screening of unusual quaternary architectures that arise from antibody engineering. We examine the CIU profiles of 2G12 and 2G12 I19R Abs and confirm that the specific footprints are determined solely by the domain-exchanged architecture of 2G12.

EXPERIMENTAL SECTION

Antibody Expression and Purification. Plasmids encoding antibody heavy and light chains of 2G12 and 2G12 I19R, b12, and trastuzumab were transiently expressed in HEK 293F cells at a density of 1×10^6 cells/mL. Transfection was done with a 1:2 construct ratio (heavy to light chain) and cells were grown for 5 days. Cells were pelleted and the supernatant was loaded onto a HiTrap Protein A column (GE Healthcare) using the manufacturer's protocol. The binding buffer was 20 mM sodium phosphate, pH 7.5, and the elution buffer was 0.1 M citric acid, pH 3. Antibodies were immediately neutralized with 1 M Tris-HCl, pH 9, prior to buffer exchange into phosphate buffered saline (PBS) using a 50 kDa cutoff Vivaspin column (Sartorius). Antibodies were stored at 4 °C until prior to analysis.

Enzyme-Linked Immunosorbent Assay. ELISAs were performed using 96-well Ni-NTA HisSorb plates (Qiagen). Plates were incubated overnight at 4 °C with 10 μg of HIV gp120, expressed as described.[35] Plates were washed four times with 0.5% PBS/Tween and blocked with 5% skim milk/PBS for 1 h at room temperature. Serial dilutions of 2G12 and 2G12 I19R antibodies in 5% skim milk/PBS were incubated for an hour. Unbound antibody was removed by washing as described above. Horseradish peroxidase conjugated antihuman goat IgG (ThermoFisher Scientific) was added at a 1:3000 dilution for 1 h. The chromogen tetramethylbenzidine was added for 5 min and then neutralized using 2 M sulfuric acid. Absorbance was measured at 450 nm.

IdeS Expression and Purification. Recombinant IdeS was expressed in *Escherichia coli* BL21 (DE3) cells. Then 500 mL of MagicMedia (Invitrogen) containing carbenicillin (100 μg/mL) was inoculated with the 12.5 mL of an overnight seed culture. Cells were grown at 30 °C with shaking at 300 rpm for 7 h and then at 25 °C for an additional 24 h. Cells were harvested by centrifugation at 4000*g* for 30 min. Bacterial lysis was performed by sonication at 4 °C in a lysis buffer of 30 mM sodium phosphate, 300 mM NaCl, and pH 8. After centrifugation at 30 000*g* for 40 min, the filtered supernatant was loaded onto a HisTrap column (GE Healthcare) with a binding buffer of 100 mM HEPES, 150 mM NaCl at pH 7.5. Sequential elutions of IdeS were performed using increasing concentrations of imidazole from 20 to 300 mM.

IdeS Cleavage. 2G12 and 2G12 I19R were incubated for 2 h at 37 °C with IdeS (IdeS-to-antibody ratio of 1:100) to generate intact Fc and Fab antibody fragments.

Ion-Mobility Mass Spectrometry and Collision-Induced Unfolding. Samples were buffer-exchanged into 200 mM ammonium acetate, pH 6.9, with Bio-Spin 6 columns (Bio-Rad). Additional desalting was accomplished by three sequential washes of 200 mM ammonium acetate using 50 kDa molecular weight cutoff (MWCO) centrifugal spin filters (GE Healthcare). Samples were diluted to 10 μM and analyzed with a Synapt G2Si mass spectrometer (Waters, Manchester, UK). A nanoelectrospray (nESI) ion source was used with gold-coated borosilicate glass capillaries prepared in-house. Instrument settings were as follows: capillary voltages, 1.4−1.8 kV; source temperature, 30 °C; sampling cone, 10 kV; trap

DOI: 10.1021/acs.analchem.8b00573
Anal. Chem. 2018, 90, 7325−7331

direct current bias, 20 V; trap gas flow, 5 mL/min; helium cell gas flow, 5 mL/min; IMS gas flow, 30 mL/min; trap wave velocity, 311 m/s; trap wave height, 4 V. Data acquisition and processing were carried out using MassLynx v4.11 and Driftscope version 2.8 software (Waters, Manchester, UK).

For CIU experiments, collision energy was applied in the trap region prior to IM. The collision voltage was increased from 4 to 200 V in 10 V increments. IM data were collected for 1 min at each collision voltage. Unfolding plots were created using PULSAR software, available from http://pulsar.chem.ox.ac.uk.[36] Mass fitter was utilized using the following parameters: mass, 150 kDa; Z_{avg} = 24; Z_{width} = 1.5; resolution, 900; and maximum intensity, 1. Experiments were grouped according to hierarchy based on collision voltages.

Drift time data were extracted using Driftscope version 2.8 (Waters, Manchester, UK), and statistical analysis was performed using python2.7 version of CIUSuite package.[37] Root-mean-squared-deviation (RMSD) values were calculated using CIUSuite stats and compare modules between 2G12 and 2G12 I19R variants.

Differential Scanning Calorimetry (DSC). Antibody samples were prepared at a concentration of 20 μM in 400 μL of PBS. DSC measurements were performed on a Malvern VP Capillary DSC (Malvern), with an unfolding temperature range of 30−110 °C.

Glycan Analysis. N-Linked glycans were enzymatically cleaved from SDS-PAGE gels using PNGaseF prepared in-house. Released N-glycans were fluorescently labeled with 2-aminobenzoic acid (2-AA) for analysis by hydrophilic interaction chromatography ultra-high-performance liquid chromatography (HILIC-UHPLC) with a Waters Acquity UPLC instrument, as previously described.[38] Glycan compositions were determined using traveling wave (TW) IM-MS measurements performed on a Synapt G2Si instrument (Waters, Manchester, UK). For each analysis 2 μL of sample was cleaned with a Nafion 117 membrane and directly infused by nanoelectrospray ionization (nano-ESI) from gold-coated borosilicate glass capillaries as previously described.[16,39] Instrument settings were as follows: capillary voltage, 0.8−1.0 kV; sample cone, 100 V; extraction cone, 25 V; cone gas, 40 L/h; source temperature, 150 °C; trap collision voltage, 4−160 V; transfer collision voltage, 4 V; trap direct current bias, 35−65 V; IMS wave velocity, 450 m/s; IMS wave height, 40 V; trap gas flow, 2 mL/min; IMS gas flow, 80 mL/min. Data were acquired and processed with MassLynx v4.1 and Driftscope version 2.8 software (Waters, Manchester, UK).

RESULTS AND DISCUSSION

Functional ELISA Assay To Characterize Domain Exchange. Prior to IM-MS and CIU experiments, an ELISA assay was used to confirm the impact of each 2G12 antibody architecture on function. The monovalent I19R variant is able to recognize HIV gp120 monomers but with a lower affinity to the divalent domain-exchanged 2G12.[15] Wells were coated with gp120 (BG505 clade A strain) which has 24 N-glycosylation sites and binding was measured from serial dilutions of the antibodies. Expectedly, binding to gp120, which displays the oligomannose-type glycans of the 2G12 epitope,[40−42] differs between 2G12 and 2G12 I19R (Figure 2). These data are consistent with the expected 2G12 and 2G12 I19R architectures and the material was suitable for subsequent MS experiments.

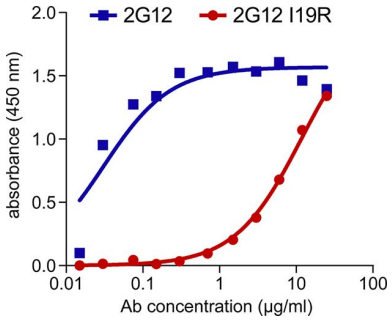


Figure 2. Binding assay of 2G12 (blue) and 2G12 I19R (red) to recombinant HIV gp120. ELISAs were performed with Ni^+-immobilized gp120 and serial Ab dilutions.

Differential Scanning Calorimetry of 2G12 and 2G12 I19R. Differential scanning calorimetry (DSC), a technique that measures unfolding in solution because of heat denaturation, was used to explore the potential for different thermodynamic properties between 2G12 mAbs. DSC revealed that 2G12 I19R unfolded in the manner previously observed for typical IgG1 antibodies,[43,44] with three transitions: first, C_H2 domain unfolding (T_m = 61.0 °C), followed by Fab unfolding (T_m = 72.9 °C) and unfolding of the C_H3 domain (T_m = 86.0 °C) (Figure 3). The domain-exchanged 2G12 yielded only two

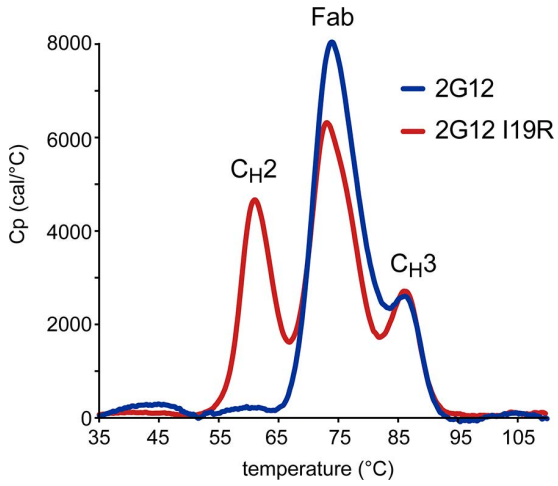


Figure 3. DSC of 2G12 (blue) and 2G12 I19R (red). Cp, specific heat capacity. Peaks are labeled based on previously reported IgG1 assignments.[43,44]

transitions, with the first unfolding state occurring at 73.7 °C and the second at 86.0 °C. The C_H2 domain and the Fab region of 2G12 unfolds at the same temperature, possibly because of the domain-exchanged architecture, while the C_H3 domain unfolds at the same temperature as 2G12 I19R. Importantly DSC unfolding profiles can range significantly among typical IgGs[43,45] but these results point to different stability properties between 2G12 and 2G12 I19R.

Native Mass Spectrometry and Ion Mobility of 2G12 and 2G12 I19R. We next tested gas-phase separation of both antibodies by ion-mobility MS. The native MS spectra of 2G12 and 2G12 I19R were characteristic of an intact folded IgG1 with charge states ranging between 22^+−26^+ (2G12) and 21^+−25^+ (2G12 I19R) (Figure 4A). The 2G12 charge-state distribution was slightly higher with 24^+ as the major charge

DOI: 10.1021/acs.analchem.8b00573
Anal. Chem. 2018, 90, 7325−7331

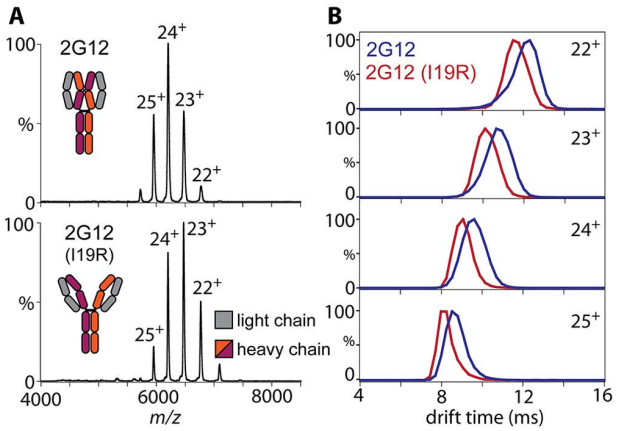


Figure 4. (A) Native mass spectra of 2G12 and 2G12 I19R. (B) Arrival time distributions of 2G12 (blue) and 2G12 I19R (red) 22^+–25^+ charge states.

state, whereas for 2G12 I19R the 23^+ charge state was foremost, followed closely by 24^+. In principle, the difference in distribution of charge could be attributed to differences in available surface area (during electrospray desolvation) exhibited by undomain-exchanged (2G12 I19R) or domain-exchanged (2G12) Fab architectures. Interestingly, the IM drift times of 2G12 were greater between all charge states (22^+–25^+) which directly infers a larger CCS compared to the canonical undomain-exchanged IgG (2G12 I19R) architecture (Figure 4B). Although subtle, the intact MS analysis indicates differences exist between 2G12 and 2G12 I19R; however, these data alone are inconclusive to recognize domain exchange among unknown IgGs.

Collision-Induced Unfolding Signatures of Domain-Exchanged Antibodies. We next assessed CIU differences between 2G12 and 2G12 I19R (Figure 5). In a CIU experiment

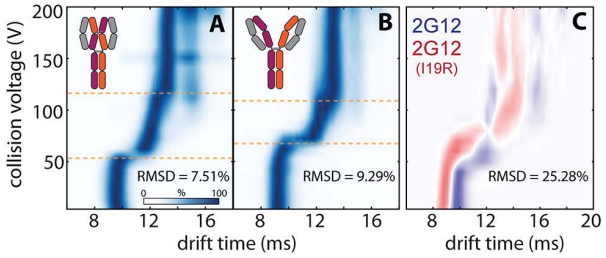


Figure 5. Collision-induced unfolding profiles of the 24^+ charge state of 2G12 (A) and 2G12 I19R (B). The orange dotted lines indicate the collision voltage at the two observed transition states. The blue gradient represents normalized ion intensities for each voltage scan. (C) Difference overlay of 2G12 (blue) and 2G12 I19R (red) CIU profiles.

the unfolding of a protein (induced by increased collision voltages) is monitored as a function of drift time. The resulting two-dimensinal data follow the progressive unfolding or change in size from the initial folded protein ion. The relative abundance of each transition is presented (Figure 5, blue line) over the course of each applied voltage. From the initial folded state, the CIU plot follows unfolded intermediates, sometimes referred to as transitions, and represents gas-phase stability of a protein complex.

The CIU profiles of 2G12 and 2G12 I19R 24^+ charge state revealed differences in gas-phase unfolding with 2G12 undergoing its first unfolding transition at a lower collision voltage (50 V, Figure 5A), whereas 2G12 I19R started to unfold at 70 V (Figure 5B). The second 2G12 unfolded intermediate was maintained to 120 V compared to 2G12 I19R which unfolded earlier (110 V). Consistent with Figure 2, the drift times of the initial folded state differed (2G12 I19R = 9.1 ms, 2G12 = 9.6 ms), but interestingly, the drift times of the second intermediate were similar for both Abs (~12.5 ms). This could be interpreted as both 2G12s unfolding toward a similar configuration or unfolding of the Fab domain during first transition. A third unfolded state was observed for the 24^+ with some low-intensity signal detected which may correspond to a fourth state (~16 ms). The choice of charge state used for antibody CIU experiments has been previously described[34] and influences the amount and features of transitions. This differential behavior between domain-exchanged and non-exchanged antibodies of the 24^+ charge state showed clear differences in gas-phase stability and unfolding existed between the two Ab forms. CIU experiments were performed using two biological replicates (two separate Ab expressions) for both 2G12 antibodies, and the root-mean-squared deviation (RMSD) values were calculated using the CIUSuite software.[37] The RMSD values were 7.51% and 9.29% for 2G12 and 2G12 I19R, respectively, suggesting minimal differences within each replicate.[32] The RMSD value for the difference plot was much greater (RMSD = 25.28%), revealing the considerable differences between the two 2G12 Abs.

Glycosylation Analysis of 2G12 and 2G12 I19R. Antibody glycosylation has been shown to affect CIU fingerprinting of antibodies[32,34] and may account for different unfolding profiles observed for the 2G12-based architectures. Although 2G12 and 2G12 I19R differ by a single amino acid substitution, it is conceivable that Fc glycosylation may account for CIU discrepancies, yet any glycan-specific-induced stability effects in the gas phase is unclear. To confirm the glycosylation status of both 2G12 and 2G12 I19R, N-glycans were released by PNGase F digestion and analyzed by HILIC-UHPLC (Figure 6). UHPLC analysis of fluorescently labeled glycans provided relative amounts of each glycan and no major differences were observed between 2G12 antibodies. Both 2G12 and 2G12 I19R glycans were equivalent and also similar to typical IgG1s with the fucosylated a-, mono-, and digalactosylated biantennary glycans being the most dominant.[46] However, 2G12 exhibited slightly more $Man_5GlcNAc_2$ glycan than 2G12 I19R but no previous influence on this glycan on CIU has been reported. Individual glycan structures were further characterized by IM-MS (Figure S1, Supporting Information).

Collision-Induced Unfolding of 2G12 and 2G12 I19R Fab and Fc domains. To confirm that the observations made from the CIU profiles in Figure 5 were solely due to the domain-exchanged architecture rather than other factors, an IgG degrading enzyme from *Streptococcus pyogenes* (IdeS) cleavage was used to isolate the two regions. IdeS specifically cleaves IgGs between the Fc and Fab at a single site at the lower hinge region, leaving both fragments intact.[48] After incubation with IdeS, the Fab and Fc regions were unfolded as before, with the Fc domains of both 2G12 and 2G12 I19R unfolding in an identical manner (Figure 7). However, CIU of the Fab regions differed and were analogous to the intact CIU results. Overall, three unfolded states were observed and the

DOI: 10.1021/acs.analchem.8b00573
Anal. Chem. 2018, 90, 7325–7331

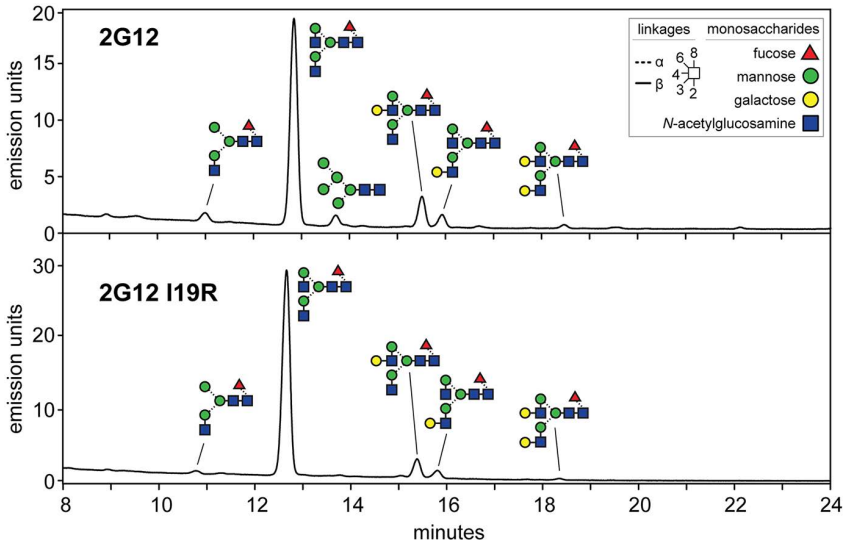


Figure 6. Glycan analysis by HILIC-UHPLC of 2AA-labeled glycans from 2G12 and 2G12 I19R. Peaks were labeled accordingly to previous assignments.[47]

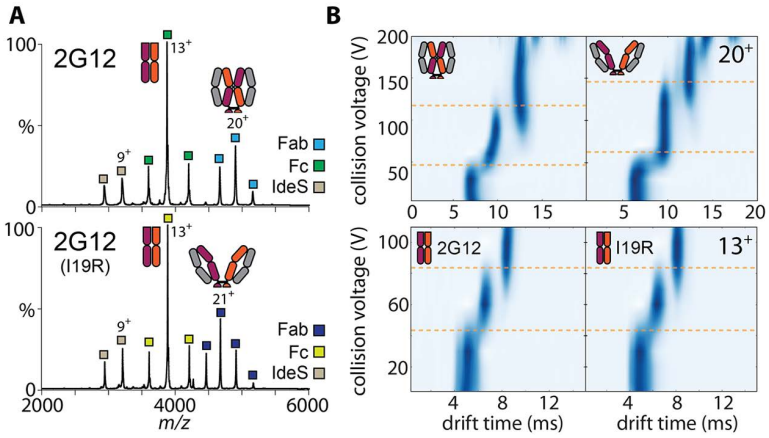


Figure 7. (A) Native mass spectra of 2G12 and 2G12 I19R treated with IdeS. (B) Unfolding plots of 2G12 Fab domains (20^+ charge state) and Fc domains (13^+ charge state). The orange dotted lines indicate the collision voltage at each observed transition states.

domain-exchanged 2G12 Fab unfolded first (55 V) while the 2G12 I19R Fab unfolded at a higher collision energy (70 V). Unfolding to the third state occurred at 125 and 150 V for 2G12 and 2G12 I19R, respectively. On the whole, CIU profiles of both the domain-exchanged intact antibody and isolated Fabs unfolded at lower collision voltages than the non-domain-exchanged IgGs.

Collision-Induced Unfolding of 2G12 I19R Is Equivalent to Standard IgG Antibodies. To confirm that the CIU profiles of 2G12 I19R were consistent with other IgG1 antibodies, two representative antibodies were tested. The first is the well-known IgG1 therapeutic antibody, trastuzumab, and the other is an HIV-neutralizing IgG1 antibody, b12[49,50] (Figure 8). The CIU profiles of the 24^+ charge state were similar to trastuzumab, b12, and 2G12 I19R unfolding first at a collision voltage of 70 V. The second transition occurred at a collision voltage of 120 V. The unfolding patterns of b12, trastuzumab, and 2G12 I19R were observed across all other major charge states (22^+, 23^+, and 25^+) (Figure S2) and differed from 2G12. As discussed above, charge-state selection for CIU fingerprinting is an important factor when comparing antibodies.

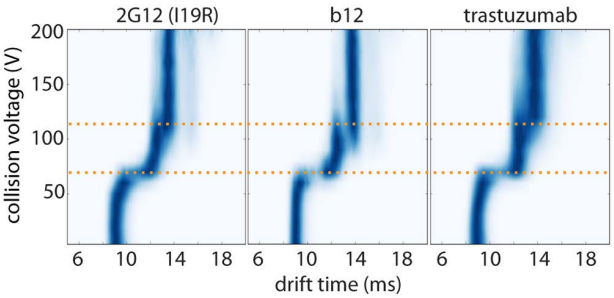


Figure 8. Collision-induced unfolding profiles of conventional IgG1 antibodies b12 and trastuzumab alongside 2G12 I19R (24^+ charge state). The dashed orange lines represent the unfolding transitions.

DOI: 10.1021/acs.analchem.8b00573
Anal. Chem. 2018, 90, 7325−7331

CONCLUSIONS

Identification of Ab structural features has generally been approached by classical techniques including X-ray crystallography and electron microscopy. Further structural details have been investigated using thermodynamic, hydrodynamic, scattering, and other spectroscopic and mass spectrometric methodologies. Here, we demonstrate that native mass spectrometry and collision-induced unfolding can complement these established methods by rapidly and sensitively distinguishing between domain-exchanged and nonexchanged antibody architectures. We further demonstrate that the change in gas-phase unfolding is localized to Fab domains, identified specifically after IdeS cleavage. The characteristics of the full-length antibody CIU profiles are consistent between canonical IgG1 structures, thereby providing a benchmark for screening for antibody features, including domain-exchanged architectures, extended CDRs, and potentially the correct assembly of multispecific antibodies. The applications for CIU in characterization of unusual antibody architectures offers the potential for this technique to become a high-throughput screening tool during the development of macromolecular drugs.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.8b00573.

Figures showing the IM-MS spectra of 2G12 and 2G12 I19R N-glycans and CIU data of b12, trastuzumab, 2G12 I19R, and 2G12 charge states not presented in the main text (PDF)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: weston.struwe@chem.ox.ac.uk (W.B.S.).
*E-mail: max.crispin@soton.ac.uk (M.C.).

ORCID

Max Crispin: 0000-0002-1072-2694
Weston B. Struwe: 0000-0003-0594-226X

Author Contributions

W.S. and M.C. conceived the project. Y.W., M.C., and W.S. designed experiments. S.V. and K.D. provided key reagents. Y.W., J.A., and W.S. performed intact mass spectrometry experiments. H.M.E.D. completed statistical analyses. G.S. performed UPLC analysis. Y.W., M.C., and W.S. produced the first draft of the paper. All authors contributed to and approved the final draft.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We kindly thank David Staunton for assisting in DSC measurements (University of Oxford, Department of Biochemistry, Biophysics Facility). M.C. and W.S. gratefully acknowledge a research grant from Against Breast Cancer (www.againstbreastcancer.org; UK Charity 1121258). H.M.E.D (studentship ALR00750-B500.1), was supported by the Wellcome Trust. M.C. was the Against Breast Cancer Fellow in Cancer Therapeutics at Oriel College, Oxford. M.C. is additionally supported by Scripps CHAVI-ID (Grant 1UM1AI100663), which provided funds to purchase the Synapt G2Si mass spectrometer.

REFERENCES

(1) Ecker, D. M.; Jones, S. D.; Levine, H. L. *mAbs* **2015**, *7*, 9−14.
(2) Carter, P. J.; Lazar, G. A. *Nat. Rev. Drug Discovery* **2017**, *17*, 197−223.
(3) Nelson, A. L.; Dhimolea, E.; Reichert, J. M. *Nat. Rev. Drug Discovery* **2010**, *9*, 767−774.
(4) Sliwkowski, M. X.; Mellman, I. *Science* **2013**, *341*, 1192−1198.
(5) Blundell, P. A.; Le, N. P. L.; Allen, J.; Watanabe, Y.; Pleass, R. J. *J. Biol. Chem.* **2017**, *292*, 12994−13007.
(6) Le, N. P. L.; Bowden, T. A.; Struwe, W. B.; Crispin, M. *Biochim. Biophys. Acta, Gen. Subj.* **2016**, *1860*, 1655−1668.
(7) Li, T.; DiLillo, D. J.; Bournazos, S.; Giddens, J. P.; Ravetch, J. V.; Wang, L. X. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 3485−3490.
(8) Jefferis, R. *J. Immunol. Res.* **2016**, *2016*, 5358272.
(9) Brinkmann, U.; Kontermann, R. E. *mAbs* **2017**, *9*, 182−212.
(10) Chames, P.; Baty, D. *mAbs* **2009**, *1*, 539−547.
(11) Fan, G.; Wang, Z.; Hao, M.; Li, J. *J. Hematol. Oncol.* **2015**, *8*, 130.
(12) Calarese, D. A.; Scanlan, C. N.; Zwick, M. B.; Deechongkit, S.; Mimura, Y.; Kunert, R.; Zhu, P.; Wormald, M. R.; Stanfield, R. L.; Roux, K. H.; Kelly, J. W.; Rudd, P. M.; Dwek, R. A.; Katinger, H.; Burton, D. R.; Wilson, I. A. *Science* **2003**, *300*, 2065−71.
(13) Scanlan, C. N.; Pantophlet, R.; Wormald, M. R.; Ollmann Saphire, E.; Stanfield, R.; Wilson, I. A.; Katinger, H.; Dwek, R. A.; Rudd, P. M.; Burton, D. R. *J. of Virol.* **2002**, *76*, 7306−7321.
(14) Sanders, R. W.; Venturi, M.; Schiffner, L.; Kalyanaraman, R.; Katinger, H.; Lloyd, K. O.; Kwong, P. D.; Moore, J. P. *J. of Virol.* **2002**, *76*, 7293−7305.
(15) Doores, K. J.; Fulton, Z.; Huber, M.; Wilson, I. A.; Burton, D. R. *J. Virol.* **2010**, *84*, 10690−10699.
(16) Dunlop, D. C.; Bonomelli, C.; Mansab, F.; Vasiljevic, S.; Doores, K. J.; Wormald, M. R.; Palma, A. S.; Feizi, T.; Harvey, D. J.; Dwek, R. A.; Crispin, M.; Scanlan, C. N. *Glycobiology* **2010**, *20*, 812−823.
(17) Scanlan, C. N.; Pantophlet, R.; Wormald, M. R.; Ollmann Saphire, E.; Stanfield, R.; Wilson, I. A.; Katinger, H.; Dwek, R. A.; Rudd, P. M.; Burton, D. R. *J. of Virol.* **2002**, *76*, 7306−7321.
(18) Murin, C. D.; Julien, J.-P.; Sok, D.; Stanfield, R. L.; Khayat, R.; Cupo, A.; Moore, J. P.; Burton, D. R.; Wilson, I. A.; Ward, A. B. *J. Virol.* **2014**, *88*, 10177−10188.
(19) Huber, M.; Le, K. M.; Doores, K. J.; Fulton, Z.; Stanfield, R. L.; Wilson, I. A.; Burton, D. R. *J. Virol.* **2010**, *84*, 10700−10707.
(20) Ehkirch, A.; D'Atri, V.; Rouviere, F.; Hernandez-Alba, O.; Goyon, A.; Colas, O.; Sarrut, M.; Beck, A.; Guillarme, D.; Heinisch, S.; Cianferani, S. *Anal. Chem.* **2018**, *90*, 1578−1586.
(21) Ferguson, C. N.; Gucinski-Ruth, A. C. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 822−833.
(22) Schachner, L.; Han, G.; Dillon, M.; Zhou, J.; McCarty, L.; Ellerman, D.; Yin, Y.; Spiess, C.; Lill, J. R.; Carter, P. J.; Sandoval, W. *Anal. Chem.* **2016**, *88*, 12122−12127.
(23) Terral, G.; Beck, A.; Cianferani, S. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2016**, *1032*, 79−90.
(24) Pacholarz, K. J.; Peters, S. J.; Garlish, R. A.; Henry, A. J.; Taylor, R. J.; Humphreys, D. P.; Barran, P. E. *ChemBioChem* **2016**, *17*, 46−51.
(25) Edgeworth, M. J.; Phillips, J. J.; Lowe, D. C.; Kippen, A. D.; Higazi, D. R.; Scrivens, J. H. *Angew. Chem., Int. Ed.* **2015**, *54*, 15156−15159.
(26) Pacholarz, K. J.; Porrini, M.; Garlish, R. A.; Burnley, R. J.; Taylor, R. J.; Henry, A. J.; Barran, P. E. *Angew. Chem., Int. Ed.* **2014**, *53*, 7765−7769.
(27) Gahoual, R.; Heidenreich, A. K.; Somsen, G. W.; Bulau, P.; Reusch, D.; Wuhrer, M.; Haberger, M. *Anal. Chem.* **2017**, *89*, 5404−5412.
(28) Pisupati, K.; Tian, Y.; Okbazghi, S.; Benet, A.; Ackermann, R.; Ford, M.; Saveliev, S.; Hosfield, C. M.; Urh, M.; Carlson, E.; Becker, C.; Tolbert, T. J.; Schwendeman, S. P.; Ruotolo, B. T.; Schwendeman, A. *Anal. Chem.* **2017**, *89*, 4838−4846.

DOI: 10.1021/acs.analchem.8b00573
Anal. Chem. 2018, 90, 7325−7331

(29) Hopper, J. T. S.; Oldham, N. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1851−1858.
(30) Huang, Y. N.; Salinas, N. D.; Chen, E.; Tolia, N. H.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 2515−2518.
(31) Wagner, N. D.; Clemmer, D. E.; Russell, D. H. *Anal. Chem.* **2017**, *89*, 10094−10103.
(32) Tian, Y.; Han, L.; Buckner, A. C.; Ruotolo, B. T. *Anal. Chem.* **2015**, *87*, 11509−11515.
(33) Marcoux, J.; Champion, T.; Colas, O.; Wagner-Rousset, E.; Corvaia, N.; Van Dorsselaer, A.; Beck, A.; Cianferani, S. *Protein Sci.* **2015**, *24*, 1210−1223.
(34) Dixit, S. M.; Polasky, D. A.; Ruotolo, B. T. *Curr. Opin. Chem. Biol.* **2018**, *42*, 93−100.
(35) Struwe, W. B.; Stuckmann, A.; Behrens, A. J.; Pagel, K.; Crispin, M. *ACS Chem. Biol.* **2017**, *12*, 357−361.
(36) Allison, T. M.; Reading, E.; Liko, I.; Baldwin, A. J.; Laganowsky, A.; Robinson, C. V. *Nat. Commun.* **2015**, *6*, 1−10.
(37) Eschweiler, J. D.; Rabuck-Gibbons, J. N.; Tian, Y.; Ruotolo, B. T. *Anal. Chem.* **2015**, *87*, 11516−11522.
(38) Pritchard, L. K.; Vasiljevic, S.; Ozorowski, G.; Seabright, G. E.; Cupo, A.; Ringe, R.; Kim, H. J.; Sanders, R. W.; Doores, K. J.; Burton, D. R.; Wilson, I. A.; Ward, A. B.; Moore, J. P.; Crispin, M. *Cell Rep.* **2015**, *11*, 1604−1613.
(39) Harvey, D. J.; Crispin, M.; Bonomelli, C.; Scrivens, J. H. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1754−1767.
(40) Behrens, A.-J.; Vasiljevic, S.; Pritchard, L. K.; Harvey, D. J.; Andev, R. S.; Krumm, S. A.; Struwe, W. B.; Cupo, A.; Kumar, A.; Zitzmann, N.; Seabright, G. E.; Kramer, H. B.; Spencer, D. I. R.; Royle, L.; Lee, J. H.; Klasse, P. J.; Burton, D. R.; Wilson, I. A.; Ward, A. B.; Sanders, R. W.; et al. *Cell Rep.* **2016**, *14*, 2695−2706.
(41) Pritchard, L. K.; Harvey, D. J.; Bonomelli, C.; Crispin, M.; Doores, K. J. *J. Virol.* **2015**, *89*, 8932−8944.
(42) Pritchard, L. K.; Spencer, D. I. R.; Royle, L.; Bonomelli, C.; Seabright, G. E.; Behrens, A.-J.; Kulp, D. W.; Menis, S.; Krumm, S. A.; Dunlop, D. C.; Crispin, D. J.; Bowden, T. A.; Scanlan, C. N.; Ward, A. B.; Schief, W. R.; Doores, K. J.; Crispin, M. *Nat. Commun.* **2015**, *6*, 1−11.
(43) Ionescu, R. M.; Vlasak, J.; Price, C.; Kirchmeier, M. *J. Pharm. Sci.* **2008**, *97*, 1414−1426.
(44) Ito, T.; Tsumoto, K. *Protein Sci.* **2013**, *22*, 1542−1551.
(45) Brader, M. L.; Estey, T.; Bai, S.; Alston, R. W.; Lucas, K. K.; Lantz, S.; Landsman, P.; Maloney, K. M. *Mol. Pharmaceutics* **2015**, *12*, 1005−1017.
(46) Stöckmann, H.; Adamczyk, B.; Hayes, J.; Rudd, P. M. *Anal. Chem.* **2013**, *85*, 8841−8849.
(47) Guile, G. R.; Rudd, P. M.; Wing, D. R.; Prime, S. B.; Dwek, R. A. *Anal. Biochem.* **1996**, *240*, 210−226.
(48) Akesson, P.; Moritz, L.; Truedsson, M.; Christensson, B.; von Pawel-Rammingen, U. *Infect. Immun.* **2006**, *74*, 497−503.
(49) Saphire, E. O.; Parren, P. W.; Pantophlet, R.; Zwick, M. B.; Morris, G. M.; Rudd, P. M.; Dwek, R. A.; Stanfield, R. L.; Burton, D. R.; Wilson, I. A. *Science* **2001**, *293*, 1155−1159.
(50) Zwick, M. B.; Parren, P. W. H. I.; Saphire, E. O.; Church, S.; Wang, M.; Scott, J. K.; Dawson, P. E.; Wilson, I. A.; Burton, D. R. *J. of Virol.* **2003**, *77*, 5863−5876.

DOI: 10.1021/acs.analchem.8b00573
Anal. Chem. 2018, 90, 7325−7331

IUCrJ
ISSN 2052-2525
BIOLOGY | MEDICINE

Dose-resolved serial synchrotron and XFEL structures of radiation-sensitive metalloproteins

Ali Ebrahim,[a,b] Tadeo Moreno-Chicano,[a]‡ Martin V. Appleby,[b] Amanda K. Chaplin,[a]§ John H. Beale,[b] Darren A. Sherrell,[b] Helen M. E. Duyvesteyn,[b,c] Shigeki Owada,[d,e] Kensuke Tono,[d,e] Hiroshi Sugimoto,[e] Richard W. Strange,[a] Jonathan A. R. Worrall,[a] Danny Axford,[b] Robin L. Owen[b]* and Michael A. Hough[a]*

[a]School of Biological Sciences, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK, [b]Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxfordshire OX11 0DE, UK, [c]Division of Structural Biology (STRUBI), The Henry Wellcome Building for Genomic Medicine, University of Oxford, Roosevelt Drive, Oxford, Oxfordshire OX3 7BN, UK, [d]RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan, and [e]Japan Synchrotron Radiation Research Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5198, Japan. *Correspondence e-mail: robin.owen@diamond.ac.uk, mahough@essex.ac.uk

Received 6 February 2019
Accepted 22 March 2019

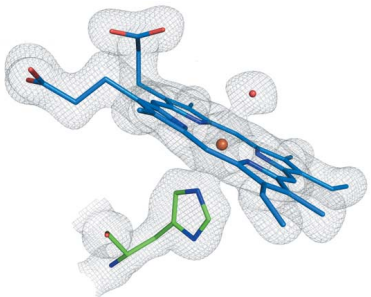
Edited by E. E. Lattman, University at Buffalo, USA

‡ Current address for TMC: Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, Jahnstrasse 29, 69120 Heidelberg, Germany.
§ Current address for AKC: Department of Biochemistry, Sanger Building, 80 Tennis Court Road, Cambridge CB2 1GA, UK.

Keywords: XFELs; microcrystals; serial femtosecond crystallography; serial synchrotron crystallography; serial millisecond crystallography; fixed targets; heme peroxidase; metalloproteins; radiation damage.

PDB references: Dye-type peroxidase Aa from *Streptomyces lividans*, MSS 78.4 kGy structure, 6i8p; 274.4 kGy structure, 6q3e; 235.2 kGy structure, 6q3d; 164 kGy structure, 6i8k; 98.4 kGy structure, 6i8i; 32.8 kGy structure, 6i7z; 39.2 kGy structure, 6i8o; 196 kGy structure, 6q34; 117.6 kGy structure, 6i8q; 65.6 kGy structure, 6i8e; 156.8 kGy structure, 6q31; 131.2 kGy structure, 6i8j; 313.6 kGy structure, 6ibn; SFX structure of damage-free ferric state of dye-type peroxidase Aa from *Streptomyces lividans*, 6i43

Supporting information: this article has supporting information at www.iucrj.org



OPEN ACCESS

An approach is demonstrated to obtain, in a sample- and time-efficient manner, multiple dose-resolved crystal structures from room-temperature protein microcrystals using identical fixed-target supports at both synchrotrons and X-ray free-electron lasers (XFELs). This approach allows direct comparison of dose-resolved serial synchrotron and damage-free XFEL serial femtosecond crystallography structures of radiation-sensitive proteins. Specifically, serial synchrotron structures of a heme peroxidase enzyme reveal that X-ray induced changes occur at far lower doses than those at which diffraction quality is compromised (the Garman limit), consistent with previous studies on the reduction of heme proteins by low X-ray doses. In these structures, a functionally relevant bond length is shown to vary rapidly as a function of absorbed dose, with all room-temperature synchrotron structures exhibiting linear deformation of the active site compared with the XFEL structure. It is demonstrated that extrapolation of dose-dependent synchrotron structures to zero dose can closely approximate the damage-free XFEL structure. This approach is widely applicable to any protein where the crystal structure is altered by the synchrotron X-ray beam and provides a solution to the urgent requirement to determine intact structures of such proteins in a high-throughput and accessible manner.

1. Introduction

Enzymology and structural biology are highly dependent on the accurate three-dimensional models obtained by X-ray crystallography. Such structures provide insight into function and can form a basis for understanding how proteins interact with each other or with small molecules. Fundamentally, the structure obtained should be representative of the native state of the protein. However, macromolecular crystallography is typically carried out at cryogenic temperatures (100 K) to minimize radiation-damage-induced structural perturbation (Garman & Weik, 2017; Holton, 2009). There is an increasing recognition of the importance of determining structures at ambient or 'room' temperature so as to be more representative of the structures and dynamics adopted by proteins *in vivo* at physiological temperature (Keedy *et al.*, 2018, 2015; Fischer *et al.*, 2015; Weik & Colletier, 2010). A major challenge in conventional synchrotron-based X-ray crystallography, particularly at room temperature, is the extremely rapid onset of radiation damage, *i.e.* changes to the structure of the protein caused by the ionizing effects of the X-ray beam (Garman &

Weik, 2017; Holton, 2009). The rapidity of site-specific radiation damage means it is present in differing levels of severity in almost all crystallographic datasets determined using synchrotron radiation, even if great care to avoid it is taken during data collection by minimizing the absorbed dose; the challenge becomes even greater if only microcrystals are available. When collecting data from microcrystals, microbeams of increased brilliance are required for optimal data collection, though the use of such beams comes with a concomitant increase in the rate of X-ray induced changes. It is critical that the site-specific effects of radiolysis are understood in detail and minimized in order to produce structures that are accurately representative of radiation-sensitive proteins *in vivo*. We note however that in some protein crystals that do not contain metals or other redox centres, radiation damage, while present, may cause little change to the observed structure.

It is estimated that approximately one-third of all proteins require a metal ion, with around half of all enzymes utilizing a metal for catalytic function (Waldron *et al.*, 2009). Heme enzymes catalyse many essential reactions in biology and understanding their structures throughout their reaction cycles is of high interest, prompting extensive efforts made to obtain 'intact' structures at high resolution (Chreifi *et al.*, 2016; Casadei *et al.*, 2014; Gumiero *et al.*, 2011; Moody & Raven, 2018). A major challenge is to obtain the higher valence states of these proteins, for example peroxidases (Fe^{III} resting state and Fe^{IV} intermediate states), as these states are phenomenally sensitive to reduction in synchrotron experiments caused by the presence of large numbers of solvated electrons or other radiolytically produced species, generated by the interaction of X-rays with the crystal (Kekilli *et al.*, 2017; Beitlich *et al.*, 2007; Denisov *et al.*, 2007). This site-specific damage is known to occur at doses much lower than those typically required to collect a dataset (Garman & Weik, 2017; Holton, 2009). Structures of peroxidases and other redox-sensitive metalloproteins obtained from synchrotron X-ray crystallography, even at 100 K, are therefore likely to represent a superposition of resting and damaged states. This site-specific damage is an extremely pressing problem if mechanistic conclusions are to be drawn from the structures obtained.

In contrast, X-ray free-electron lasers (XFELs) promise damage and artefact-free crystallography, provided the pulse duration is short enough (Schlichting, 2015; Nass *et al.*, 2015; Lomb *et al.*, 2011). For serial femtosecond crystallography (SFX), data collection from each microcrystal can be completed before site-specific and global radiation damage occurs, but at the expense of longer term crystal destruction, such that a new crystal must be presented to every pulse. Radiation-induced changes have been detected in SFX data with pulse durations as short as 40 fs (Nass *et al.*, 2015) or 70 fs (Lomb *et al.*, 2011). In contrast, data measured with pulse durations of 10 fs or shorter are considered to be free of typical radiation-induced site-specific radiation damage (Halsted *et al.*, 2018; Andersson *et al.*, 2017). Direct comparison of XFEL and synchrotron structures of the same protein presents many challenges because of differences between these methods, such as crystal size, mosaicity, temperature, cryoprotection, crystallization conditions and resolution. The epitome of this being that almost all XFEL structures are obtained from tens of thousands of room-temperature microcrystals while most synchrotron structures are obtained from a single crystal held at 100 K.

We report a new method based on a highly efficient fixed-target silicon nitride chip system (Oghbaey *et al.*, 2016; Mueller *et al.*, 2015). This system allows for data to be measured at room temperature from microcrystals in the same manner using synchrotron or XFEL radiation. Our fixed-target approach enables time- and sample-efficient data collection by both SFX and serial synchrotron crystallography (SSX), and simultaneously minimizes any differences in structure by eliminating the experimental variables outlined above. It is also well suited to tracking functionally relevant changes in redox enzymes as X-ray generated solvated electrons drive the enzyme along the catalytic pathway in the crystal when exposed to the X-ray beam. Multiple serial structures (MSS) can be obtained from a set of crystals on a single fixed target, as sequential exposure events to each crystal are binned and processed as individual dose-dependent datasets. This is analogous to the multiple structures from one crystal (MSOX) approach (Horrell *et al.*, 2016, 2018), previously applied to the measurement of repeated datasets from the same exposed region of a single large crystal to produce a dose series. In comparison, our new approach exposes each crystal to the X-ray beam for only a few tens of milliseconds and is well suited for high-throughput structure solution from microcrystals held at room temperature, using XFEL or synchrotron radiation sources.

Herein, we have chosen to use an extracellular dye-type heme peroxidase found in *Streptomyces lividans* and referred to as DtpAa. Using DtpAa as the exemplar, we describe the application of our method of combined SFX and MSS experiments, though the method can be used for any redox-sensitive system. Starting in the catalytic resting state, our approach reveals multiple well resolved structural states of the enzyme, with a low-dose synchrotron MSS structural series showing clearly resolved changes to the active-site region of the enzyme within tens of milliseconds. Extrapolation of varying structural parameters to zero dose produced a close match to the damage-free structure determined using SFX. Thus, a low-dose series of synchrotron MSS is anchored by a damage-free SFX structure, both being determined using the same fixed-target serial sample-delivery system.

We present this approach as a general method to efficiently collect both SFX and SSX data under near-identical conditions, characterize subtle site-specific changes caused by X-rays in proteins and allow direct comparison of, and extrapolation to, damage-free XFEL structures from low-dose synchrotron models.

2. Materials and methods

2.1. Sample preparation

The SLI_2602 gene encoding DtpAa was amplified from the genomic DNA of *S. lividans* strain 1326 (*S. lividans* stock

number 1326, John Innes Centre) by polymerase chain reaction (the primers used for amplification are reported in the Supporting information). The gene was subsequently cloned into the NdeI and HindIII sites of a pET28a vector (Novagen) to create an N-terminal His_6-tagged construct (pET2602) for overexpression in *Escherichia coli*. The pET2602 vector was transformed into *E. coli* BL21 (DE3) cells. Overnight precultures [low-salt Luria–Bertani (LB) medium; Melford] were successively used to inoculate 1.4 l of high-salt LB medium with 50 mg ml^{-1} kanamycin and were grown at 37°C, 180 rev min^{-1}. At an OD_{600} of 1.0–1.2, 5-aminolevulinic acid (0.25 m*M* final concentration) and iron citrate (100 µ*M* final concentration) were added consecutively for their use as a heme precursor and iron supplement, respectively. Cultures were then induced by adding isopropyl *β*-D-thiogalactopyranoside (Melford) to a final concentration of 0.5 m*M*, and carbon monoxide gas was bubbled through the culture for 30–60 s. Flasks were then sealed and incubated for a further 18 h at 30°C and 100 rev min^{-1}. Cells were harvested via centrifugation (10 000*g*, 10 min, 4°C) and the cell pellet resuspended in 50 m*M* Tris–HCl, 500 m*M* NaCl (Fisher) and 20 m*M* imidazole (Sigma) at pH 8 (buffer A). The resuspended cell suspension was lysed using an EmulsiFlex-C5 cell disrupter (Avestin) followed by centrifugation (22 000*g*, 30 min, 4°C). The clarified supernatant was loaded onto a 5 ml nickel–nitrilotriacetic acid–sepharose column (GE Healthcare) equilibrated with buffer A and eluted by a linear imidazole gradient using buffer B (buffer A with 500 m*M* imidazole). The DtpAa peak eluting at approximately 30–40% buffer B was pooled and concentrated using a Centricon (VivaSpin) with a 10 kDa cut-off at 4°C followed by application to an S200 Sephadex column (GE Healthcare) equilibrated with 20 m*M* NaPi, 100 m*M* NaCl, pH 7. A major peak eluted consistent with a monomeric species with fractions assessed by SDS–PAGE then concentrated and stored at −20°C. DtpAa concentrations were determined by UV–vis spectroscopy (Varian Cary 60 UV–vis spectrophotometer) using an extinction coefficient at 280 nm of 46 075 M^{-1} cm^{-1}.

Microcrystals were grown in batch (typically 0.4–0.5 ml total volume) by mixing in a 1:1 ratio a 6.5 mg ml^{-1} DtpAa protein solution with a precipitant solution containing 20% PEG 6000, 100 m*M* HEPES pH 7.0, to typical dimensions of 15 µm. Silicon nitride fixed-target chips with either 7, 12, 14 or 37 µm apertures at their narrowest opening and a nominal capacity of 25 600 crystals were loaded as previously described (Ebrahim *et al.*, 2019), with an identical loading protocol used both at Diamond and the SPring-8 Ångstrom Free Electron Laser (SACLA) XFEL. In brief, chips were loaded with 100–200 µl of microcrystal suspension within a humidity enclosure (Solo Containment, Cheshire, England) and sealed between two layers of 6 µm thick Mylar.

2.2. Data collection and fixed-target motion

SFX data were measured at SACLA beamline BL2 EH3 using an X-ray energy of 10.0 keV, a pulse length 10 fs and a repetition rate of 30 Hz, with the beam attenuated to 13% of full flux. Chips were translated within the interval between X-ray pulses, ensuring that the chip had stopped at the centre of each crystal position (the centre of the aperture) and was exposed only once to X-rays, before moving to the next position during the next pulse interval. Data were typically collected from all 25 600 positions on a chip in 14 min using the SACLA MPCCD detector (Kameshima *et al.*, 2014), with experiments performed in a helium chamber to minimize air scatter. A modified custom entry port to the helium chamber permitted rapid exchange of chips, meaning that measurement from all positions with subsequent sample exchange and alignment interval of <5 min between data collections allowed a sustained data-collection rate of just over 3 chips per hour. While sufficient data for structure solution and refinement were obtained from crystals mounted on only 2 chips (*ca* 13 000 hits), for the structure described here data were collected from a total of 11 chips, still in under 4 h of beam time, in order to increase the redundancy of the data and the quality of the maps obtained.

Data collection at beamline I24, Diamond Light Source was carried out using an unattenuated X-ray beam of energy 12.8 keV and a Pilatus3 6M detector in shutterless mode. To form a dose-dependent series of DtpAa structures, 5 (MSS-1) and 10 (MSS-2) sequential diffraction patterns were measured at each crystal position each with an exposure time of 10 ms and subsequently binned into one dataset per dose interval (Fig. 1). The series of exposures at each chip position was individually triggered via a Keysight 33500B signal generator which was itself triggered by a DeltaTau Geobrick LV-IMS-II stage controller when each desired crystal position had been attained. The X-ray shutter was not closed between apertures on a chip and remained open for the duration of the experiment. X-ray fluxes were measured using a silicon PIN diode as previously described (Owen *et al.*, 2006) and were 3.2×10^{12} and 3.0×10^{12} photons s^{-1} for MSS-1 and MSS-2, respectively. The corresponding beamsizes (measured using a knife-edge scan) were 7×7 and 9×8 µm, respectively. Absorbed doses were estimated using *RADDOSE-3D* (Zeldin *et al.*, 2013), with dose increments corresponding to the total dose accumulated within the exposure time of the first image, and are detailed in Table S1 in the Supporting information. We note that crystals will be subjected to a small additional absorbed dose during deceleration of the stages prior to the time when the detector starts recording the first diffraction image. While challenging to accurately determine, we estimate that an upper bound for this dose is ∼3 kGy. These experiments were carried out using the same fixed-target chips and translation system as used at SACLA. Details of the datasets are given in Table S1.

2.3. Data analysis

For data measured at SACLA, initial hit finding at the beamline was carried out in *CHEETAH* (Barty *et al.*, 2014). Peak-finding, integration and merging were all performed in *CrystFEL* (White *et al.*, 2016). Data from Diamond beamline I24 were indexed using *DIALS* (version 1.8.5) (Winter *et al.*,

2018) with subsequent scaling and merging performed using *PRIME* (Uervirojnangkoorn *et al.*, 2015). MSS data from beamline I24 consisted of cbf image files numbered sequentially. These were binned into dose points using a simple partitioning script. Multiple lattices were allowed during indexing.

In both cases resolution limits were assessed using $CC_{1/2}$ and R_{split} parameters (White *et al.*, 2013, 2016) together with behaviour in refinement. Structures were solved by molecular replacement using a starting model obtained from a small number of larger DtpAa crystals mounted between two layers of thin film (Axford *et al.*, 2012; Doak *et al.*, 2018) and used to obtain rotation wedges. Water molecules were removed from this model prior to refinement. Structures were refined initially using *REFMAC5* (Murshudov *et al.*, 2011) within the *CCP*4 suite (Winn *et al.*, 2011) and later in *PHENIX* (Adams *et al.*, 2010) and rebuilt between refinement cycles using *Coot* (Emsley *et al.*, 2010). Atoms not well supported by electron density (primarily surface side chains) were deleted from the model. Validation was performed using *MolProbity* (Richardson *et al.*, 2018), *QCCheck* and tools within *Coot* and *PHENIX*. Estimates of bond-length error were calculated from the coordinate diffraction precision index as described (Gurusaran *et al.*, 2014) using the online diffraction precision indicator (DPI) server (Kumar *et al.*, 2015).

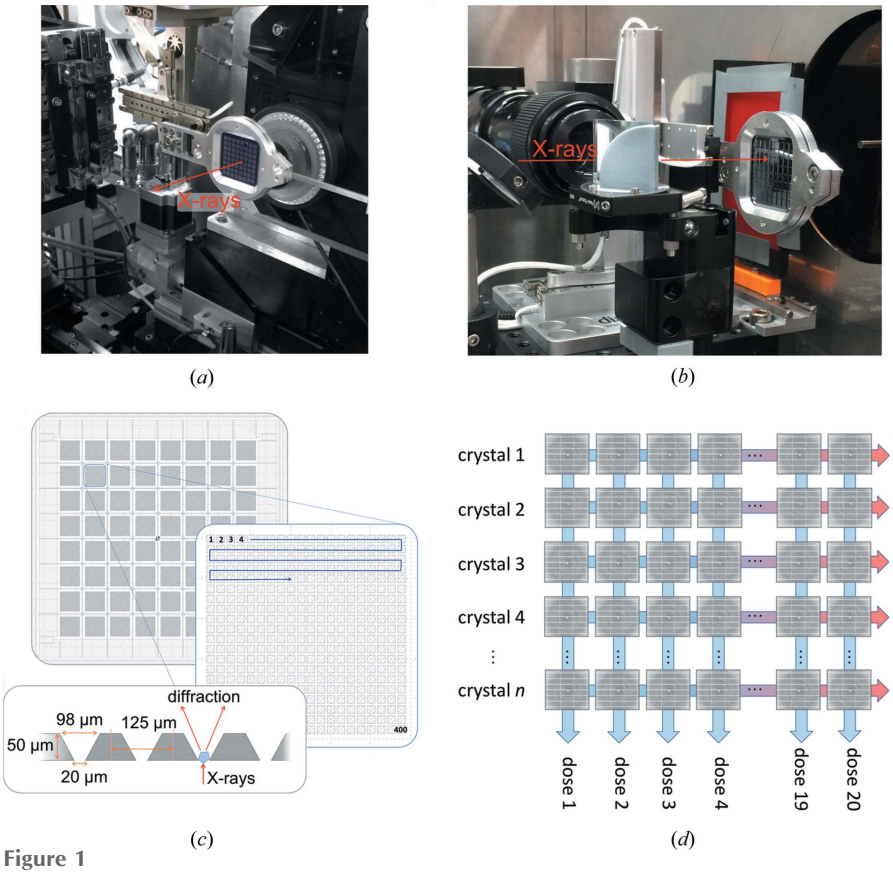


Figure 1
Fixed-target instrumentation in place at (*a*) beamline I24, Diamond Light Source and (*b*) beamline BL2 EH3, SACLA. (*c*) Schematic of fixed target used showing layout of 8 × 8 'city blocks', each comprising 20 × 20 apertures. Shown is a zoomed-in view of a single city block with motion path followed and chip cross-section. (*d*) Formation of dose-resolved datasets by collecting multiple images at each chip aperture. For XFEL data collection, only a single dose point is recorded at each position.

3. Results

3.1. Sample- and time-efficient serial data collection at synchrotron microfocus and XFEL beamlines using silicon nitride fixed-target chips

We used high-capacity silicon nitride fixed targets or 'chips' each containing 25 600 apertures based on those described previously (Oghbaey *et al.*, 2016) to hold the microcrystals used to determine room-temperature serial crystallography structures of DtpAa. Importantly, this sample-delivery system was used in a near-identical manner for both the SFX and the MSS experiments, (Fig. 1), allowing for a direct comparison of the resulting structures. Typically hit rates (we define hit rate as the percentage of frames collected that could be indexed) of ∼30% were achieved on each chip allowing structures to be determined in a highly time- and sample-efficient manner. The volume of microcrystal suspension required per chip was typically 100–200 µl. A schematic of the chip setup and methodological approach is shown in Fig. 1.

3.2. Damage-free DtpAa structure using serial femtosecond crystallography

To produce an 'anchor' structure of DtpAa, *i.e.* resting state ferric, free of any effects of the X-ray beam on the structure, we used the SACLA XFEL (Ishikawa *et al.*, 2012) beamline BL2 EH3 to perform SFX with an X-ray energy of 10 keV, a pulse length of 10 fs with a 1.25 × 1.34 µm beam and a pulse energy of 289 µJ pulse^{-1}. The chip was translated between apertures in the 33 ms separating the 30 Hz XFEL pulses, with a single image recorded at each position. The SFX structure was determined to a resolution of 1.88 Å from a total of 72 615 indexed and merged diffraction patterns [Fig. 2(*a*), Table S1]. The overall structure reveals a ferredoxin-like fold typical of dye decolourizing peroxidases (Sugano, 2009) with two DtpAa monomers in the crystallographic asymmetric unit. The structure was of high quality (Table S1) and refined to an R_{work} and R_{free} of 13.2% and 16.7%, respectively. The refined model exhibited a mean-determined *B* factor of 34.7 Å^2.

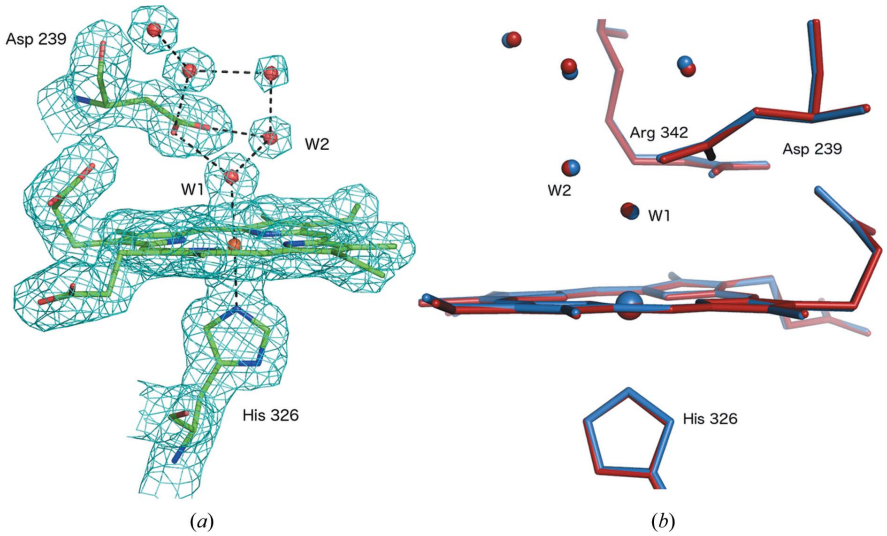


Figure 2
(*a*) $2F_o$–F_c electron-density map contoured at 1σ for the damage-free SFX structure of DtpAa at 1.88 Å resolution, showing the clear and well resolved water network within the heme pocket. Water molecules interact extensively with the pocket residue Asp239 as well as with Arg369 (omitted for clarity). (*b*) Superposition of the SFX structure (blue) with the 32.8 kGy SSX structure (red). Small changes to the heme-pocket water network are apparent even at this low dose.

The heme Fe is six-coordinate with residue His326 acting as the proximal ligand with an Fe—N bond length of 2.19 Å (we note here that monomer B appears to be inactive and is consequently not discussed further). The distal heme coordination site is occupied by a well defined, full occupancy water molecule (W1), bound to the Fe at a distance of 2.40 Å. A number of further, well defined water molecules occupy the remainder of the heme distal pocket [Fig. 2(*a*)]. W1 is hydrogen bonded to a second water, W2, at a distance of 2.68 Å and also interacts with the charged side chains of Asp239 (2.92 Å) and Arg342 (2.74 Å). Interestingly, the side chains of these two amino acids are only 3.13 Å apart (Arg $N_{\eta1}$ to Asp $O_{\delta2}$) suggesting a charge-based interaction.

3.3. DtpAa structures from serial synchrotron crystallography

Serial synchrotron crystallography was carried out at Diamond Light Source beamline I24 at an X-ray energy of 12.8 keV using the same chip and translation system as used for SFX at SACLA. The beam size and flux were measured immediately prior to each experiment, see Materials and methods for details, with approximate values of 7 × 7 µm and 3.1×10^{12} photons s^{-1}. Following each translation of the chip to bring a fresh aperture/crystal into the beam, a series of 10 ms exposures were recorded using the PILATUS3 detector in shutterless mode. This allowed multiple successive snapshots of the same microcrystal within 100 ms. Following exposure of a crystal to the X-ray beam, the chip was translated to the next aperture position and the process repeated [shown schematically in Fig. 1(*d*)]. Using this approach, the total experimental time per fully loaded chip for ten dose points is 45 min, but the total exposure (and hence the absorbed dose) of any individual microcrystal is low and multiple time- (dose-) resolved structures are obtained from a single fixed target. We note here that the 10 ms minimum exposure time was imposed by the maximum frame rate of the current detector available (PILATUS3 6M) and not by limitations arising from the fixed-target movement or synchronization of the target and the X-ray beam. Diffraction images were indexed and integrated independently using *DIALS* (`dials.stills_process`) (Winter *et al.*, 2018) with a simple image-binning procedure used to assign the resulting data to dose bins [Fig. 1(*d*)]. Data within each dose bin were then scaled and merged together using *PRIME* (Uervirojnangkoorn *et al.*, 2015) to form dose-resolved datasets. Using this approach, a complete dataset for each X-ray dose was formed and the corresponding structure refined using the methods described above. The scaling and refinement statistics for each structure are given in Table S1. We first describe an MSS experiment series comprising five dose points with a dose increment of 32.8 kGy (MSS1). An increase in unit-cell volume and trends in scaling statistics clearly indicate the onset of global radiation damage resulting from disorder in the crystalline lattice as dose is accumulated (Fig. S2). The initial resolution was 1.78 Å with only a limited loss of diffracting power/resolution during the 50 ms of total exposure for each microcrystal. Dataset 1 of this series (MSS1-ds1, 32.8 kGy) reveals a six-coordinate heme with a slightly lengthened Fe—O bond at 2.48 Å compared with the SFX structure [Fig. 2(*b*)]. A superposition of MSS1-ds1 with the SFX structure is shown in Fig. 2(*b*). With increasing dose, distinct changes occur around the heme pocket consistent with reduction of the heme iron by X-ray generated solvated photoelectrons (Beitlich *et al.*, 2007; Kekilli *et al.*, 2017). In MSS1-ds2, the Fe—O bond is 2.70 Å with this continuing to lengthen until the last dataset (MSS1-ds5, 164.0 kGy) where it reaches a value of 2.97 Å (Table S3).

In order to provide additional dose points and obtain higher dose SSX structures, a second MSS series was measured with an increased incremental dose value (MSS2). In this case, 10 × 10 ms exposures were measured per crystal position (Table S1 and Fig. S3) with a dose interval of 39.2 kGy. The initial dataset was refined to a resolution of 1.70 Å with the resolution remaining as high as 1.93 Å by dose point 6. After this point (60 ms exposure) the resolution limit decayed, with structure refinement only carried out to 2.18 Å resolution by the last dataset (MSS2-ds8). For comparison, dataset 10 reached only 2.7 Å resolution. In this series the first dataset (MSS2-ds1), associated with a dose of 39.2 kGy, exhibited a Fe–W1 distance of 2.50 Å. This distance increased in successive dose point structures, reaching 2.64 Å in ds3, 2.91 Å in MSS2-ds5 and 3.76 Å in MSS2-ds8 (Fig. 3, Tables S1 and S2).

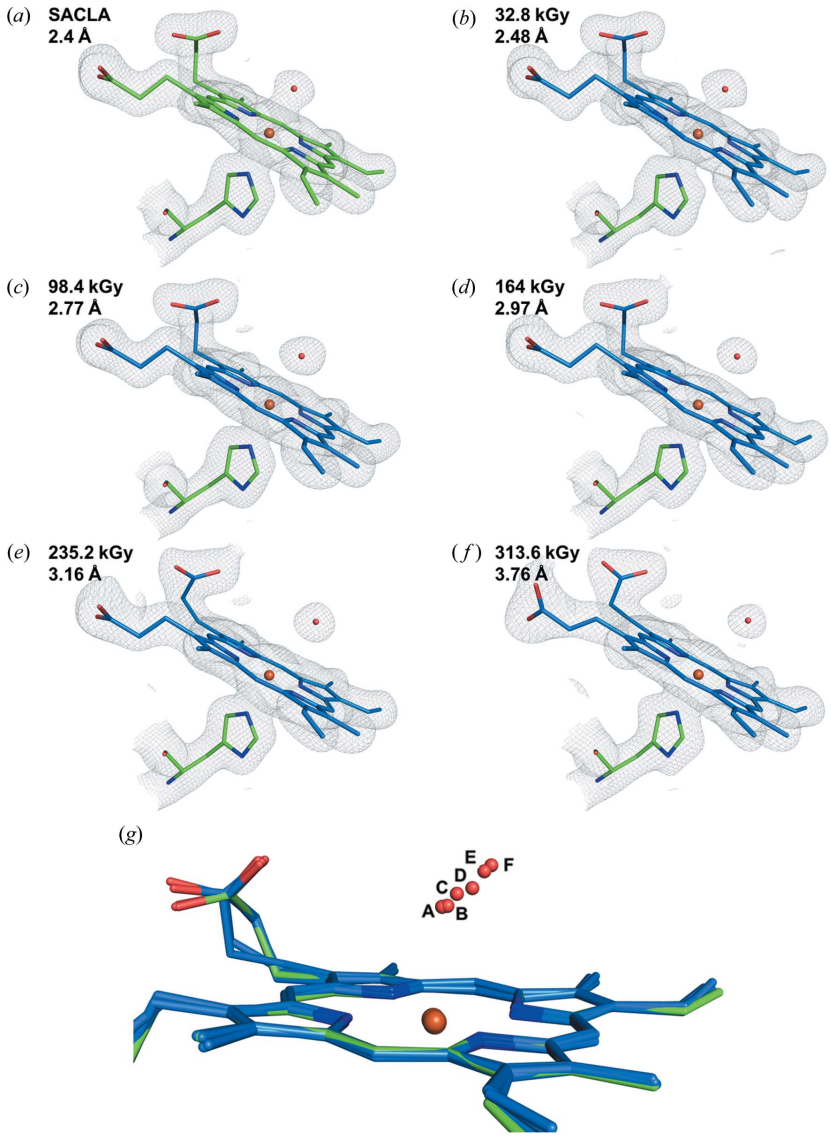


Figure 3
$2F_o$–F_c electron-density maps contoured at 1σ for the heme environment of DtpAa in (*a*) the SFX dataset from SACLA and (*b*–*f*) selected structures from the two MSS series. (*g*) Superposition of selected structures revealing the dose-dependent migration of the water molecule W1 away from the heme Fe. The SFX structure is shown in green with MSS in blue.

Additional structural changes were evident in the heme pocket, with rearrangement of water structures and a flip of a heme propionate as dose was accumulated (Figs. S4 and S5).

The Fe–O distance in all structures from both MSS series is plotted in Fig. 4 and migration of the water away from the heme Fe is shown in Fig. 3(*g*).

It is of considerable interest to compare low-dose synchrotron structures and damage-free XFEL structures determined under near-identical experimental conditions, and to explore if dose-series data may be used to extrapolate back to the 'native' state present prior to X-ray exposure, a so-called 'zero-dose extrapolation'. This approach is analogous to the zero-dose extrapolation of diffraction intensities within conventional single-crystal datasets that has been described previously (Diederichs *et al.*, 2003; Diederichs, 2006; Diederichs & Junk, 2009). In this way, the SFX structure provides a starting point from which synchrotron datasets (inevitably incurring radiation damage and consequent structural change) may be interpreted. A vivid example is shown in a plot of the Fe–W1 distance in the SFX 'anchor' structure and both MSS series, Fig. 4. A near-linear relationship is observed, demonstrating that water migration away from the Fe is dose-dependent under the conditions used. A linear fit to the data yields an intercept (*i.e.* extrapolated to zero dose) of 2.37 Å, which is very close to the value in the SFX structure (2.40 Å) at comparable resolution and within the experimental error for this bond length in the room-temperature structures. The SFX and MSS datasets were deposited in the Protein Data Bank with accession codes as indicated in the supplementary tables in the Supporting information.

4. Discussion

To our knowledge, this is the first reported method for directly comparing dose-resolved serial synchrotron and XFEL structures of radiation-sensitive metalloproteins using the same microcrystal preparations and sample-delivery system. The resolutions achieved with each X-ray source are comparable (1.88 Å SFX and 1.70 Å SSX) allowing the direct comparison of structural features. The SSX data collection allowed sequences of 5–10 MSS dose points to be measured in a time of tens to hundreds of milliseconds per microcrystal. The effective resolution remained high for a substantial proportion of each series. Interestingly, the MSS structures of DtpAa showed well resolved water molecules (Figs. 2 and 3) in all structures, indicating that the progression of reactions within the exposed crystal volume is relatively uniform.

Determining a sequence of dose-dependent structures from the same microcrystals allows subtle and relatively rapidly occurring structural changes to be resolved. In the case of DtpAa, an elongation of the Fe—water bond and eventual bond breakage were observed during tens of milliseconds of exposure to an intense microfocus X-ray beam. By obtaining MSS throughout the process, sufficient data points were recorded in order to be able to fit a function with confidence and allow a zero-dose extrapolation to be made. This provided a close approximation to the structure determined by SFX, providing an alternative approach to obtain a good approx-

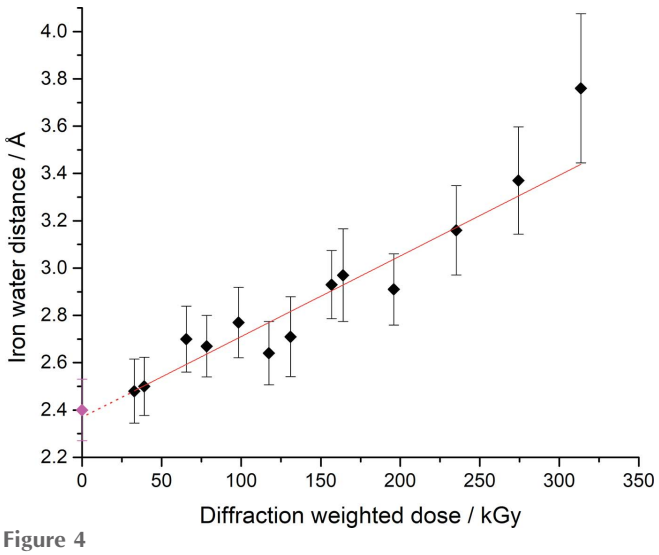


Figure 4
Plot of Fe–W1 distance as a function of X-ray dose from the two measured MSS series. The SFX structure determined using SACLA is plotted as the zero-dose point (magenta). The elongation of the bond length with dose is well fitted by a linear function (red line). The deviation at higher doses is associated with a dissociation of W1 from the immediate vicinity of the heme Fe. The extrapolation to zero dose (dashed red line) gives a value of 2.37 (±0.05) Å which is very close to the 2.40 (±0.13) Å value of this parameter in the SFX structure. Error bars shown are the estimated standard uncertainty in bond length obtained from the DPI value of the Fe and W1 atoms (see Materials and methods).

imation of the 'gold standard' damage-free structure obtained using an XFEL.

4.1. How close can we get to the damage-free enzyme structure using synchrotron radiation?

Despite a relatively low absorbed dose of 32.8 kGy in the MSS1-ds1 dataset, the structure is not identical to that determined by SFX [Fig. 2(*b*)]. Notably, the iron—water bond in MSS1-ds1, the shortest out of all the SSX structures, is elongated compared with the SFX structure. A simple linear fit of the plot of iron—water bond length as a function of dose allowed an extrapolation of the SSX data to zero dose (*y*-axis intercept), yielding a comparable distance to that observed in the SFX structure (Fig. 4).

While not the main focus of this report, the MSS series we present reveal a number of structural states populated during the initial response of DtpAa to X-rays. The elongation of the iron—water bond is consistent with Fe^{III} to Fe^{II} reduction. This reduction is consistent with the generation of solvated electrons by the interaction of X-rays with solvent molecules in the crystal (Kwon *et al.*, 2017; Moody & Raven, 2018). In contrast to the situation at 100 K, where X-ray generated radicals are largely immobilized, room-temperature reactions that involve mass transport may allow such radicals to contribute to the structural changes that our methods allow to be resolved. The relevance of these structures to the function of this class of enzymes will be explored in detail elsewhere.

An additional advantage of our approach is that the same methodology and sample-delivery system is used at synchrotron and XFEL sources/beamlines. This allows for an effective comparison of the structures produced by each X-ray source, allowing the use of comparable crystal sizes, temperatures and sample-delivery methods, factors that might otherwise cause structural heterogeneity.

In summary, we have shown that microcrystals loaded into fixed-target silicon nitride chips can be efficiently employed for data collection at both synchrotron and XFEL sources, allowing near-identical conditions for experiments. Using this technology, we have characterized subtle site-specific changes caused by X-rays in proteins, and directly compared low-dose synchrotron models with, and extrapolation to, damage-free SFX structures. Our method has the potential to be applied to a wide range of enzymes and other proteins especially those that are highly sensitive to radiation damage, including the characterization of electron-driven mechanistic steps in detail through a dose series such as redox reactions in redox metalloenzymes. On a practical level, our approach can be used to extract functionally relevant features of damage-free SFX structures (which require access to scarce beam time at XFELs), reconstructed from extrapolation of MSS determined at multiple low-dose points. Notably, the time interval per MSS structure will be reduced by at least an order of magnitude with upcoming advances in detectors and synchrotron brilliance.

Acknowledgements

XFEL experiments were performed at BL2 EH3 of SACLA with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposal No. 2017B8014). We thank the SACLA beamline staff for technical assistance. We acknowledge the contributions of Drs Minoru Kubo and Takashi Nomura (University of Hyogo) and Takehiko Tosha (RIKEN). The authors thank the beamline staff of BL3 at SACLA. Synchrotron data were measured on beamline I24 at Diamond Light Source under proposal NT14493. Support for travel from the UK XFEL Hub at Diamond Light Source is gratefully acknowledged. We thank Thomas White for helpful discussions regarding the processing of SFX data in *CrystFEL*. R. L. Owen, J. A. R. Worrall, R. W. Strange, D. Axford and M. A. Hough designed the experiment. A. Ebrahim, T. Moreno-Chicano, A. K. Chaplin and J. A. R. Worrall were involved in sample preparation. A. Ebrahim, T. Moreno-Chicano, R. W. Strange, J. A. R. Worrall, M. V. Appleby, J. H. Beale, D. Axford, D. A. Sherrell, H. M. E. Duyvesteyn, H. Sugimoto, K. Tono, S. Owada, R. L. Owen and M. A. Hough participated in data collection. A. Ebrahim, T. Moreno-Chicano, H. M. E. Duyvesteyn, M. V. Appleby, D. Axford, R. L. Owen and M. A. Hough analyzed the data. A. Ebrahim, R. L. Owen and M. A. Hough wrote the manuscript with contributions from all authors.

Funding information

T. Moreno-Chicano was supported by the Leverhulme Trust award RPG-2014-355 to M. A. Hough and R. W. Strange. M. A. Hough acknowledges support from BBSRC award BB/

M022714/1 and J. A. R. Worrall, D. A. Sherrell, D. Axford, R. L. Owen, M. A. Hough and R. W. Strange acknowledge support from a BBSRC Japan–UK International Partnering Award, BB/R021015/1.

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* D**66**, 213–221.

Andersson, R., Safari, C., Dods, R., Nango, E., Tanaka, R., Yamashita, A., Nakane, T., Tono, K., Joti, Y., Båth, P., Dunevall, E., Bosman, R., Nureki, O., Iwata, S., Neutze, R. & Brändén, G. (2017). *Sci. Rep.* **7**, 4518.

Axford, D., Owen, R. L., Aishima, J., Foadi, J., Morgan, A. W., Robinson, J. I., Nettleship, J. E., Owens, R. J., Moraes, I., Fry, E. E., Grimes, J. M., Harlos, K., Kotecha, A., Ren, J., Sutton, G., Walter, T. S., Stuart, D. I. & Evans, G. (2012). *Acta Cryst.* D**68**, 592–600.

Barty, A., Kirian, R. A., Maia, F. R. N. C., Hantke, M., Yoon, C. H., White, T. A. & Chapman, H. (2014). *J. Appl. Cryst.* **47**, 1118–1131.

Beitlich, T., Kühnel, K., Schulze-Briese, C., Shoeman, R. L. & Schlichting, I. (2007). *J. Synchrotron Rad.* **14**, 11–23.

Casadei, C. M., Gumiero, A., Metcalfe, C. L., Murphy, E. J., Basran, J., Concilio, M. G., Teixeira, S. C., Schrader, T. E., Fielding, A. J., Ostermann, A., Blakeley, M. P., Raven, E. L. & Moody, P. C. (2014). *Science*, **345**, 193–197.

Chreifi, G., Baxter, E. L., Doukov, T., Cohen, A. E., McPhillips, S. E., Song, J., Meharenna, Y. T., Soltis, S. M. & Poulos, T. L. (2016). *Proc. Natl Acad. Sci. USA*, **113**, 1226–1231.

Denisov, I. G., Victoria, D. C. & Sligar, S. G. (2007). *Radiat. Phys. Chem.*, **76**, 714–721.

Diederichs, K. (2006). *Acta Cryst.* D**62**, 96–101.

Diederichs, K. & Junk, M. (2009). *J. Appl. Cryst.* **42**, 48–57.

Diederichs, K., McSweeney, S. & Ravelli, R. B. G. (2003). *Acta Cryst.* D**59**, 903–909.

Doak, R. B., Nass Kovacs, G., Gorel, A., Foucar, L., Barends, T. R. M., Grünbein, M. L., Hilpert, M., Kloos, M., Roome, C. M., Shoeman, R. L., Stricker, M., Tono, K., You, D., Ueda, K., Sherrell, D. A., Owen, R. L. & Schlichting, I. (2018). *Acta Cryst.* D**74**, 1000–1007.

Ebrahim, A., Appleby, M. V., Axford, D., Beale, J., Moreno-Chicano, T., Sherrell, D. A., Strange, R. W., Hough, M. A. & Owen, R. L. (2019). *Acta Cryst.* D**75**, 151–159.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Fischer, M., Shoichet, B. K. & Fraser, J. S. (2015). *ChemBioChem*, **16**, 1560–1564.

Garman, E. F. & Weik, M. (2017). *Methods Mol. Biol.* **1607**, 467–489.

Gumiero, A., Metcalfe, C. L., Pearson, A. R., Raven, E. L. & Moody, P. C. (2011). *J. Biol. Chem.* **286**, 1260–1268.

Gurusaran, M., Shankar, M., Nagarajan, R., Helliwell, J. R. & Sekar, K. (2014). *IUCrJ*, **1**, 74–81.

Halsted, T. P., Yamashita, K., Hirata, K., Ago, H., Ueno, G., Tosha, T., Eady, R. R., Antonyuk, S. V., Yamamoto, M. & Hasnain, S. S. (2018). *IUCrJ*, **5**, 22–31.

Holton, J. M. (2009). *J. Synchrotron Rad.* **16**, 133–142.

Horrell, S., Antonyuk, S. V., Eady, R. R., Hasnain, S. S., Hough, M. A. & Strange, R. W. (2016). *IUCrJ*, **3**, 271–281.

Horrell, S., Kekilli, D., Sen, K., Owen, R. L., Dworkowski, F. S. N., Antonyuk, S. V., Keal, T. W., Yong, C. W., Eady, R. R., Hasnain, S. S., Strange, R. W. & Hough, M. A. (2018). *IUCrJ*, **5**, 283–292.

Ishikawa, T., Aoyagi, H., Asaka, T., Asano, Y., Azumi, N., Bizen, T., Ego, H., Fukami, K., Fukui, T., Furukawa, Y., Goto, S., Hanaki, H., Hara, T., Hasegawa, T., Hatsui, T., Higashiya, A., Hirono, T., Hosoda, N., Ishii, M., Inagaki, T., Inubushi, Y., Itoga, T., Joti, Y., Kago, M., Kameshima, T., Kimura, H., Kirihara, Y., Kiyomichi, A., Kobayashi, T., Kondo, C., Kudo, T., Maesaka, H., Marechal, X. M., Masuda, T., Matsubara, S., Matsumoto, T., Matsushita, T., Matsui, S., Nagasono, M., Nariyama, N., Ohashi, H., Ohata, T., Ohshima, T., Ono, S., Otake, Y., Saji, C., Sakurai, T., Sato, T., Sawada, K., Seike, T., Shirasawa, K., Sugimoto, T., Suzuki, S., Takahashi, S., Takebe, H., Takeshita, K., Tamasaku, K., Tanaka, H., Tanaka, R., Tanaka, T., Togashi, T., Togawa, K., Tokuhisa, A., Tomizawa, H., Tono, K., Wu, S. K., Yabashi, M., Yamaga, M., Yamashita, A., Yanagida, K., Zhang, C., Shintake, T., Kitamura, H. & Kumagai, N. (2012). *Nat. Photonics* **6**, 540–544.

Kameshima, T., Ono, S., Kudo, T., Ozaki, K., Kirihara, Y., Kobayashi, K., Inubushi, Y., Yabashi, M., Horigome, T., Holland, A., Holland, K., Burt, D., Murao, H. & Hatsui, T. (2014). *Rev. Sci. Instrum.* **85**, 033110.

Keedy, D. A., Hill, Z. B., Biel, J. T., Kang, E., Rettenmaier, T. J., Brandao-Neto, J., Pearce, N. M., von Delft, F., Wells, J. A. & Fraser, J. S. (2018). *Elife*, **7**, e36307.

Keedy, D. A., Kenner, L. R., Warkentin, M., Woldeyes, R. A., Hopkins, J. B., Thompson, M. C., Brewster, A. S., Van Benschoten, A. H., Baxter, E. L., Uervirojnangkoorn, M., McPhillips, S. E., Song, J., Alonso-Mori, R., Holton, J. M., Weis, W. I., Brunger, A. T., Soltis, S. M., Lemke, H., Gonzalez, A., Sauter, N. K., Cohen, A. E., van den Bedem, H., Thorne, R. E. & Fraser, J. S. (2015). *Elife*, **4**, e07574.

Kekilli, D., Moreno-Chicano, T., Chaplin, A. K., Horrell, S., Dworkowski, F. S. N., Worrall, J. A. R., Strange, R. W. & Hough, M. A. (2017). *IUCrJ*, **4**, 263–270.

Kumar, K. S. D., Gurusaran, M., Satheesh, S. N., Radha, P., Pavithra, S., Thulaa Tharshan, K. P. S., Helliwell, J. R. & Sekar, K. (2015). *J. Appl. Cryst.* **48**, 939–942.

Kwon, H., Smith, O., Raven, E. L. & Moody, P. C. E. (2017). *Acta Cryst.* D**73**, 141–147.

Lomb, L., Barends, T. R. M., Kassemeyer, S., Aquila, A., Epp, S. W., Erk, B., Foucar, L., Hartmann, R., Rudek, B., Rolles, D., Rudenko, A., Shoeman, R. L., Andreasson, J., Bajt, S., Barthelmess, M., Barty, A., Bogan, M. J., Bostedt, C., Bozek, J. D., Caleman, C., Coffee, R., Coppola, N., DePonte, D. P., Doak, R. B., Ekeberg, T., Fleckenstein, H., Fromme, P., Gebhardt, M., Graafsma, H., Gumprecht, L., Hampton, C. Y., Hartmann, A., Hauser, G., Hirsemann, H., Holl, P., Holton, J. M., Hunter, M. S., Kabsch, W., Kimmel, N., Kirian, R. A., Liang, M. N., Maia, F. R. N. C., Meinhart, A., Marchesini, S., Martin, A. V., Nass, K., Reich, C., Schulz, J., Seibert, M. M., Sierra, R., Soltau, H., Spence, J. C. H., Steinbrener, J., Stellato, F., Stern, S., Timneanu, N., Wang, X. Y., Weidenspointner, G., Weierstall, U., White, T. A., Wunderer, C., Chapman, H. N., Ullrich, J., Struder, L. & Schlichting, I. (2011). *Phys. Rev. B*, **84**, 214111.

Moody, P. C. E. & Raven, E. L. (2018). *Acc. Chem. Res.* **51**, 427–435.

Mueller, C., Marx, A., Epp, S. W., Zhong, Y., Kuo, A., Balo, A. R., Soman, J., Schotte, F., Lemke, H. T., Owen, R. L., Pai, E. F., Pearson, A. R., Olson, J. S., Anfinrud, P. A., Ernst, O. P. & Dwayne Miller, R. J. (2015). *Struct. Dyn.* **2**, 054302.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Nass, K., Foucar, L., Barends, T. R. M., Hartmann, E., Botha, S., Shoeman, R. L., Doak, R. B., Alonso-Mori, R., Aquila, A., Bajt, S., Barty, A., Bean, R., Beyerlein, K. R., Bublitz, M., Drachmann, N., Gregersen, J., Jönsson, H. O., Kabsch, W., Kassemeyer, S., Koglin, J. E., Krumrey, M., Mattle, D., Messerschmidt, M., Nissen, P., Reinhard, L., Sitsel, O., Sokaras, D., Williams, G. J., Hau-Riege, S., Timneanu, N., Caleman, C., Chapman, H. N., Boutet, S. & Schlichting, I. (2015). *J. Synchrotron Rad.* **22**, 225–238.

Oghbaey, S., Sarracini, A., Ginn, H. M., Pare-Labrosse, O., Kuo, A., Marx, A., Epp, S. W., Sherrell, D. A., Eger, B. T., Zhong, Y., Loch, R., Mariani, V., Alonso-Mori, R., Nelson, S., Lemke, H. T., Owen, R. L., Pearson, A. R., Stuart, D. I., Ernst, O. P., Mueller-Werkmeister, H. M. & Miller, R. J. D. (2016). *Acta Cryst.* D**72**, 944–955.

Owen, R. L., Rudiño-Piñera, E. & Garman, E. F. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 4912–4917.
Richardson, J. S., Williams, C. J., Hintze, B. J., Chen, V. B., Prisant, M. G., Videau, L. L. & Richardson, D. C. (2018). *Acta Cryst.* D**74**, 132–142.
Schlichting, I. (2015). *IUCrJ*, **2**, 246–255.
Sugano, Y. (2009). *Cell. Mol. Life Sci.* **66**, 1387–1403.
Uervirojnangkoorn, M., Zeldin, O. B., Lyubimov, A. Y., Hattne, J., Brewster, A. S., Sauter, N. K., Brunger, A. T. & Weis, W. I. (2015). *Elife*, **4**, e05421
Waldron, K. J., Rutherford, J. C., Ford, D. & Robinson, N. J. (2009). *Nature*, **460**, 823–830.
Weik, M. & Colletier, J.-P. (2010). *Acta Cryst.* D**66**, 437–446.
White, T. A., Barty, A., Stellato, F., Holton, J. M., Kirian, R. A., Zatsepin, N. A. & Chapman, H. N. (2013). *Acta Cryst.* D**69**, 1231–1240.
White, T. A., Mariani, V., Brehm, W., Yefanov, O., Barty, A., Beyerlein, K. R., Chervinskii, F., Galli, L., Gati, C., Nakane, T., Tolstikova, A., Yamashita, K., Yoon, C. H., Diederichs, K. & Chapman, H. N. (2016). *J. Appl. Cryst.* **49**, 680–689.
Winn *et al.* (2011). *Acta Cryst.* D**67**, 235–242
Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K. & Evans, G. (2018). *Acta Cryst.* D**74**, 85–97.
Zeldin, O. B., Gerstel, M. & Garman, E. F. (2013). *J. Synchrotron Rad.* **20**, 49–57.

IUCrJ
ISSN 2052-2525
BIOLOGY | MEDICINE

High-throughput structures of protein–ligand complexes at room temperature using serial femtosecond crystallography

Tadeo Moreno-Chicano,[a]‡ Ali Ebrahim,[a,b] Danny Axford,[b] Martin V. Appleby,[b] John H. Beale,[b] Amanda K. Chaplin,[a]§ Helen M. E. Duyvesteyn,[b,c] Reza A. Ghiladi,[d] Shigeki Owada,[e,f] Darren A. Sherrell,[b] Richard W. Strange,[a] Hiroshi Sugimoto,[e] Kensuke Tono,[e,f] Jonathan A. R. Worrall,[a] Robin L. Owen[b]* and Michael A. Hough[a]*

[a]School of Life Sciences, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, England, [b]Diamond Light Source, Harwell Science and Innovation Campus, Didcot OX11 0DE, England, [c]Division of Structural Biology (STRUBI), University of Oxford, The Henry Wellcome Building for Genomic Medicine, Roosevelt Drive, Oxford OX3 7BN, England, [d]Department of Chemistry, North Carolina State University, Raleigh, NC 27695-8204, USA, [e]RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan, and [f]Japan Synchrotron Radiation Research Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5198, Japan. *Correspondence e-mail: robin.owen@diamond.ac.uk, mahough@essex.ac.uk

Received 8 May 2019
Accepted 21 August 2019

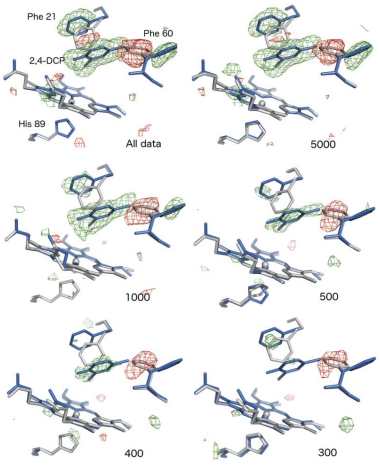
Edited by K. Moffat, University of Chicago, USA

‡ Current address: Institut de Biologie Structurale, 71 Avenue des Martyrs, 38000 Grenoble, France.
§ Current address: Department of Biochemistry, Sanger Building, 80 Tennis Court Road, Cambridge CB2 1GA, England.

Keywords: serial femtosecond crystallography; ligand binding; high throughput; X-ray crystallography; damage-free structures; X-ray free-electron lasers.

PDB references: dehaloperoxidase B, complex with 2,4-dichlorophenol, 6i7f; complex with 5-bromoindole, 6i6g; dye-type peroxidase Aa, complex with imidazole, 6i7c; copper nitrite reductase, complex with nitrite, 6qwg

Supporting information: this article has supporting information at www.iucrj.org



OPEN ACCESS

High-throughput X-ray crystal structures of protein–ligand complexes are critical to pharmaceutical drug development. However, cryocooling of crystals and X-ray radiation damage may distort the observed ligand binding. Serial femtosecond crystallography (SFX) using X-ray free-electron lasers (XFELs) can produce radiation-damage-free room-temperature structures. Ligand-binding studies using SFX have received only modest attention, partly owing to limited beamtime availability and the large quantity of sample that is required per structure determination. Here, a high-throughput approach to determine room-temperature damage-free structures with excellent sample and time efficiency is demonstrated, allowing complexes to be characterized rapidly and without prohibitive sample requirements. This yields high-quality difference density maps allowing unambiguous ligand placement. Crucially, it is demonstrated that ligands similar in size or smaller than those used in fragment-based drug design may be clearly identified in data sets obtained from <1000 diffraction images. This efficiency in both sample and XFEL beamtime opens the door to true high-throughput screening of protein–ligand complexes using SFX.

1. Introduction

The accurate determination of the structures of protein–ligand complexes is essential for drug discovery, enzymology and biotechnology. Developments in the automation of protein crystallization, ligand soaking, harvesting, structure determination, ligand modelling and structural refinement have allowed the high-throughput screening of soaked crystals at synchrotron X-ray beamlines (Collins *et al.*, 2018; Pearce, Krojer, Bradley *et al.*, 2017; Pearce, Krojer & von Delft, 2017). For important classes of proteins, the binding of ligands may be affected by X-ray-driven changes either in the oxidation state of redox centres within the protein or to amino-acid side chains involved in protein–ligand interactions. In these cases, there is a premium on structure determination using low-dose methods. Prime examples of this are heme enzymes, where the iron centre in resting iron(III) and high-valent iron(IV) states is exquisitely prone to reduction by solvated photoelectrons generated by the interaction of synchrotron X-rays with solvent in the crystal (see, for example, Beitlich *et al.*, 2007; Kekilli *et al.*, 2017). Heme enzymes, such as the cytochrome

P450s, are involved in the metabolism/breakdown of approximately 90% of small-molecule drugs, and are more generally themselves drug targets in yeast, fungi and tuberculosis infections (McLean & Munro, 2017; Guengerich *et al.*, 2016; Rendic & Guengerich, 2015). Importantly, the determination of protein–ligand complexes at room temperature is likely to better reflect *in vivo* conditions than crystals cryogenically cooled to 100 K (for an interesting example, see Fischer *et al.*, 2015). Indeed, significant differences in binding have been observed at room temperature (RT) compared with 100 K (Keedy *et al.*, 2018). Furthermore, ligand soaking into microcrystals (1–20 µm) has the theoretical potential to be more effective than soaking into larger crystals (>50 µm) (McPherson, 2019). The distance that a ligand needs to penetrate into the crystal to reach its centre is proportionately shorter for smaller crystals, potentially leading to higher occupancy rates.

These issues in combination place a high value on protein–ligand complexes determined from microcrystals at RT that are free of observable effects of radiation damage. The only current approach that can deliver this is serial femtosecond crystallography (SFX) at X-ray free-electron lasers (XFELs; Schlichting, 2015) using short (<20 fs) X-ray pulses (Inoue *et al.*, 2016; Lomb *et al.*, 2011; Nass *et al.*, 2015; Nass, 2019). Ligand-binding studies using SFX have received little attention, largely owing to the scarcity of beamtime and high sample requirements in typical sample-delivery systems such as the gas dynamic virtual nozzle (GDVN) injectors (Schlichting, 2015). The drive to obtain damage-free, RT structures is balanced against the strong practical driver to minimize sample consumption per obtained structure, and the desire to collect data from multiple candidate ligands in a short time period.

A limited number of studies have sought to address the challenge of obtaining damage-free, RT crystal structures of protein–ligand complexes in a manner that is efficient both in sample and in data-collection time. An early study examined ligand binding to a P-type ATPase membrane protein in microcrystals delivered to the beam using a liquid microjet injector (Bublitz *et al.*, 2015). This work demonstrated the applicability of SFX to ligand-binding studies, showing that ligands could be clearly resolved even if the high-resolution data collected are weak and statistically poor. A more recent study (Naitow *et al.*, 2017) explored the feasibility of SFX ligand-binding studies using microcrystals of the model system thermolysin delivered by a high-viscosity water- or oil-based injector. The small-molecule ligand was readily resolved in electron-density maps, with clear differences in binding modes observed between the room-temperature SFX and 100 K synchrotron-radiation (SR) structures.

Here, we describe the rapid determination of protein–ligand complexes at RT. Microcrystals were mounted in silicon fixed targets or 'chips' at the SPring-8 Ångstrom Free Electron Laser (SACLA), Hyogo, Japan (Ishikawa *et al.*, 2012). The fixed-target sample-delivery approach minimizes sample consumption, provides high hit rates and allows multiple high-quality data sets to be measured in a very short time, an important advantage given the limited availability of XFEL beamtime. The chip system also allows rapid switching experiments in which crystals of different targets are soaked with different ligands. Moreover, the short time between soaking, chip loading and the completion of data collection reduces the need for long-term protein–ligand crystal stability that is required for a typical injector experiment. This also ensures that crystals are exposed to the soaked ligand for a similar length of time.

We have applied this approach to crystals of two heme peroxidase enzymes: a multifunctional dehaloperoxidase from the marine annelid *Amphitrite ornata* (DHP-B; Barrios *et al.*, 2014; Franzen *et al.*, 2012; McCombs, Moreno-Chicano, *et al.*, 2017; McCombs, Smirnova *et al.*, 2017) and a dye-decolourizing peroxidase (Sugano, 2009) of industrial relevance (Colpa *et al.*, 2014) from *Streptomyces lividans* (DtpAa). We also examine the challenging case of detecting nitrite binding to copper nitrite reductase from *Achromobacter cycloclastes* (*Ac*NiR; Horrell *et al.*, 2017), where the ligand displaces a water molecule bound in the active site. The enzyme and crystal systems used are of cubic (high), orthorhombic (medium) and monoclinic (low) symmetry space groups, as well as exhibiting full to partial ligand occupancies within the same crystallographic asymmetric units. The complexes investigated include ligands directly binding to the heme, together with those occupying a binding pocket but not bound to the iron, with ligand sizes of 3–10 non-H atoms (Supplementary Fig. S1). We note that the typical molecular weight of the fragments used in fragment-based drug design is approximately 150–250 Da, with a typical size of 200 Da (Price *et al.*, 2017).

We explore the potential of this approach for rapid SFX screening of ligands/drug candidates, examining the minimum number of merged diffraction patterns required to reliably detect ligand binding and the future potential of this approach at current and planned XFEL beamlines. We assess several metrics for ligand fit to electron density with the data sets presented in the light of the recent debate around ligand validation (Smart *et al.*, 2018). Remarkably, data sets comprising of <1000 merged diffraction patterns allowed clear and unambiguous identification of ligand-binding modes, despite extremely poor merging and refinement statistics. The number of crystals required for complete data is lowered by the bandwidth of the XFEL beam. Our work thus demonstrates that high-throughput screening is eminently practicable using SFX, with modest requirements for sample quantity and experimental time.

2. Materials and methods

2.1. Protein production and crystallization

Dye-type peroxidase Aa (DtpAa) from *S. lividans* was expressed and purified as described previously (Ebrahim, Moreno-Chicano *et al.*, 2019). Crystals were grown in batch using a modification of the crystallization conditions used for growing large single crystals, consisting of 25%(*w*/*v*) PEG

1500 and 100 m*M* MIB buffer (Hampton Research, comprising MES, boric acid and imidazole pH 8). The final protein concentration in the batches ranged from 6.5 to 2.1 mg ml^{-1}. Crystals grew in 1–2 days to approximate dimensions of 20–30 µm and were transported to SACLA at ambient temperature (the crystals were transported in hand luggage without cooling). Dehaloperoxidase B (DHP) from *A. ornata* was expressed and purified as described previously (McCombs, Moreno-Chicano *et al.*, 2017). Batch microcrystallization was used, mixing 30 mg ml^{-1} DHP in 20 m*M* MES pH 6.0 with 40%(*w*/*v*) PEG 4000, 200 m*M* ammonium sulfate in a 1:4 ratio in a total volume of 250–500 µl. DHP microcrystals grew in 3–5 days at 4°C to typical dimensions of 20–30 µm and were transported to SACLA at 4°C. 5-Bromoindole (5BR) and 2,4-dichlorophenol (DCP) (Sigma) were each dissolved in 100% DMSO and 20 µl of the resulting solution was added to a 200 µl crystal suspension to yield final ligand concentrations of 5 m*M* DCP and 50 m*M* 5BR. Microcrystals were soaked in batches for 3–5 min immediately prior to loading onto the silicon chip. *Ac*NiR microcrystals were grown as described previously (Ebrahim, Appleby *et al.*, 2019) and were soaked in 100 m*M* potassium nitrite for approximately 20 min prior to loading onto the chip.

2.2. Data collection and processing

Microcrystals were loaded into fixed-target chips as described previously (Ebrahim, Appleby *et al.*, 2019; Oghbaey *et al.*, 2016). The chips were fabricated commercially (Southampton Nanofabrication Centre; https://www.southampton-nanofab.com) using a method based on that described previously (Oghbaey *et al.*, 2016). Typically, 100–200 µl of microcrystal suspension was loaded onto a chip containing 25 600 apertures and excess liquid was removed using a weak vacuum applied to the underside of the chip surface. For DHP microcrystals around 1.5 mg of protein was loaded in each chip, requiring around 4.5 mg for a complete data set (three or four chips), while *Ac*NiR microcrystals were loaded in quantities of around 2 mg for a complete data set (two chips at 1 mg per chip). In the case of DtpAa even less protein was needed: only 0.45–6.0 mg per chip and around 1.80 mg for a complete data set. SFX data were measured on SACLA (Ishikawa *et al.*, 2012) beamline BL2 EH3 with a photon energy of 10.0 keV, a repeat rate of 30 Hz and a pulse length of 10 fs. The beam, with a 1.25 × 1.34 µm spot size (FWHM) and a pulse energy of 289 µJ per pulse (pre-attenuation), was attenuated to 13% of full flux to minimize detector overloads. The SACLA beam was in SASE mode, with FWHM bandwidth ∼70 eV. The fixed-target chip was translated between X-ray pulses such that each crystal position was exposed only once, and the measurement of all 25 600 positions on a chip took 14 min. The hit rate during data collection was monitored using *Cheetah* (Barty *et al.*, 2014), while peak finding, indexing and merging of data were performed using *CrystFEL* v.0.6.4 (White *et al.*, 2016). Structures were refined using starting models of ligand-free structures from which water and other solvent molecules had been removed. Refinement was initially carried out in *REFMAC*5 (Murshudov *et al.*, 2011) within the *CCP*4 suite and completed in *PHENIX* (Adams *et al.*, 2010). All structures were validated using *MolProbity* (Williams *et al.*, 2018), the *JCSG Quality Control Check* server and tools within *PHENIX* (Adams *et al.*, 2010) and *Coot* (Emsley *et al.*, 2010).

To explore the limits of ligand identification in SFX data sets, randomly selected images from the indexed data (`*.stream` files from *CrystFEL*) formed data subsets with defined, variable numbers of images. These were scaled and merged in the same manner as the data sets containing all images and were used in refinement versus the model for the appropriate complex determined using all data, from which the ligand had been removed. OMIT maps were generated using torsion-based simulated-annealing refinement in *phenix.refine* (Adams *et al.*, 2010) in order to minimize model bias. As an additional validation step, selected subsets were refined against the structure of the native enzymes (where the ligands were not present) using the procedure described above.

2.3. Ligand modelling

Ligands were initially modelled into the all-image data sets based on the $mF_o - DF_c$ difference electron-density maps. In all cases, ligand density was unambiguous and ligands were modelled with near-full occupancy in one of the two subunits of the homodimeric enzymes (for DHP and DtpAa) or in the single subunit of *Ac*NiR in the crystallographic asymmetric unit. The ligands were straightforwardly located in an automated manner using the 'Find Ligands' feature of *Coot*. The second monomer in the DHP asymmetric unit contained a lower occupancy ligand (5BR) or very weak ligand density (DCP), while in DtpAa the second monomer did not show a bound exogenous ligand in the active site. Restraints for nonstandard ligands were produced using *ACEDRG* (Long *et al.*, 2017). For the data subsets, the data were refined by two parallel approaches to avoid model bias. Firstly, the data were refined against the 'all-images' structures, from which the ligands had been removed, using simulated annealing in *phenix.refine* to remove bias. As an additional test that bias was not present, selected structures were refined against the native, ligand-free structures of the enzymes and simulated-annealing (SA) OMIT difference maps were generated. The known position of the ligand from the 'all-images' models was then compared with the difference density map generated from that subset. The quality of the fit of modelled ligands to the electron-density maps was determined using *EDIAscorer* (Meyder *et al.*, 2017). The 'Find Ligands' feature of *Coot* was also used for each subset, in this case searching the $mF_o - DF_c$ SA OMIT map for suitable hits.

$F_o - F_o$ isomorphous difference maps between the DHP–5BR and DHP–DCP data sets were generated in *PHENIX* with the native ligand-free DHP structure (see below) used to phase the data sets (although near-identical results were generated if either of the above ligand-bound structures were used for phasing).

Table 1
Data-collection, processing and refinement statistics for full SFX data sets for enzyme–ligand complexes.

Values in parentheses are for the outermost resolution shell.

Structure	DHP–DCP	DHP–5BR	DtpAa–imidazole	*Ac*NiR–nitrite
Data collection and processing				
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_1$	$P2_13$
Unit-cell parameters (Å, °)	a = 60.9, b = 67.2, c = 68.7, $\alpha = \beta = \gamma$ = 90	a = 61.0, b = 67.3, c = 68.8, $\alpha = \beta = \gamma$ = 90	a = 72.5, b = 68.0, c = 73.5, $\alpha = \gamma$ = 90, β = 105.6	a = 97.6, b = 97.6, c = 97.6, $\alpha = \beta = \gamma$ = 90
Chips used	3	4	4	2
Images collected	76800	102800	102800	51200
Indexed images merged	32618	53662	20316	16586
Unique reflections	24749	24840	56220	24729
Resolution (Å)	37.7–1.85 (1.90–1.85)	45.6–1.85 (1.90–1.85)	70.8–1.88 (1.93–1.88)	43.7–1.90 (1.93–1.90)
Completeness (%)	100 (100)	100 (100)	100 (100)	100 (100)
Multiplicity	579 (340)	907.7 (524.0)	101.6 (64.2)	3281.4 (2299.1)
$CC_{1/2}$	0.99 (0.66)	1.00 (0.65)	0.96 (0.60)	0.99 (0.63)
R_{split} (%)	6.6 (61.9)	5.5 (66.6)	15.8 (63.9)	9.73 (58.61)
Refinement				
Resolution range (Å)	34.4–1.85	45.6–1.85	35.3–1.88	43.7–1.90
R_{work} (%)	16.8	16.7	13.9	13.7
R_{free} (%)	19.9	18.9	17.6	17.2
R.m.s.d., bond lengths (Å)	0.010	0.005	0.010	0.006
R.m.s.d., bond angles (°)	1.23	0.96	0.87	0.90
Ramachandran most favoured (%)	98.2	98.9	98.5	98.8
PDB code	6i7f	6i6g	6i7c	6qwg

3. Results

3.1. Determination of protein–ligand complex structures by SFX in a time- and sample-efficient manner

SFX structures for each enzyme–ligand complex were determined from data measured from either two (*Ac*NiR), three (DHP–DCP) or four (DHP–5BR and DtpAa–imidazole) chips. This took approximately 14 min of data collection and ∼16 min of beamtime per chip (sample-change, hutch-search and alignment time are included). In each case, structure solution was by molecular replacement and the resolution and data quality were sufficient to clearly define essentially all main-chain and most side-chain atoms together with well defined networks of water molecules. The quality of the data sets and structures is given in Table 1. For each structure, clear positive difference density was evident for the ligands, which were unambiguously located. The chemical structures of the ligands used in this study are shown in Supplementary Fig. S1.

3.2. SFX structures of ligand-bound complexes

In each DHP structure, clear $F_o - F_c$ electron density was apparent in the heme pocket consistent with a high-occupancy bound ligand in one monomer of the dimer and a second lower occupancy binding site in the other. This difference in occupancy is consistent with previous single-crystal structures of DHP complexes with a range of different ligands in this space group (see, for example, McCombs, Moreno-Chicano *et al.*, 2017). DCP exhibited a binding site that was virtually identical to those previously observed for the guaiacol substrates 4-bromoguaiacol (PDB entry 6cke), 4-nitroguaiacol (PDB entry 6ch5) and 4-methoxyguaiacol (PDB entry 6ch6) (McGuire *et al.*, 2018), while the 5BR complex was consistent with a computationally hypothesized binding site (Barrios *et al.*, 2014), with both results together demonstrating that SFX provides accurate substrate-binding orientations.

The details of the binding modes themselves are beyond the scope of this manuscript and will be described in detail in a separate publication. Strong electron-density peaks were present for the two Cl atoms of DCP and the single Br atom of 5BR, allowing the ligand orientation to be easily confirmed, although it is important to note that the electron density was well defined for all atoms of the ligands. For both DHP–ligand structures one monomer had near-full occupancy, but significantly lower occupancy (as refined in *PHENIX*; Adams *et al.*, 2010) was observed in the second monomer of the homodimer [Figs. 1(*a*) and 1(*b*)]. This feature allowed us to examine the effect of ligand occupancy on ligand detectability in maps derived from SFX data (see below).

The SFX structure of DtpAa was determined in space group $P2_1$ to 1.88 Å resolution (Table 1). The overall structure of the enzyme homodimer was highly similar to that of ferric DtpAa crystallized in a condition that did not contain imidazole (Ebrahim, Moreno-Chicano *et al.*, 2019). The examination of $F_o - F_c$ difference maps indicated that an imidazole ligand was coordinated *via* an N atom to the distal position of the heme iron in one monomer of the DtpAa dimer with full occupancy. The Fe—N (imidazole) bond was 2.2 Å, while imidazole also formed two hydrogen bonds (2.7 and 2.9 Å) to Asp239 [Fig. 1(*c*)], and the heme pocket also contains several well ordered water molecules. Comparison with the ligand-free ferric DtpAa structure also obtained by SFX (Ebrahim, Moreno-Chicano *et al.*, 2019) revealed that the imidazole displaces the distal water molecule from the heme and induces a number of modest structural rearrangements in the heme pocket (Supplementary Fig. S2). A second imidazole molecule is bound to the protein away from the heme pocket, forming a 2.7 Å bond to Thr351 and interacting via a bridging water with Glu283. In contrast, for monomer *A* no imidazole ligand was observed in the distal heme pocket and instead a water molecule is bound at a distance of 2.4 Å in a similar manner to

that in the ferric DtpAa structure (Ebrahim, Moreno-Chicano *et al.*, 2019).

The structure of *Ac*NiR in complex with nitrite was determined to 1.90 Å resolution (Table 1). The type 2 Cu site, which is the site of ligand binding, displayed clear electron density for a bound nitrite molecule with a bidentate O-binding geometry as previously described in multiple 100 K and room-temperature structures obtained from single crystals (Meyder *et al.*, 2017; Horrell *et al.*, 2016) [Fig. 1(*d*)]. The positions of the ligand-binding sites within the protein fold for each complex are shown in Supplementary Fig. S3.

3.3. Ligand density features as a function of the number of diffraction patterns in a data set

As described above, all three ligands were unambiguously identified in maps derived from the full data sets, demonstrating that ample diffraction patterns had been included in the merged data sets, which had good data-quality metrics (Table 1). To test the lower limits of the number of diffraction patterns that would allow us to identify bound ligands in high-throughput SFX experiments, the data were partitioned into subsets of decreasing size to produce independent merged data sets containing progressively fewer diffraction patterns, (Supplementary Tables S1–S4). OMIT difference maps were generated by simulated-annealing refinement in *PHENIX* (Adams *et al.*, 2010) using the all-data structure with ligand atoms omitted in the initial model. As expected, the merging and refinement statistics, and consequently the resolution cut-off, progressively deteriorated as the number of merged patterns was reduced (Fig. 2, Supplementary Tables S1–S4, Supplementary Fig. S4). The OMIT map quality was, as might be reasonably expected, proportional to the number of images in the data set. As a simple practical test to emulate a typical crystallographic workflow, the 'Find Ligands' feature of *Coot* (Debreczeni & Emsley, 2012; Emsley, 2017) was used to test whether each ligand could be correctly fitted into the simulated-annealing $F_o - F_c$ map without bias from the experimentalist's prior knowledge of the correct pose.

3.4. 2,4-Dichlorophenol–DHP complex

We first examined the effect of the number of crystals included in a data set on the resulting electron-density maps for the complex between DHP and DCP. Very clear $F_o - F_c$ simulated-annealing OMIT map features for the ligand were evident in subsets considerably smaller than the 'full' data sets. For example, a subset of 5000 crystals showed merging statistics that would still be considered acceptable by standard assessments [R_{split} = 0.17 (0.73), $CC_{1/2}$ = 0.95 (0.56) to 1.95 Å

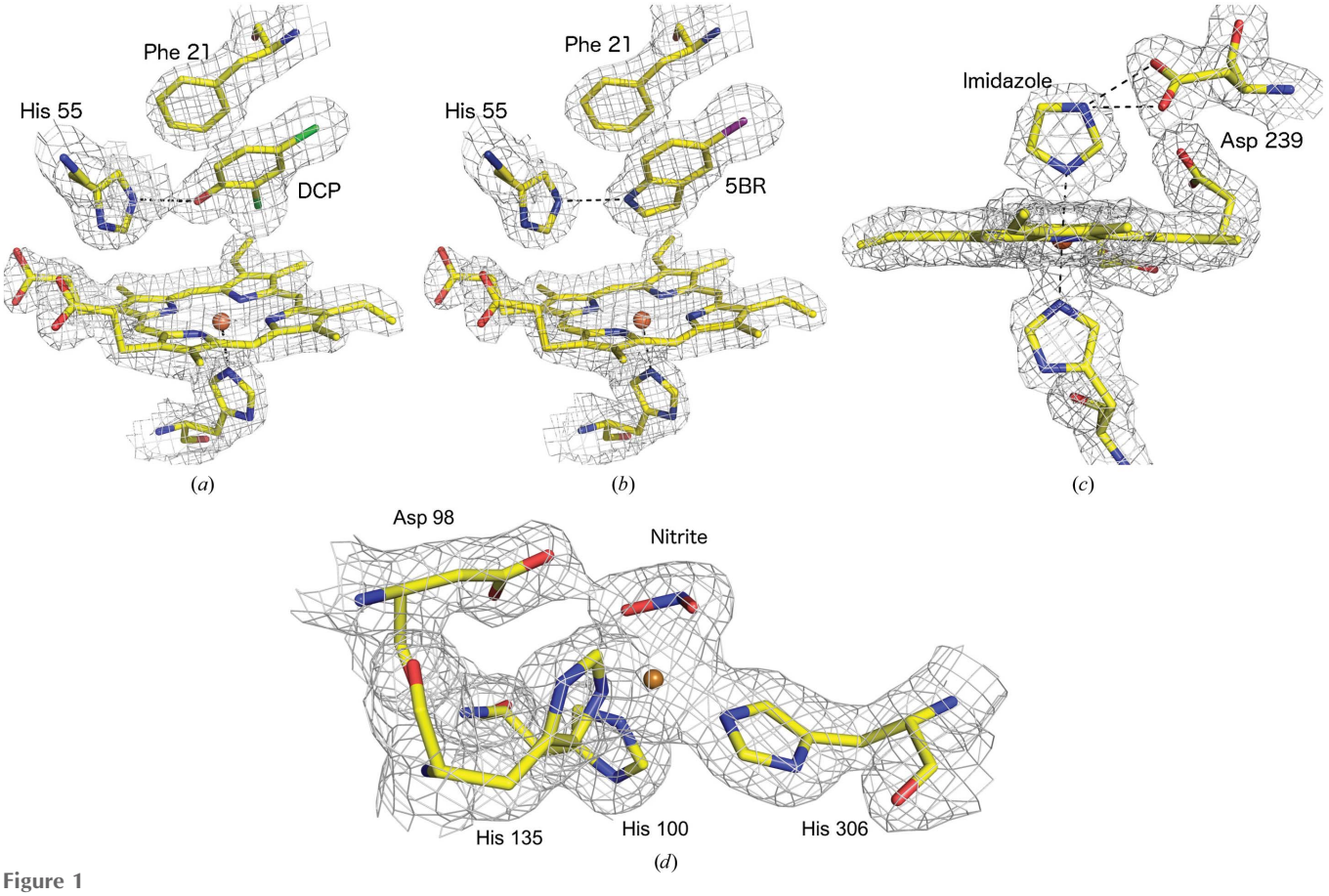


Figure 1
$2F_o - F_c$ electron-density maps, contoured at 1σ, showing the complexes of DHP with (*a*) DCP with the Cl atoms shown in green and (*b*) 5BR with the Br atom shown in purple, (*c*) the complex of DtpAa with imidazole and (*d*) the complex of *Ac*NiR with nitrite. In each case, the active site of the monomer with the highest ligand occupancy is shown. The maps in (*a*)–(*d*) were generated using the all-image data sets.

resolution] and unsurprisingly ligand finding was straightforward. When the data set was reduced to containing only 1000 crystals the merging statistics were poor [R_{split} = 0.39 (0.65), $CC_{1/2}$ = 0.72 (0.57) to 2.2 Å resolution], and with 500 images these metrics indicated very poor or even meaningless data quality [R_{split} = 0.56 (0.92), $CC_{1/2}$ = 0.57 (0.42) to 2.2 Å resolution]. The refinement statistics also deteriorated with decreasing data-set size (Supplementary Table S1).

Remarkably, data sets comprising far fewer than 1000 indexed patterns displayed very clear features in simulated-annealing OMIT maps of the distal pocket, covering all atoms of the best-ordered DCP ligand (in monomer *B*). Examples are shown in Supplementary Fig. S5, where the $F_o - F_c$ OMIT map allowed all atoms of the ligand to be unambiguously modelled, even when the merging statistics were very poor and refinement *R* factors were high (Supplementary Table S1). Because of the poor merging statistics with <1000 images, it was not possible to use these metrics to assess the resolution limit in merging for these data; however, refinement using the same resolution limit as the 1000-image set still allowed straightforward ligand placement.

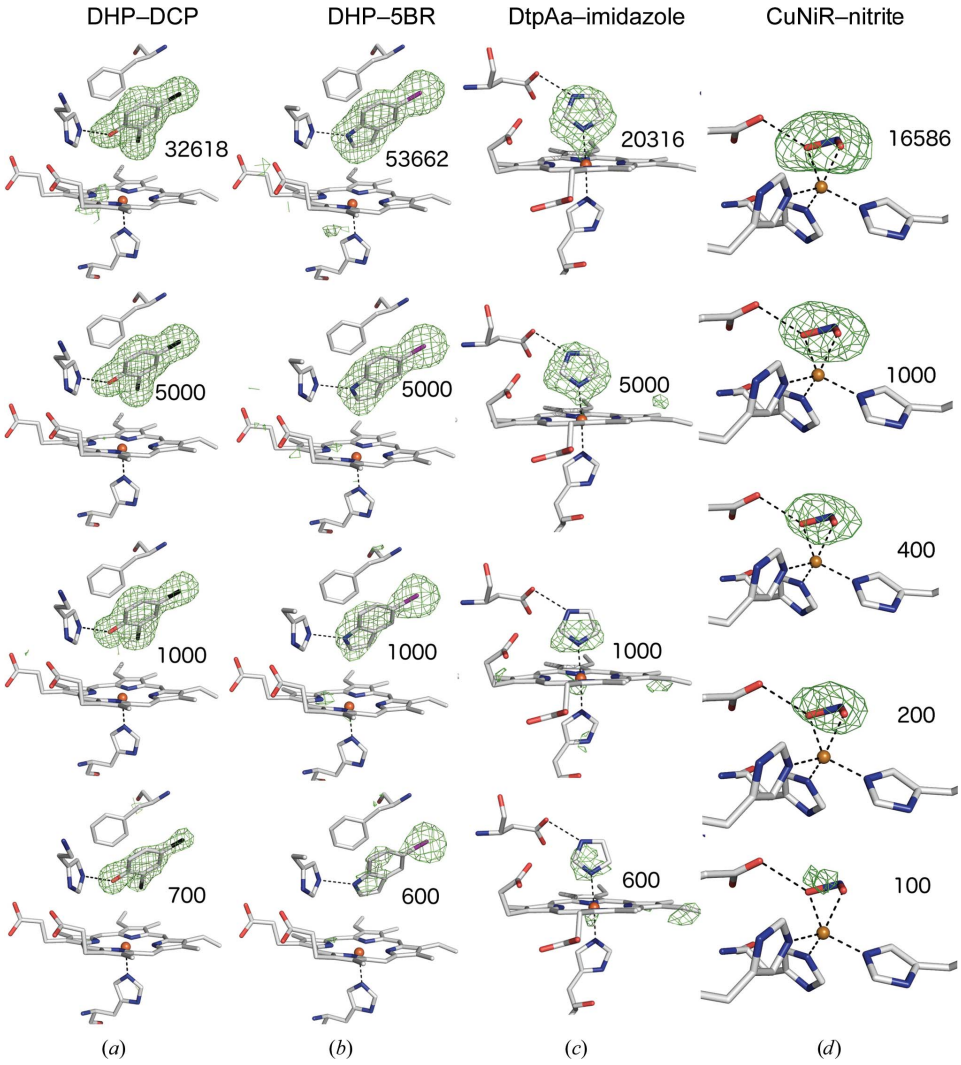


Figure 2
$F_o - F_c$ simulated-annealing OMIT maps contoured at 3σ for the heme region from selected data subsets for (*a*) DHP–DCP, (*b*) DHP–5BR, (*c*) DtpAa–imidazole and (*d*) *Ac*NiR–nitrite, each superposed on the refined structure from all data. For (*a*) and (*b*) the highest occupancy ligand monomer of the homodimer is shown. Additional subsets are shown in Supplementary Figs. S5, S6 and S8–S11.

For data sets produced from <400 crystals, difference map quality rapidly deteriorated (Supplementary Fig. S5). This deterioration of the maps appears to approximately coincide with a loss of data completeness and redundancy in these data sets. The lower occupancy ligand present in the second monomer failed to be located in data-subset OMIT maps of decreasing size more rapidly than was the case for the fully occupied ligand (Supplementary Fig. S6). *EDIAscorer* (Meyder *et al.*, 2017) electron-density analysis is shown in Fig. 3 and Supplementary Fig. S7, showing the excellent quality of the difference map for all ligand atoms down to very low image numbers.

3.5. 5-Bromoindole–DHP complex

Data and map quality followed a similar pattern with reducing crystal numbers to that described above for DCP (Table 1, Fig. 2, Supplementary Figs. S8 and S9). In this case, the lower occupancy of the two 5BR ligands was automatically found in *Coot* with a data set from 2000 images, but this step failed with 1000 images. For the higher occupancy 5BR ligand, the correct pose was found down to a data set of 800 images, while in data sets comprising 500, 600 or 700 images an incorrect pose was found by *Coot*, although manual re-orientation was straightforward based on the $F_o - F_c$ map. A remarkable observation is that even in a data set comprising only 200 images (with 75.3% data completeness) the heavier Br atom of the ligand was clearly identified, with OMIT map peaks of 6.8σ (1.83 $e^- Å^{-3}$) in monomer *B* and 4.3σ (1.16 $e^- Å^{-3}$) in monomer *A* at its position (Supplementary Figs. S8 and S9).

3.6. Imidazole complex of DtpAa

The imidazole ligand provided a more challenging case owing to its smaller size in comparison to DCP and 5BR and

because of the lower symmetry space group ($P2_1$) of the DtpAa crystals. For the latter reason, the merging statistics deteriorated more rapidly than for DHP (Supplementary Table S3). In particular, data completeness began to deteriorate, with the 2000-image data set being essentially complete, while this was not the case for the 1000-crystal data set. With subsets of 5000 images or larger, *Coot* was able to successfully locate both the heme-coordinated imidazole and the second imidazole ligand located in the inter-monomer cleft [Supplementary Fig. S3(*c*)]. With smaller subsets, the latter ligand was not found, although the heme-bound imidazole was located in data sets of as few as 800 images. In these very small data sets the imidazole ring as positioned by *Coot* was sometimes rotated around its normal axis while still fitting the symmetrical electron-density feature well, but this was readily corrected by applying simple chemical knowledge, *i.e.* that N atoms rather than C atoms should be forming the coordination bond to the Fe atom and be oriented towards the Asp residue in the heme pocket. Simulated-annealing OMIT maps for the DtpAa–imidazole complex are shown in Fig. 2 and Supplementary Fig. S10, with electron-density statistics in Fig. 3 and Supplementary Fig. S7. We note that for all three ligands, even when automated ligand finding failed, significant ligand density was present that could allow manual identification in cases where the binding pocket was known in advance.

3.7. *Ac*NiR complex with nitrite

Although nitrite is the smallest ligand of interest used in this study, *Ac*NiR has the inherent characteristic of crystallizing in a high-symmetry space group ($P2_13$), resulting in fewer data being required for a complete data set owing to the high redundancy of the data collected (Table 1). Again, very clear $F_o - F_c$ simulated-annealing OMIT map features for the ligand were evident in subsets of small numbers of diffraction patterns, despite exhibiting merging statistics that would typically be considered rather poor (Figs. 2 and 3, Supplementary Table S4 and Supplementary Fig. S11). *Coot* successfully located nitrite binding at the type 2 Cu active site in subsets of very few crystals, with 200 being the lowest number of crystals that were needed to successfully auto-find the nitrite ligand. Although *Coot* was unsuccessful at determining the ligand in the lowest crystal subset of only 100 crystals, positive electron density is still identifiable at the site where ligand binding is expected, although this did not allow for reasonable modelling of a ligand.

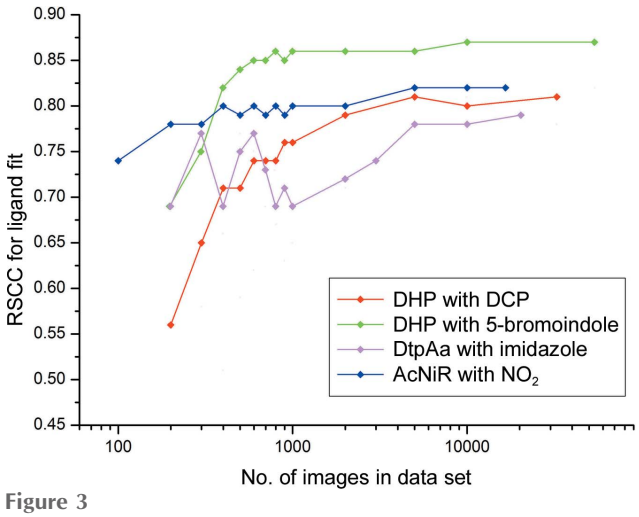


Figure 3
Real-space correlation coefficient (RSCC) values from *EDIAscorer* (Meyder *et al.*, 2017) as a function of the number of images per subset. Data are shown for the highest occupancy binding site for each complex. A plot including values for additional binding sites is shown in Supplementary Fig. S7.

3.8. OMIT maps from simulated-annealing refinement against ligand-free structures of the native enzymes

Although simulated-annealing refinement as described above would reasonably be expected to remove all model bias, as an additional validation step selected subsets were refined against the corresponding native structures obtained by SFX (Ebrahim, Moreno-Chicano *et al.*, 2019; Moreno-Chicano *et al.*, manuscript in preparation), where the ligands were not present. OMIT map generation followed an identical procedure to that described above, with the exception of the input coordinate file used. The resulting OMIT maps are shown in Fig. 4 and Supplementary Figs. S12, S13 and S14) for data subsets of differing sizes. The results of this process corresponded well with the previously described OMIT maps, suggesting that model bias was not significant in the previous procedure for any of the complexes. Notably, for the two DHP ligand structures, in addition to very clear OMIT map density for the ligands themselves the map features clearly define the movements of heme-pocket residues that are necessary to accommodate the ligand (Fig. 4 and Supplementary Fig. S12). This provides further evidence of the information content within these data sets, despite the low numbers of diffraction patterns and extremely poor statistics. Importantly, we used *Ac*NiR–nitrite as a very challenging case to test the limitations of our approach as the nitrite ligand contains only three atoms and also displaces a water molecule upon binding (Antonyuk *et al.*, 2005). In addition, the water density in the native structure is disordered, with the presence of a second water molecule a possibility. Notably, refinement of *Ac*NiR data and subsets versus the native *Ac*NiR SFX structure produced clear positive difference map features for the nitrite atoms that are separated from the water molecule present in the native structure (Supplementary Fig. S14). For comparison, SA OMIT maps produced from refinement of the same subsets against the native *Ac*NiR SFX model with the copper-coordinated water molecule deleted are shown in Supplementary Fig. S15.

3.9. Detection of differences between ligands from $F_o - F_o$ isomorphous difference maps

For the DHP case, in which two different ligands bind in a similar binding pose to the same enzyme pocket, we tested the ability to distinguish between these ligands using $F_o - F_o$ isomorphous difference maps. For the full data sets, an F_o(DHP–5BR) $- F_o$(DHP–DCP) map is shown in Fig. 5. Strong positive density (a 32σ peak) is present where the Br

atom of 5BR occupies a similar position to a Cl atom of DCP, consistent with the larger number of electrons on the Br atom. A negative feature is present over the second Cl atom of DCP, consistent with a C atom occupying a similar position in 5BR. Finally, a positive peak is present for the C5 atom of 5BR where no equivalent atom is present in DCP. As the number of crystals in a data subset decreases, the map features become less prominent, with the C5 feature disappearing in subsets of 1000 crystals or smaller. However, the features corresponding to the Br and Cl atoms are remarkably still evident, albeit much weaker, in subsets comprised of as few as 200 crystals (Fig. 5).

4. Discussion

4.1. High-throughput determination of ligand-bound SFX structures using fixed targets

Our results indicate that high-quality SFX crystal structures that allow unambiguous ligand identification may be achieved using our fixed-target approach. This can be achieved using a small quantity of enzyme sample with high throughput and rapid switchover between different proteins and ligands. Typical data-collection times for complexes were 30–60 min using all data measured, and using these data sets ligand modelling was clear and unambiguous. In comparison to previously published data for ligand-binding experiments (Naitow *et al.*, 2017; Bublitz *et al.*, 2015), the fixed-target approach allows the high-throughput production of multiple intact enzyme–ligand complex structures. In addition, the soaking and data-collection strategy can be easily adapted and optimized for a synchrotron beamline using the same sample-loading and mounting system.

Our results cover several different ligand-binding scenarios, such as coordinate-bond formation to a heme iron (imidazole) or copper (nitrite) and noncoordinate ligand binding in a pocket (DHP ligands), with the latter being highly relevant to the binding of ligands to pharmacologically important proteins such as cytochromes P450. In each structure, binding sites are present with different occupancies, allowing a further test of the ability of the method to locate high- or low-occupancy ligands.

The limits of our ability to identify ligand binding were tested using the small ligands nitrite (46 Da) and imidazole (68 Da). Both of these are much smaller than the fragments used in fragment-based drug design (FBDD), where 200 Da is a typical molecular weight (Price *et al.*, 2017). In the case of *Ac*NiR, a particular challenge was that nitrite displaces a water molecule on binding. In *Ac*NiR structures determined from single crystals, distinguishing between the electron-density features of active-site waters and nitrite is challenging and requires high-

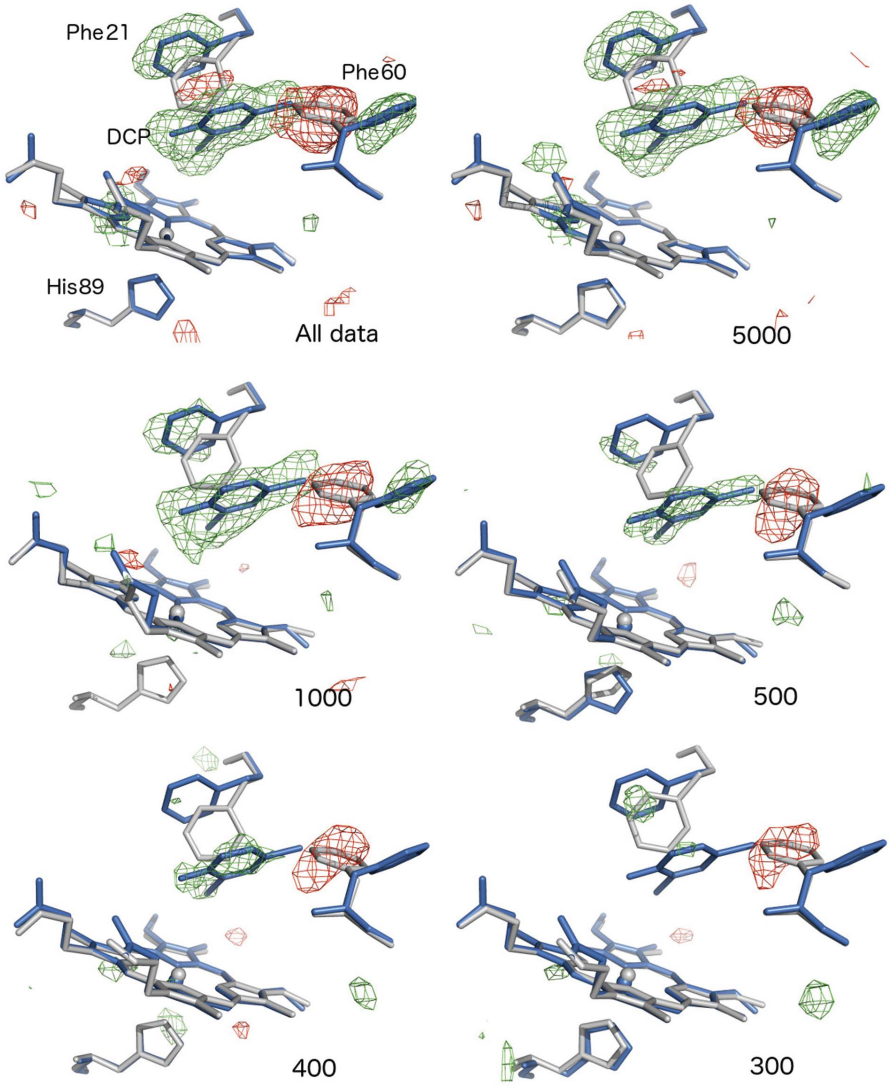


Figure 4
Difference map features produced by simulated-annealing refinement against ligand-free native structures clearly reveal ligand binding and active-site rearrangements in the absence of the risk of model bias. $F_o - F_c$ OMIT maps, contoured at 3σ, are shown for DCP data subsets refined versus the native DHP structure. In each case, the native DHP structure from OMIT refinement versus a particular subset is shown in grey, while the superimposed structure of the ligand complex is shown in blue. Positive difference map features are shown in green, with negative features in red. Note that the flips of Phe21 and Phe60 to accommodate ligand binding, together with the ligand density itself, are very clearly defined in the data set obtained from all data and this is maintained in the 5000-image subset. Clear OMIT map features are apparent for Phe60 and DCP in data sets with as few as 400 images, while this was no longer the case in the 300-image subset.

resolution data (Antonyuk *et al.*, 2005; Horrell *et al.*, 2018). Nonetheless, our method allowed the identification of these ligands in subsets comprising a small fraction of the full data sets. Our results are therefore strongly indicative that the ligands used in FBDD will be readily detected using our approach.

4.2. What is the minimal quantity of data required to identify ligand-binding modes?

Analysis of simulated-annealing OMIT maps generated from data subsets containing only a subset of merged diffraction patterns clearly demonstrates that only a small fraction of the total data-collection time that we used is in fact necessary to locate ligands in the correct binding pose. For example, for the 5BR complex of DHP a subset of just 800 indexed images (∼1.5% of the total number of images in the full data set) was sufficient to correctly model the ligand using a careful strategy to preclude the possibility of model bias. A conservative approach of measuring several times this minimal number would still require only a small proportion of the 25 600 crystal positions on each chip. Our data also show that useful information is contained in data sets obtained from extremely small numbers of microcrystals; for example, the Br atom of 5BR was identified in a data set of only 200 crystals (<0.4% of the total data set).

In a lower symmetry space group (DtpAa; $P2_1$), the ability to detect ligand binding in data sets of <2000 images was compromised by a lack of data completeness at higher resolution, although ligand finding was still achieved with 800 images. Notably, for the DHP structures in space group $P2_12_12_1$ data completeness remained good in very small data sets; for example, for the DCP complex the 400-image data set retained >90% completeness in the highest resolution shell.

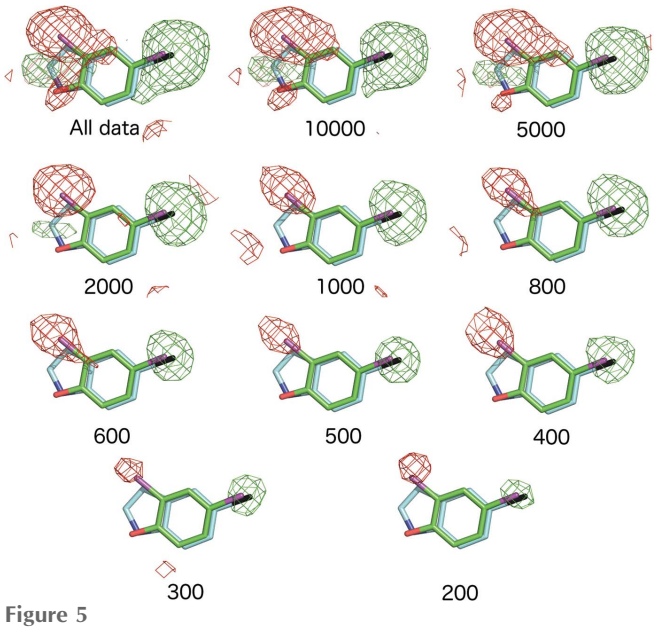


Figure 5
$F_o - F_o$ isomorphous difference maps comparing the 5BR and DCP ligand complexes of DHP. Maps are F_o(5BR) − F_o(DCP) contoured at 3σ. With all data included, the map shows a clear positive peak near to the position of the Br atom of 5BR (black) and one of the Cl atoms of DCP (magenta), consistent with the greater number of electrons in bromine. A negative peak is present at the position of the second Cl atom of DCP, where the closest atom of 5BR is a carbon. An additional but weaker positive peak is present close to the C5 atoms of 5BR where no atom is present in DCP.

The high completeness of data sets formed from (relatively) small numbers of crystals parallels the success in forming complete data sets from multiple thin wedges in virus crystallography (Fry *et al.*, 1999). The completeness of the final data set is a function of the number of wedges collected and the point group of the crystals used, with the prerequisite for each approach being that the crystals must be randomly orientated. The completeness of the data obtained from small numbers of crystals here illustrates that this is the case for DtpAa, DHP and *Ac*NiR crystals on silicon chips. The bandwidth of the XFEL beam allows complete data to be obtained from fewer crystals than would be the case with a more monochromatic beam, yet still requires many more crystals than might be required in a wide-bandpass Laue experiment (Meents *et al.*, 2017). Our data strongly suggest that data completeness is the key metric for assessing the suitability of data sets for ligand-binding studies and that very poor values of other typically used metrics of data quality (for example $CC_{1/2}$ and R_{split}) still allow successful ligand characterization provided that the data are complete. For *Ac*NiR, with cubic symmetry, the data remained essentially complete in all of the subset sizes analysed, with density for the nitrite ligand remaining apparent down to <200 indexed images. We note that substantially more diffraction patterns would be required to obtain complete data on a monochromatic beamline.

More broadly, our data clearly show that substantial information content is present in noisy and apparently low-quality data sets derived from small numbers of merged diffraction patterns with very poor merging and refinement statistics. For example, a data set formed of 200 patterns revealed a very clear peak for the Br atom of the 5BR ligand (outer shell completeness 70.9% in DCP). Importantly, refinement of data subsets against native structures unambiguously showed not only clear density for ligands, but also any movements of the active-site residues needed to accommodate ligand binding (Fig. 4 and Supplementary Fig. S12). This provides conclusive evidence that the ligand density that we describe is not owing to model bias from prior knowledge of the binding mode.

4.3. Future potential of the 'chip-soak' approach for high-throughput structure determination of protein–ligand complexes

In this work, SFX structures were recorded from two (*Ac*NiR), three (DHP–DCP) or four (DHP–5BR and DtpAa–imidazole) chips, aiming for 1–2 structures per hour. The number of chips used for a single structure was subsequently seen to err significantly on the side of caution, as in all cases sufficient data for unambiguous ligand identification were available from significantly less than half a chip. Crucially, careful data analysis demonstrated that data sets comprising of no more than a few hundred to a few thousand indexed

images are sufficient to correctly model ligands into clear difference density features. Thus, without modification of the approach or changes to the experimental conditions, an approximately 4–5-fold increased throughput of multiple protein–ligand structures per hour could easily be realized. Rapid on-site data analysis should allow on-the-fly decision making as to whether sufficient data have been collected for a particular soak and if a ligand is indeed bound. A key advantage of the fixed-target sample-delivery method is that switching between samples of different protein–ligand soaks is no more time-consuming than continuing with a chip of the same sample. With typical loading rates of approximately 30%, multiple ligand soaks could be carried out on a single redesigned chip, again drastically increasing throughput. As a further example, for systems where approximately 1000 hits would be sufficient, at the latter hit rate some eight ligand complexes could be characterized on a single chip.

The sample quantity required for our approach (in the range of 1.35–6.0 mg protein per data set) is less than required in liquid-jet approaches, although higher than has been reported for high-viscosity (LCP) injection systems at XFEL (Weierstall *et al.*, 2014) and synchrotron (Weinert *et al.*, 2017) beamlines. An additional factor is ligand consumption. In our case, without optimization to minimize sample consumption, the typical ligand quantities used were in the range 4–40 µmol.

Our system of work is applicable at other current and future XFEL sources, such as PAL (60 Hz repetition rate), SwissFEL (100 Hz) and LCLS (120 Hz), as well as SACLA (30 Hz). However, XFEL sources with very high repetition rates or complex pulse patterns (for example EuXFEL and LCLS-II) may require a modified or different approach. We have demonstrated that at a source with a modest repetition rate sufficient data for multiple, unrelated, protein–ligand structures may be obtained within a couple of hours. Increasing this level of throughput to ∼5–20 structures per hour at higher repetition-rate sources, or collecting fewer images per complex (see above) as is practical, would allow, for example, >200 structures to be determined in a single 12 h shift, similar to dedicated synchrotron beamlines. Fixed targets are also well suited to time-resolved crystallography of, for example, protein–ligand complexes using laser pump–probe methods (Schulz *et al.*, 2018) and it is important to note that in time-resolved experiments significantly more data may be required as crystals may contain a mix of states.

Another key advantage is that the chip approach allows us to test soaking protocols at synchrotron beamlines under identical conditions to those used at the XFEL in order to ensure that soaking does not damage crystals and also that ligands are bound, albeit in a radiation-damaged structure. At such high rates of sample delivery, automation of chip loading and robotic sample exchange will of course become increasingly important. Our work demonstrates the feasibility of high-throughput room-temperature ligand screening by SFX using microcrystals and is highly applicable to drug-discovery efforts, including in fragment-based drug design. Our approach would be of particular importance in cases where only small weakly diffracting crystals are obtained or when the enzyme–ligand complexes are radiation-sensitive. We have demonstrated the ability to identify ligand binding by our high-throughput approach using 'conventional' approaches to both refinement and ligand finding. Further data-analysis improvements to the ability to identify in particular low-occupancy ligands in FBDD could be achieved using a multi-data-set approach, for example in *PanDDa*, with subtraction of the ligand-free ground state (Pearce, Krojer, Bradley *et al.*, 2017) and with refinement against a composite of the ligand-free and ligand-bound structures (Pearce, Krojer & von Delft, 2017).

In conclusion, we demonstrate (i) a method to rapidly measure SFX data sets from protein–ligand complexes and to rapidly switch between ligands during beamtime, (ii) that data sets comprised of hundreds to a few thousands of diffraction patterns can be sufficient for unambiguous ligand identification and (iii) that even ligands smaller than those used in fragment-based drug design may be located using our approach. These data demonstrate the feasibility of high-throughput structure determination of protein–ligand complexes at XFEL sources.

Acknowledgements

XFEL experiments were performed on BL2 EH3 at SACLA with the approval of the Japan Synchrotron Radiation Research Institute (JASRI; Proposal No. 2017B8014). We acknowledge the contributions of Drs Minoru Kubo and Takashi Nomura (University of Hyogo) and Dr Takehiko Tosha (RIKEN). Support for travel from the UK XFEL Hub at Diamond Light Source is gratefully acknowledged. We are grateful to an anonymous reviewer of the manuscript for the suggestion of generating isomorphous difference maps for the two DHP ligand-bound structures. Author contributions are as follows. RLO, JARW, RWS, RAG and MAH designed the experiment. AE, TM-C, AKC and JARW were involved in sample preparation. AE, TM-C, RWS, JARW, MVA, JHB, DA, DAS, HMED, HS, KT, SO, RLO and MAH participated in data collection. AE, TM-C, HMED, MVA, DA, RLO and MAH analysed the data. AE, TM-C, RLO and MAH wrote the manuscript with contributions from all authors.

Funding information

AE was supported by a joint Diamond–University of Essex PhD studentship. TM-C was supported by Leverhulme Trust award RPG-2014-355 to MAH and RWS. MH acknowledges support from BBSRC award BB/M022714/1 and JARW, DAS, DA, RLO, MAH and RWS acknowledge support from BBSRC Japan–UK International Partnering Award BB/R021015/1. HMED was supported by The Wellcome Trust (award ALR00750). RAG was supported by National Science Foundation Award CHE-1609446. The University of Essex and Diamond Light Source contributed equally to this work.

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-

Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* D**66**, 213–221.
Antonyuk, S. V., Strange, R. W., Sawers, G., Eady, R. R. & Hasnain, S. S. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 12041–12046.
Barrios, D. A., D'Antonio, J., McCombs, N. L., Zhao, J., Franzen, S., Schmidt, A. C., Sombers, L. A. & Ghiladi, R. A. (2014). *J. Am. Chem. Soc.* **136**, 7914–7925.
Barty, A., Kirian, R. A., Maia, F. R. N. C., Hantke, M., Yoon, C. H., White, T. A. & Chapman, H. (2014). *J. Appl. Cryst.* **47**, 1118–1131.
Beitlich, T., Kühnel, K., Schulze-Briese, C., Shoeman, R. L. & Schlichting, I. (2007). *J. Synchrotron Rad.* **14**, 11–23.
Bublitz, M., Nass, K., Drachmann, N. D., Markvardsen, A. J., Gutmann, M. J., Barends, T. R. M., Mattle, D., Shoeman, R. L., Doak, R. B., Boutet, S., Messerschmidt, M., Seibert, M. M., Williams, G. J., Foucar, L., Reinhard, L., Sitsel, O., Gregersen, J. L., Clausen, J. D., Boesen, T., Gotfryd, K., Wang, K.-T., Olesen, C., Møller, J. V., Nissen, P. & Schlichting, I. (2015). *IUCrJ*, **2**, 409–420.
Collins, P. M., Douangamath, A., Talon, R., Dias, A., Brandao-Neto, J., Krojer, T. & von Delft, F. (2018). *Methods Enzymol.* **610**, 251–264.
Colpa, D. I., Fraaije, M. W. & van Bloois, E. (2014). *J. Ind. Microbiol. Biotechnol.* **41**, 1–7.
Debreczeni, J. É. & Emsley, P. (2012). *Acta Cryst.* D**68**, 425–430.
Ebrahim, A., Appleby, M. V., Axford, D., Beale, J., Moreno-Chicano, T., Sherrell, D. A., Strange, R. W., Hough, M. A. & Owen, R. L. (2019). *Acta Cryst.* D**75**, 151–159.
Ebrahim, A., Moreno-Chicano, T., Appleby, M. V., Chaplin, A. K., Beale, J. H., Sherrell, D. A., Duyvesteyn, H. M. E., Owada, S., Tono, K., Sugimoto, H., Strange, R. W., Worrall, J. A. R., Axford, D., Owen, R. L. & Hough, M. A. (2019). *IUCrJ*, **6**, 543–551.
Emsley, P. (2017). *Acta Cryst.* D**73**, 203–210.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Fischer, M., Shoichet, B. K. & Fraser, J. S. (2015). *ChemBioChem*, **16**, 1560–1564.
Franzen, S., Thompson, M. K. & Ghiladi, R. A. (2012). *Biochim. Biophys. Acta*, **1824**, 578–588.
Fry, E. E., Grimes, J. & Stuart, D. I. (1999). *Mol. Biotechnol.* **12**, 13–23.
Guengerich, F. P., Waterman, M. R. & Egli, M. (2016). *Trends Pharmacol. Sci.* **37**, 625–640.
Horrell, S., Antonyuk, S. V., Eady, R. R., Hasnain, S. S., Hough, M. A. & Strange, R. W. (2016). *IUCrJ*, **3**, 271–281.
Horrell, S., Kekilli, D., Sen, K., Owen, R. L., Dworkowski, F. S. N., Antonyuk, S. V., Keal, T. W., Yong, C. W., Eady, R. R., Hasnain, S. S., Strange, R. W. & Hough, M. A. (2018). *IUCrJ*, **5**, 283–292.
Horrell, S., Kekilli, D., Strange, R. W. & Hough, M. A. (2017). *Metallomics*, **9**, 1470–1482.
Inoue, I., Inubushi, Y., Sato, T., Tono, K., Katayama, T., Kameshima, T., Ogawa, K., Togashi, T., Owada, S., Amemiya, Y., Tanaka, T., Hara, T. & Yabashi, M. (2016). *Proc. Natl Acad. Sci. USA*, **113**, 1492–1497.
Ishikawa, T., Aoyagi, H., Asaka, T., Asano, Y., Azumi, N., Bizen, T., Ego, H., Fukami, K., Fukui, T., Furukawa, Y., Goto, S., Hanaki, H., Hara, T., Hasegawa, T., Hatsui, T., Higashiya, A., Hirono, T., Hosoda, N., Ishii, M., Inagaki, T., Inubushi, Y., Itoga, T., Joti, Y., Kago, M., Kameshima, T., Kimura, H., Kirihara, Y., Kiyomichi, A., Kobayashi, T., Kondo, C., Kudo, T., Maesaka, H., Marechal, X. M., Masuda, T., Matsubara, S., Matsumoto, T., Matsushita, T., Matsui, S., Nagasono, M., Nariyama, N., Ohashi, H., Ohata, T., Ohshima, T., Ono, S., Otake, Y., Saji, C., Sakurai, T., Sato, T., Sawada, K., Seike, T., Shirasawa, K., Sugimoto, T., Suzuki, S., Takahashi, S., Takebe, H., Takeshita, K., Tamasaku, K., Tanaka, H., Tanaka, R., Tanaka, T., Togashi, T., Togawa, K., Tokuhisa, A., Tomizawa, H., Tono, K., Wu, S. K., Yabashi, M., Yamaga, M., Yamashita, A., Yanagida, K., Zhang, C., Shintake, T., Kitamura, H. & Kumagai, N. (2012). *Nat. Photonics*, **6**, 540–544.
Keedy, D. A., Hill, Z. B., Biel, J. T., Kang, E., Rettenmaier, T. J., Brandao-Neto, J., Pearce, N. M., von Delft, F., Wells, J. A. & Fraser, J. S. (2018). *Elife*, **7**, e36307.
Kekilli, D., Moreno-Chicano, T., Chaplin, A. K., Horrell, S., Dworkowski, F. S. N., Worrall, J. A. R., Strange, R. W. & Hough, M. A. (2017). *IUCrJ*, **4**, 263–270.
Lomb, L., Barends, T. R. M., Kassemeyer, S., Aquila, A., Epp, S. W., Erk, B., Foucar, L., Hartmann, R., Rudek, B., Rolles, D., Rudenko, A., Shoeman, R. L., Andreasson, J., Bajt, S., Barthelmess, M., Barty, A., Bogan, M. J., Bostedt, C., Bozek, J. D., Caleman, C., Coffee, R., Coppola, N., DePonte, D. P., Doak, R. B., Ekeberg, T., Fleckenstein, H., Fromme, P., Gebhardt, M., Graafsma, H., Gumprecht, L., Hampton, C. Y., Hartmann, A., Hauser, G., Hirsemann, H., Holl, P., Holton, J. M., Hunter, M. S., Kabsch, W., Kimmel, N., Kirian, R. A., Liang, M. N., Maia, F. R. N. C., Meinhart, A., Marchesini, S., Martin, A. V., Nass, K., Reich, C., Schulz, J., Seibert, M. M., Sierra, R., Soltau, H., Spence, J. C. H., Steinbrener, J., Stellato, F., Stern, S., Timneanu, N., Wang, X. Y., Weidenspointner, G., Weierstall, U., White, T. A., Wunderer, C., Chapman, H. N., Ullrich, J., Strüder, L. & Schlichting, I. (2011). *Phys. Rev. B*, **84**, 214111.
Long, F., Nicholls, R. A., Emsley, P., Gražulis, S., Merkys, A., Vaitkus, A. & Murshudov, G. N. (2017). *Acta Cryst.* D**73**, 112–122.
McCombs, N. L., Moreno-Chicano, T., Carey, L. M., Franzen, S., Hough, M. A. & Ghiladi, R. A. (2017). *Biochemistry*, **56**, 2294–2303.
McCombs, N. L., Smirnova, T. & Ghiladi, R. A. (2017). *Catal. Sci. Technol.* **7**, 3104–3118.
McGuire, A. H., Carey, L. M., de Serrano, V., Dali, S. & Ghiladi, R. A. (2018). *Biochemistry*, **57**, 4455–4468.
McLean, K. J. & Munro, A. W. (2017). *Drug Discov. Today*, **22**, 566–575.
McPherson, A. (2019). *Acta Cryst.* F**75**, 132–140.
Meents, A., Wiedorn, M. O., Srajer, V., Henning, R., Sarrou, I., Bergtholdt, J., Barthelmess, M., Reinke, P. Y. A., Dierksmeyer, D., Tolstikova, A., Schaible, S., Messerschmidt, M., Ogata, C. M., Kissick, D. J., Taft, M. H., Manstein, D. J., Lieske, J., Oberthuer, D., Fischetti, R. F. & Chapman, H. N. (2017). *Nat. Commun.* **8**, 1281.
Meyder, A., Nittinger, E., Lange, G., Klein, R. & Rarey, M. (2017). *J. Chem. Inf. Model.* **57**, 2437–2447.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Naitow, H., Matsuura, Y., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Tanaka, R., Tanaka, T., Sugahara, M., Kobayashi, J., Nango, E., Iwata, S. & Kunishima, N. (2017). *Acta Cryst.* D**73**, 702–709.
Nass, K. (2019). *Acta Cryst.* D**75**, 211–218.
Nass, K., Foucar, L., Barends, T. R. M., Hartmann, E., Botha, S., Shoeman, R. L., Doak, R. B., Alonso-Mori, R., Aquila, A., Bajt, S., Barty, A., Bean, R., Beyerlein, K. R., Bublitz, M., Drachmann, N., Gregersen, J., Jönsson, H. O., Kabsch, W., Kassemeyer, S., Koglin, J. E., Krumrey, M., Mattle, D., Messerschmidt, M., Nissen, P., Reinhard, L., Sitsel, O., Sokaras, D., Williams, G. J., Hau-Riege, S., Timneanu, N., Caleman, C., Chapman, H. N., Boutet, S. & Schlichting, I. (2015). *J. Synchrotron Rad.* **22**, 225–238.
Oghbaey, S., Sarracini, A., Ginn, H. M., Pare-Labrosse, O., Kuo, A., Marx, A., Epp, S. W., Sherrell, D. A., Eger, B. T., Zhong, Y., Loch, R., Mariani, V., Alonso-Mori, R., Nelson, S., Lemke, H. T., Owen, R. L., Pearson, A. R., Stuart, D. I., Ernst, O. P., Mueller-Werkmeister, H. M. & Miller, R. J. D. (2016). *Acta Cryst.* D**72**, 944–955.
Pearce, N. M., Krojer, T., Bradley, A. R., Collins, P., Nowak, R. P., Talon, R., Marsden, B. D., Kelm, S., Shi, J., Deane, C. M. & von Delft, F. (2017). *Nat. Commun.* **8**, 15123.
Pearce, N. M., Krojer, T. & von Delft, F. (2017). *Acta Cryst.* D**73**, 256–266.
Price, A. J., Howard, S. & Cons, B. D. (2017). *Essays Biochem.* **61**, 475–484.

Rendic, S. & Guengerich, F. P. (2015). *Chem. Res. Toxicol.* **28**, 38–42.
Schlichting, I. (2015). *IUCrJ*, **2**, 246–255.
Schulz, E. C., Mehrabi, P., Müller-Werkmeister, H. M., Tellkamp, F., Jha, A., Stuart, W., Persch, E., De Gasparo, R., Diederich, F., Pai, E. F. & Miller, R. J. D. (2018). *Nat. Methods*, **15**, 901–904.
Smart, O. S., Horský, V., Gore, S., Svobodová Vařeková, R., Bendová, V., Kleywegt, G. J. & Velankar, S. (2018). *Acta Cryst.* D**74**, 228–236.
Sugano, Y. (2009). *Cell. Mol. Life Sci.* **66**, 1387–1403.
Weierstall, U., James, D., Wang, C., White, T. A., Wang, D., Liu, W., Spence, J. C. H., Doak, R. B., Nelson, G., Fromme, P., Fromme, R., Grotjohann, I., Kupitz, C., Zatsepin, N. A., Liu, H., Basu, S., Wacker, D., Han, G. W., Katritch, V., Boutet, S., Messerschmidt, M., Williams, G. J., Koglin, J. E., Seibert, M. M., Klinker, M., Gati, C., Shoeman, R. L., Barty, A., Chapman, H. N., Kirian, R. A., Beyerlein, K. R., Stevens, R. C., Li, D., Shah, S. T., Howe, N., Caffrey, M. & Cherezov, V. (2014). *Nat. Commun.* **5**, 3309.
Weinert, T., Olieric, N., Cheng, R., Brünle, S., James, D., Ozerov, D., Gashi, D., Vera, L., Marsh, M., Jaeger, K., Dworkowski, F., Panepucci, E., Basu, S., Skopintsev, P., Doré, A. S., Geng, T., Cooke, R. M., Liang, M., Prota, A. E., Panneels, V., Nogly, P., Ermler, U., Schertler, G., Hennig, M., Steinmetz, M. O., Wang, M. & Standfuss, J. (2017). *Nat. Commun.* **8**, 542.
White, T. A., Mariani, V., Brehm, W., Yefanov, O., Barty, A., Beyerlein, K. R., Chervinskii, F., Galli, L., Gati, C., Nakane, T., Tolstikova, A., Yamashita, K., Yoon, C. H., Diederichs, K. & Chapman, H. N. (2016). *J. Appl. Cryst.* **49**, 680–689.
Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S. & Richardson, J. S. (2018). *Protein Sci.* **27**, 293–315.