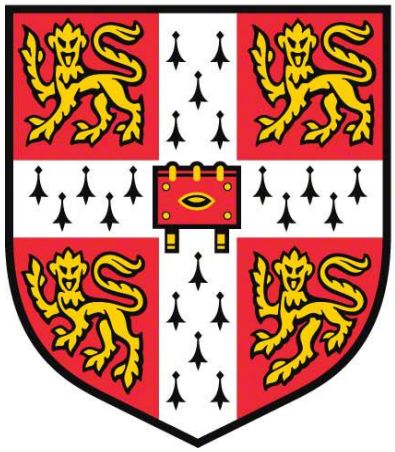
Proteomic Investigation of Interferon Alpha Stimulation and Viral Regulation to Identify Candidate Antiviral Restriction Factors



Lior Vered Soday

Department of Medicine

Cambridge Institute for Medical Research

Jesus College

This thesis is submitted for the degree of

Doctor of Philosophy

University of Cambridge

May 2020

DECLARATION

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Lior Vered Soday

May 2020

Proteomic Investigation of Interferon Alpha Stimulation and Viral Regulation to Identify Candidate Antiviral Restriction Factors

Lior Vered Soday

SUMMARY

Antiviral restriction factors (ARFs) are host proteins that play key roles in inhibiting viral infection. The plasma membrane provides a critical interface between the cell and the virus, meaning that proteins present here are well situated to act as ARFs. An understating of these ARFs and how viruses interact with them, is crucial to our knowledge of infection and immunity, and provides potential therapeutic targets. In this thesis, I apply quantitative, multiplexed proteomic technologies to explore two characteristic features of ARFs: induction by interferon (IFN), and downregulation by viral infection.

Proteomics was used to investigate cell surface changes in primary monocytes and CD4+ T-cells upon stimulation with IFNα, across five donors. The cell surface proteomes were characterised, and found to be remarkably invariant between donors, whilst the effects of IFN were more variable between donors. IFN stimulation of several proteins was validated by flow cytometry, and TMEM123 was identified as the only protein, aside from MHC class I molecules and the known restriction factor Tetherin, to be consistently upregulated by IFN at the cell surface in all donors for both cell types.

Additional proteomic screens investigated the effects of vaccinia virus (VACV) infection, quantifying ~9000 human proteins over a single replication cycle. This revealed downregulation of 265 proteins including innate immune proteins, collagens and cadherins. Overlap with a previous proteomic investigation of human cytomegalovirus (HCMV) infection demonstrated that many classes of proteins were commonly targeted, suggesting an important role in infection and immunity. Of the proteins downregulated

by VACV, ~70 % were proteasomally degraded. In addition to human proteins, ~80 % of all VACV proteins were quantified. These were classified into four distinct temporal classes, which correlated well with previous literature, and with known protein functions. Host-virus interactions were investigated by matching temporal profiles of human and viral proteins. HDAC5 was found to be targeted for proteasomal degradation by the viral protein C6. Overexpression and knockdown of HDAC5 demonstrated its restriction of both VACV and herpes simplex virus-1 replication.

Finally, both datasets were considered alongside published proteomic data on other viral infections, in particular HCMV and HIV, to identify putative ARFs. Expression of endothelin converting enzyme 1 (ECE1) was stimulated by IFN at the surface of CD4+ T-cells and was downregulated during HCMV infection. Follow-up studies to determine if ECE1 restricts HCMV have thus far led to inconclusive results.

Overall, these investigations into IFN stimulation and VACV infection have yielded very useful data, and provide valuable, comprehensive resources for future research.

ACKNOWLEDGEMENTS

There are many individuals I would like to thank for enabling me to conduct this research, and for making my PhD experience so enjoyable. First and foremost, I would like to thank my supervisor, Dr Mike Weekes, for his advice, guidance and support throughout the PhD. He has been instrumental in my development as a researcher, and I am extremely grateful for the opportunity to perform this research in his laboratory.

I would also like to thank all the members of the Weekes lab. The friendly atmosphere and everyone's willingness to help has made it a great place to work. In particular, I would like to thank Dr Katie Nightingale, Dr Luis Nobre and Alice Fletcher-Etherington for providing me with many of the cell lines used in this research, and Dr Colin Davies for fractionating my samples. Aside from scientific help, I would also like to thank them for creating such a lovely environment in the lab. They have kept me smiling throughout the joys and the struggles of the PhD, and I am so fortunate to have spent my time in the lab with them.

There are a number of others I would like to acknowledge who have also contributed to this work. I would like to thank Dr Robin Antrobus, from the CIMR Proteomics Core Facility, and Dr James Williamson from the Lehner Lab, who ran the proteomic samples discussed in this thesis on the mass spectrometer. The members of the Smith Lab, in particular Dr Jonas Albarnaz and Dr Yongxu Lu, who enabled all the research on vaccinia virus presented here. The vaccinia project has been both interesting and enjoyable, and I am extremely grateful for the opportunities the collaboration has provided. Publishing this work as a joint first author, and presenting it at the poxvirus conference in Taiwan, were highlights of my PhD. I would also like to thank the Wellcome Trust for the funding that made all this possible.

I would not be here without the help and encouragement I have received throughout my studies; from teachers at The Latymer School, to supervisors at Newnham, and all the scientists I have been so fortunate to work with during summer research projects,

undergraduate projects, and lab rotations. I am so grateful for everyone that has taught and guided me.

Finally, thank you to my friends and family. The Newnham girls, who have been a massive part of what first made me love being in Cambridge. Tori and Krittika for being here throughout the ups and downs of the last four years, and always with cheese. El for making Marshall Road a true home. Life in Cambridge would not have been the same without you all. Thank you to Rajiev, Chadwell Heath is not where I imagined I would be writing the last of this thesis, but there is no where I would rather have spent the COVID lockdown. I am so glad I get to finish 'our' PhD here with you. Thank you to all my family, Booba and Zada, those in Israel, and especially to Alan and Adi for always being there. Finally, to Mum. There are not the words to say how grateful I am for everything you do. Anything I achieve is because of the opportunities you have given me, and your encouragement, love and support. Thank you.

PUBLICATIONS

Key manuscript referenced in thesis

Soday L, Lu Y, Albarnaz JD, Davies CTR, Antrobus R, Smith GL, Weekes MP. "Quantitative Temporal Proteomic Analysis of Vaccinia Virus Infection Reveals Regulation of Histone Deacetylases by an Interferon Antagonist". **Cell Reports (2019)** 27: 1920-1933.e7.

Other articles published during PhD

Ravenhill BJ, **Soday L**, Houghton J, Antrobus R, Weekes MP. "Comprehensive cell surface proteomics defines markers of classical, intermediate and non-classical monocytes". **Sci Rep (2020)** 10: 1–11.

Nobre L, Nightingale K, Ravenhill BJ, Antrobus R, **Soday L**, Nichols J, Davies J, Seirafian S, Wang ECY, Davison AJ, Wilkinson GWG, Stanton RJ, Huttlin EL & Weekes MP. "Human cytomegalovirus interactome analysis identifies degradation hubs, domain associations and viral protein functions". **Elife (2019)**, 8: 1–35

Nightingale K, Lin KM, Ravenhill BJ, Davies C, Nobre L, Fielding CA, Ruckova E, Fletcher-Etherington A, **Soday L**, Nichols H, Sugrue D, Wang ECYY, Moreno P, Umrania Y, Huttlin EL, Antrobus R, Davison AJ, Wilkinson GWGG, Stanton RJ, Tomasec P, Weekes MP. "High-Definition Analysis of Host Protein Stability during Human Cytomegalovirus Infection Reveals Antiviral Factors and Viral Evasion Mechanisms". **Cell Host Microbe (2018)**, 24: 447-460 e11

Ersing I, Nobre L, Wang LW, **Soday L**, Ma Y, Paulo JA, Narita Y, Ashbaugh CW, Jiang C, Grayson NE, Kieff E, Gygi SP, Weekes MP & Gewurz BE. "A Temporal Proteomic Map of Epstein-Barr Virus Lytic Replication in B Cells". **Cell Rep (2017)**, 19: 1479–1493

TABLE OF CONTENTS

TABLE OF CONTENTS ..... **vii**
LIST OF FIGURES ..... **xii**
LIST OF TABLES ..... **xv**
ABBREVIATIONS ..... **xvii**

CHAPTER 1: INTRODUCTION **1**
1.1 Introduction 1
1.2 Aim: Identify Candidate Antiviral Restriction Factors 2
1.2.1 Defining Antiviral Restriction Factors 2
1.2.2 Known Restriction Factors 4
1.2.3 Therapeutic Potential of Restriction Factors 6
1.3 Screen 1: Interferon 10
1.3.1 Types of Interferon 10
1.3.2 Type I interferon signalling pathways 12
1.3.3 Effects of Interferon 13
1.3.4 IFN Responses in Leukocytes 15
1.4 Screen 2: Systems Analyses of Viral Infection 19
1.4.1 Vaccinia Virus 19
1.4.2 Human Cytomegalovirus 23
1.4.3 HIV 29
1.4.4 Other TMT Based Quantitative Temporal Analyses of Viral Infection 32
1.5 Key Technology Used: Proteomics 32
1.5.1 Basic Methodology for Bottom-Up Proteomics 33
1.5.2 Tandem Mass Tags for Quantitative Multiplexed Proteomics 34
1.5.3 Plasma Membrane Profiling 35
1.6 Aims of this Thesis 38

CHAPTER 2: MATERIALS AND METHODS **41**
2.1 Ethics Statements 41
2.2 Cell Culture 41
2.2.1 Cell lines 41
2.2.2 Cell Culture for Maintenance 42
2.3 Primary Cell Enrichment and Culture 43

2.3.1 CD4+ T cells 43
2.3.2 CD14+ and Pan- Monocytes 44
2.4 Generation of Cell Lines 45
2.4.1 Overexpression Cell Line 45
2.4.2 shRNA Knockdown Cell Lines 46
2.4.3 CRISPR Cell Lines 48
2.4.4 shRNA Knockdown and Overexpression in Two Colour System 50
2.4.5 Transformation 50
2.4.6 Confirmation of Insert 51
2.4.7 Stable Cell Line Production by Lentiviral Transduction 51
2.5 Molecular Biology Techniques 53
2.5.1 Flow Cytometry 53
2.5.2 Immunoblotting 55
2.5.3 RT-qPCR 56
2.6 Viruses 57
2.7 HCMV Restriction Assays 58
2.7.1 shRNA, Overexpression and CRISPR Cell Restriction Assays 58
2.7.2 Two Colour Restriction Assay 59
2.7.3 Plaque Assay 59
2.8 Additional Methods from Chapter 4 on VACV 60
2.8.1 Cell Culture 60
2.8.2 Production of Cell Lines 60
2.8.3 Viruses 61
2.8.4 Infection 61
2.8.5 Immunoblots 62
2.9 Proteomics 63
2.9.1 Protein Extraction 64
2.9.2 Peptide Labelling with Tandem Mass Tags 67
2.9.3 Fractionation 70
2.9.4 LC-MS3 72
2.9.5 Data Processing 73
2.9.6 Data Availability 75
2.10 Data Analysis and Statistics 76
2.10.1 Chapter 3: IFN Screen 76
2.10.2 Chapter 4: Vaccinia Virus Screen 77
2.10.3 Chapter 5: Candidate Antiviral Restriction Factors 78

2.11 Software Used 79

CHAPTER 3: INTERFERON SCREEN - The Effect of Interferon at the Surface of Primary Leukocytes 81

3.1 Introduction and Aims 81

3.1.1 Introduction 81

3.1.2 Summary 82

3.1.3 Aims 83

3.2 Validation of the Workflow for the IFN Proteomic Screen 83

3.2.1 Experimental Outline 83

3.2.2 Validation of the Workflow on THP-1s 85

3.3 Enrichment of Primary Leukocytes 86

3.3.1 Optimisation of Primary Monocyte Enrichment 86

3.3.2 Purity of Enrichment 88

3.4 Mass Spectrometry Analysis 89

3.5 Quantitative Analysis of the Cell Surface Proteome 90

3.5.1 Monocyte and T cell Surface Proteomes 90

3.5.2 Comparison of Surface Proteomes 91

3.5.3 Variation between Donors 93

3.6 The Effects of IFN at the Surface of Primary Leukocytes 94

3.6.1 Monocytes 94

3.6.2 CD4+ T Cells 102

3.6.3 Comparison of Monocytes and T cells 105

3.6.4 Variation between Donors 108

3.7 Validation 110

3.8 TMEM123 112

3.8.1 IFN Stimulation of TMEM123 112

3.8.2 TMEM123 112

3.8.3 Testing of TMEM123 Antibodies 113

3.9 Discussion 119

3.9.1 Cell Surface Proteome 119

3.9.2 The Effects of IFN 120

CHAPTER 4: VACCINIA VIRUS SCREEN - Quantitative Temporal Proteomic Analysis of Vaccinia Virus Infection 125

4.1 Introduction and Aims 125

4.1.1 Introduction 125

4.1.2 Summary 126
4.1.3 Aims 127
4.2 Quantitative Temporal Analysis of VACV Infection 128
4.2.1 Optimisation of the Infection Time Course 128
4.2.2 Experimental Outline 128
4.3 Modulation of Host Proteins by VACV Infection 130
4.4 Downregulation of Cell Surface Proteins and Antiviral Factors 133
4.4.1 Enrichment Analysis of Downregulated Proteins 133
4.4.2 Overlap with HCMV 135
4.4.3 Downregulation of Innate Immune Proteins 137
4.4.4 Modulation of Candidate Immunoreceptors 138
4.5 Temporal Analysis of VACV Viral Protein Expression 141
4.5.1 There were Four Temporal Classes of Viral Protein Expression 141
4.5.2 Comparison of viral protein classes to previous studies 144
4.6 Systematic Analysis of Protein Degradation During VACV Infection 147
4.6.1 Experimental Outline 147
4.6.2 Proteasomal Degradation of Human Proteins 148
4.6.3 Proteasomal Degradation of Viral Proteins 150
4.7 Identification of Candidate Virus-Host Interactions: HDAC5 151
4.7.1 HDAC5 was Targeted by VACV C6 Protein 152
4.7.2 HDAC5 Restricts VACV and HSV-1 Infection 154
4.8 Discussion 157
4.8.1 Downregulation of Host Proteins 157
4.8.2 Temporal Regulation of Viral Proteins 161

CHAPTER 5: CANDIDATE ANTIVIRAL RESTRICTION FACTORS 165
5.1 Introduction and Aims 165
5.1.1 Identifying Candidate Antiviral Restriction Factors 165
5.1.2 Aims 165
5.2 Identifying Candidate Antiviral Restriction Factors 166
5.2.1 Overlap of IFN Simulation and Viral Infection Data 166
5.2.2 Candidate Antiviral Restriction Factors 168
5.3 Endothelin Converting Enzyme 1 (ECE1) 170
5.3.1 Isoforms 170
5.3.2 Function 171
5.3.3 Regulation 174

5.3.4 ECE1 in Disease and Infection 175
5.3.5 Potential Roles as a Restriction Factor 175
5.4 ECE1: A Candidate Antiviral Restriction Factor 176
5.4.1 Validation of IFN Stimulation 176
5.4.2 How Does HCMV Downregulate ECE1? 176
5.5 ECE1: Reagents 178
5.5.1 ECE1 Antibodies 178
5.5.2 Testing ECE1 Antibodies by Flow Cytometry 179
5.5.3 Testing ECE1 Antibodies by Immunoblot 181
5.6 ECE1: Does it Restrict HCMV Infection? 182
5.6.1 Restriction Assay Methodology 182
5.6.2 Overexpression 184
5.6.3 shRNA knockdown 186
5.6.4 CRISPR knockdown 188
5.6.5 Two Colour Restriction Assays 194
5.6.6 HCMV Assays in Novel ECE1 shRNA Cell Lines 199
5.7 Summary of ECE1 Results 205
5.8 Discussion 206
5.8.1 How Does HCMV Downregulate ECE1? 206
5.8.2 Does ECE1 Restrict Viral Infection? 206
5.8.3 Other Considerations from the Data 209

CHAPTER 6: DISCUSSION AND FUTURE DIRECTIONS 211
6.1 IFN Screen 211
6.1.1 Summary 211
6.1.2 Discussion and Future Directions 212
6.2 VACV Screen 214
6.2.1 Summary 214
6.2.2 Discussion and Future Directions 215
6.3 ECE1 216
6.4 Identification of ARFs 216
6.5 Concluding Remarks 218

BIBLIOGRAPHY 221
APPENDIX 249

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Figure 1.1 Type I IFN induction and signalling pathways 13

Figure 1.2 Blood lineages derived from hematopoietic stem cells..... 17

Figure 1.3 VACV life cycle..... 21

Figure 1.4 HCMV life cycle 26

Figure 1.5 Schematic of HCMV time-course proteomics experiment 28

Figure 1.6 Plasma membrane profiling 37

CHAPTER 3: INTERFERON SCREEN

Figure 3.1 Experimental workflow for PM proteomic analysis of primary leukocytes . 84

Figure 3.2 Cell surface changes in THP-1 cells upon IFN stimulation..... 85

Figure 3.3 Comparison of methods for enrichment of primary CD14+ Monocytes by flow cytometry 87

Figure 3.4 Analysis of purity of primary monocytes and CD4+ T cells 88

Figure 3.5 Cell surface proteome of primary monocytes and T cells..... 91

Figure 3.6 Comparison of monocyte and T cell surface proteomes..... 92

Figure 3.7 Variation of cell surface proteomes between donors 93

Figure 3.8 IFN-induced changes at the surface of primary monocytes..... 97

Figure 3.9 IFN-induced changes in primary monocytes and cultured THP-1s..... 99

Figure 3.10 IFN induced changes at the surface of primary pan-monocytes..... 100

Figure 3.11 Comparison of proteins modulated by IFN in transcriptomic and proteomic data 102

Figure 3.12 IFN induced changes at the surface of primary CD4+ T cells..... 104

Figure 3.13 Overlap between IFN-stimulated proteins in monocytes and T cells 105

Figure 3.14 Comparison of proteomics data to transcriptomic meta-analysis 107

Figure 3.15 IFN-induced changes in individual donors of modulated proteins 109

Figure 3.16 Variation in IFN induced changes between donors 110

Figure 3.17 Validation of proteomic data on IFN stimulation of primary leukocytes . 111

Figure 3.18 Validation of IFN stimulation of TMEM123..... 112

Figure 3.19 shRNA knockdowns of TMEM123 113

Figure 3.20 Testing TMEM123 antibodies on shRNA knockdowns THP-1 cells by cell surface flow cytometry 115

Figure 3.21 Testing TMEM123 antibodies on shRNA knockdown THP-1 cells by intracellular flow cytometry 116

Figure 3.22 Testing TMEM123 Antibodies on shRNA knockdown THP-1 cells by Immunoblot 118

CHAPTER 4: VACCINIA VIRUS SCREEN

Figure 4.1 Schematic of workflow for proteomic analysis of VACV infection..... 129

Figure 4.2 Hierarchical clustering of VACV infection time-course 130

Figure 4.3 Quantitative temporal analysis of VACV infection 132

Figure 4.4 Downregulation of multiple collagens, protocadherins and innate immune mediators during VACV infection 134

Figure 4.5 Co-regulation of proteins by VACV and HCMV 136

Figure 4.6 Modulation of known and putative immune ligands by VACV infection .. 139

Figure 4.7 Definition of temporal classes of VACV gene expression 143

Figure 4.8 Comparison of viral protein temporal classes to previous studies..... 146

Figure 4.9 Schematic of MG132 experiment examining VACV induced protein degradation 147

Figure 4.10 Systematic analysis of proteasomal degradation of host proteins..... 149

Figure 4.11 Effect of MG132 treatment on viral protein expression 150

Figure 4.12 HDAC5 expression during VACV infection 151

Figure 4.13 HDAC5 was targeted by the VACV protein C6 153

Figure 4.14 HDAC5 restricted infection with VACV and HSV-1..... 155

CHAPTER 5: CANDIDATE ANTIVIRAL RESTRICTION FACTORS

Figure 5.1 Isoforms of ECE1 and gene organisation..... 171

Figure 5.2 ECE1 involvement in substance P degradation and receptor recycling..... 173

Figure 5.3 Validation of IFN stimulation of ECE1 176

Figure 5.4 ECE1 is downregulation during infection with HCMV..... 177

Figure 5.5 The HCMV proteins US2 is necessary and sufficient for degradation of ECE1 178

Figure 5.6 Testing ECE1 antibodies for cell surface flow cytometry on HFFF-TERTs 180

Figure 5.7 Testing ECE1 antibodies for intracellular flow cytometry on HFFF-TERTS 181

Figure 5.8 Immunoblot analysis of ECE1 antibodies..... 182

Figure 5.9 Schematic of HCMV restriction assay..... 183

Figure 5.10 Immunoblot confirming overexpression of ECE1 in HFFF-TERTs 184

Figure 5.11 HCMV restriction assay on HFFF-TERTs overexpressing ECE1..... 185

Figure 5.12 Validation of ECE1 knockdown in HFFF-TERTs by RT-qPCR..... 186

Figure 5.13 HCMV restriction assay on ECE1 shRNA knockdown HFFF-TERTs 187

Figure 5.14 ECE1 knockdown in HFFF-TERTs was diminished following passaging188

Figure 5.15 Immunoblot showing knockdown in polyclonal ECE1 CRISPR populations 189

Figure 5.16 HCMV restriction assay on ECE1 polyclonal CRISPR populations..... 190

Figure 5.17 Immunoblot for ECE1 in monoclonal CRISPR populations 191

Figure 5.18 HCMV restriction assay on monoclonal ECE1 CRISPR populations..... 192

Figure 5.19 HCMV restriction assay comparing percent infection between different monoclonal CRISPR control guide populations..... 194

Figure 5.20 Schematic of two colour restriction assay for shRNA knockdown cells.. 195

Figure 5.21 ECE1 cell lines for HCMV restriction assay in two colour system..... 196

Figure 5.22 Two colour restriction assays on ECE1 shRNA knockdown cells 198

Figure 5.23 Two colour restriction assay on ECE1 overexpressing cells 199

Figure 5.24 Immunoblot of shRNA and overexpressing cell lines 200

Figure 5.25 Repeat HCMV restriction assay on shRNA cell lines 201

Figure 5.26 Repeat HCMV restriction assay on overexpression cell lines 202

Figure 5.27 Plaque assay on ECE1 shRNA and overexpression cells 204

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Table 1. Summary of characteristic of a selection of well characterised antiviral restriction factors 8

CHAPTER 2: MATERIALS AND METHODS

Table 2.1 Target sequences for shRNAs 47

Table 2.2 CRISPR guide sequences 49

Table 2.3 Sequencing primers to confirm inserts in lentiviral expression vectors..... 51

Table 2.4 Plasmids for production of cell lines 52

Table 2.5 Table of primary antibodies used for flow cytometry experiments 54

Table 2.6 Table of secondary antibodies used for flow cytometry experiments..... 54

Table 2.7 Antibodies used for immunoblots 56

Table 2.8 Antibodies used for immunoblots in chapter 4 (vaccinia virus screen) 62

Table 2.9 Proteomic screens performed, and MS runs analysed in final dataset 63

Table 2.10 Sample labelling for PM proteomics samples..... 69

Table 2.11 Sample labelling for WCL proteomic samples 70

Table 2.12 Table of software used 79

CHAPTER 3: INTERFERON SCREEN

Table 3.1 Pre- and post- enrichment cell purity 89

Table 3.2 Number of proteins quantified in each cell type 89

Table 3.3 Proteins modulated by IFN in primary monocytes and T cells..... 95

Table 3.4 Proteins commonly modulated by IFN in monocytes and T cells..... 105

Table 3.5 Table of antibodies trialled for detection of TMEM123 114

CHAPTER 4: VACCINIA VIRUS SCREEN

Table 4.1 Table of peptides and proteins quantified in each proteomic experiment on VACV discussed in this chapter 129

Table 4.2 Number of host proteins modulated by VACV infection..... 131

Table 4.3 Candidate immunoreceptors modulated by VACV infection 140

CHAPTER 5: CANDIDATE ANTIVIRAL RESTRICTION FACTORS

Table 5.1 Effect of viral infection on proteins which were IFN stimulated in monocytes 167

Table 5.2 Effect of viral infection on proteins which were IFN stimulated in T cells . 168

Table 5.3 Table of antibodies trialled for ECE1 detection 179

Table 5.4 Table summarising results of ECE1 restriction and plaque assays 205

APPENDIX

Appendix Table 1. Oligonucleotides for generation of cell lines……248

Appendix Table 2. Proteins downregulated during VACV infection…….249

Appendix Table 3. Proteins upregulated during VACV infection……..256

Appendix Table 4. Proteomic and transcriptional assignment of VACV protein classes……...258

ABBREVIATIONS

2DE	2-dimensional gel electropohoresis
AcN	Acetonitrile
AGC	Automatic gain control
AIDS	Acquired immunodeficiency disease syndrome
AP	Activator protein
APC	Antigen presenting cell
APOBEC3	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3
APOC3	Apolipoprotein C3
AraC	Cytosine arabinoside
ARF	Antiviral restriction factor
ARID5A	AT-rich interactive domain-containing protein 5A
ATP1A1	Sodium/potassium-transporting ATPase subunit alpha-1
AU	Arbitrary units
BAC	Bacterial artificial chromosome
BCA	Bicinchoninic acid
BH	Benjamini-Hochberg
BST2	Bone marrow stromal cell antigen 2
C1R	Complement C1r subcomponent
C5AR1	C5a anaphylatoxin chemotactic receptor 1
Cas9	CRISPR-associated protein 9
CCR5	C-C chemokine receptor type 5
CCR7	C-C chemokine receptor type 7
cDNA	Complementary DNA
CDR2	Cerebellar degeneration-related protein 2
CEV	Cell-associated enveloped virus
cGAS	Cyclic GMP-AMP synthase
CGRP	Calcitonin gene related peptide
CID	Collision induced dissociation

CLEC12A	C-type lectin domain family 12 member A
CLR/RAMP	Calcitonin receptor-like receptor / receptor activity-modifying protein 1
CNP	C-type natriuretic peptide
COL1A2	Collagen alpha-2
COL6A2	Collagen alpha-2(VI) chain
COVID-19	Coronavirus disease 2019
CRISPR	Clustered regularly interspaced short palindromic repeat
Cul3	Cullin 3
CXCR4	C-X-C motif chemokine receptor 4
Daxx	Death domain-associated protein 6
DC	Dendritic cell
DE	Delayed early
DMEM	Dulbecco's Modified Eagle Medium
DNAM-1	DNAX accessory molecule-1
DNA-PK	DNA-dependent protein kinase
ds	Double stranded
DTT	Dithiothreitol
E	Early
EBV	Epstein-Barr virus
ECE1	Endothelin converting enzyme
EEV	Extracellular enveloped virus
EGR1	Early growth response protein 1
ELIZA	Enzyme-linked immunosorbent assay
Env	Envelope protein
EPHB2	Epherin receptor B2
ERK1/2	Extracellular signal-regulated protein kinase 1/2
ET	Endothelin
ETA	Endothelin receptor type A
ETB	Endothelin receptor type B
EV	Empty vector
EV12A	Ecotropic viral integration site 2A protein homolog

FA	Formic acid
FBS	Foetal bovine serum
FC	Fold change
FCM	Flow cytometry
FDR	False discovery rate
Gag	Group specific antigens
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GAW	Control gateway plasmid
GFP	Green fluorescent protein
GLUT	Glucose transporter
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GO	Gene ontology
GOPC	Golgi associated PDZ and coiled-coil motif containing protein
gp120	Envelope glycoprotein 120
gp41	Glycoprotein 41
GPA33	Cell surface A33 antigen
GLP-1	Glucagon-like peptide 1
GPR171	Probable G-protein coupled receptor 171
gRNA	Guide RNA
H5KO	HDAC5 knockout
HCD	High energy CID
HCMV	Human cytomegalovirus
HCV	Hepatitis C virus
HDAC4	Histone deacetylase 4
HDAC5	Histone deacetylase 5
HEK293T	Human emrbyonic kidney cells
HFFF	Human foetal foreskin fibroblasts
HFFF-TERT	HFFFs immortalised with human telomerase
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
hpi	Hours post infection

HpRp	High pH reversed-phase
HPV	Human papillomavirus
HSV	Herpes simplex virus
I	Intermediate
IAA	Iodoacetamide
iBAQ	Intensity based absolute quantification
ICAM	Intracellular adhesion molecule
IE	Immediate early
IEV	Intracellular enveloped virus
IFIT	Interferon-induced proteins with tetratricopeptide repeats
IFITM	Interferon induced transmembrane proteins
IFN	Interferon
IFNAR	IFNα/β receptor
IFNGR	IFNγ-receptor
Ig	Immunoglobulin
IL	Interleukin
IL1RN	IL-1 receptor antagonist
IL2RA	Interleukin 2 receptor subunit alpha
IL6R	IL-6 receptor
IMV	Intracellular mature virus
IRF	IFN regulatory factor
iRFP	Near-infrared fluorescent protein
ISG	IFN-stimulated gene
ISGF3	ISG factor 3
ISRE	IFN-stimulated response element
ITGAL	Integrin alpha-L
ITGAX	Integrin alpha-X
ITGB2	Integrin beta-2
JAK	Janus kinase
KCNA3	Potassium voltage-gated channel subfamily A member 3
KSHV	Kaposi sarcoma-associated herpesvirus

L	Late
LAIR-1	Leukocyte-associated Ig-like receptor-1
LB	Lysogeny broth
LC	Liquid chromatography
LFA-1	Leukocyte function-associated antigen-1
Lir-1	Leukocyte immunoglobulin-like receptor 1
LMP1	Latent membrane protein-1
LPS	Lipopolysaccharide
LRRC25	Leucine rich repeat-containing 25
m/z	Mass to charge
MAPK	Mitogen activated protein kinase
MAVS	Mitochondrial antiviral-signalling protein
MCP1	Monocyte chemotactic protein-1
MDA5	Melanoma differentiation-associated gene 5
MEF2	Myocyte enhancer factor 2
MEK	Dual specificity MAPK-kinase
MEM	Minimum essential media
MERS	Middle East respiratory syndrome
MHC	Major histocompatibility complex
MICA	MHC class I polypeptide-related sequence A
MOI	Multiplicity of infection
MS	Mass spectrometry
MS3	Triple-stage mass spectrometry
MYD88	Myeloid differentiation factor 88
NCE	Normalised collision energy
ND10	Nuclear domain-10
Nef	Negative factor
NF-kB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NHSBT	NHS blood and transplant
NK	Natural killer
NK1R	Neurokinin 1 receptor

NOD	Nucleotide-binding and oligomerization domain
NPTN	Neuroplastin
ns	Not significant
NS1	Non-structural protein 1
Nur77	Nuclear receptor subfamily 4 group A member 1
P2A	Porcine teschovirus-1 2A
PAGE	Polyacrylamide gel electrophoresis
PAMP	Pathogen associated molecular pattern
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PD-L1	Programmed Death Ligand–1
pen/strep	Penicillin / streptomycin
PM	Plasma membrane
PML	Promyelocytic leukaemia protein
Pol	Polymerase
PP2A	Protein phosphatase 2A
ppm	Parts per million
PR	Post-replicative
PRR	Pattern recognition receptor
PTPRC	Receptor-type tyrosine-protein phosphatase C
PVDF	Polyvinylidene difluoride
RA	Rheumatoid arthritis
RANTES	Regulated on activation, normal T cell expressed and secreted
RBX1	RING-box protein 1
RIG-I	Retinoic acid-inducible gene-I
RIPA	Radioimmunoprecipitation assay
RNA-seq	RNA sequencing
RNF149	RING finger protein 149
RT-qPCR	Reverse transcription-quantitative PCR
S:N	Signal-to-noise

SAMHD1	SAM domain and HA domain-containing protein 1
SARS	Severe acute respiratory syndrome
SBSN	Suprabasin
SCX	Solid-phase cation exchange
SDC	Syndecan
SDS	Sodium dodecyl sulfate
SD	Standard deviation
SELL	L-selectin
SEM	Standard error of the mean
SERINC	Probable serine incorporator
shRNA	Short hairpin RNA
SIGLEC1	Sialic acid binding Ig-like lectin 1
SILAC	Stable isotope labelling with amino acids in cell culture
SIRPA	Signal regulatory protein alpha
SLC25A2	ADP/ATP translocase 2
SLC2A3	Solute family carrier 2, facilitated glucose transporter member 3
SLC4A7	Sodium bicarbonate cotransporter 3
SLC5A3	Sodium/myo-inositol cotransporter
SLE	Systemic lupus erythematosus
SOCS	Suppressor of cytokine signalling
SP	Substance P
Sp100	Nuclear antigen Sp-100
SPN	Leukosialin
SSTR2	Somatostatin receptor type 2
STAT	Signal transducer and activator of transcription
TAT	Trans-activator of transcription
TBK1	TANK-binding kinase
TBST	Tris-buffered saline with tween
tet	Tetracycline
Th	Thomson
TLR	Toll-like receptor

TMEM123	Transmembrane prtotein 123
TMT	Tandem mass tag
TNF	Tumour necrosis factor
TNFRSF10B	TNF receptor superfamily member 10B
Tp	Temporal profile
Tp53	Cellular tumour antigen p53
TRAM	Translocating chain-associated membrane protein
TRIF	Toll/IL-1R resistance domain–containing adapter-inducing IFN-β
TRIM	Tripartite motif-containing
TWEAK	TNF-like weak inducer of apoptosis
TYK2	Tyrosine kinase 2
ULBP2	UL16 binding protein 2
USP18	Ubiquitin-specific peptidase 18
v/v	Volume per volume
VACV	Vaccinia virus
Vif	Viral infectivity factor
Vpu	Viral protein U
Vpx	Viral protein X
VSVg	Vesicular stomatitis virus glycoprotein G
VZV	Varicella-zoster virus
w/v	Weight per volume
WCL	Whole cell lysate
WHO	World heath organisation
WR	Western Reserve
WT	Wild type
XRCC5	X-ray repair cross-complementing protein 5
XRN1	5'-3' exoribonuclease 1

CHAPTER 1: INTRODUCTION

1.1 Introduction

Until the end of the twentieth century, infectious diseases were the biggest global cause of premature death. Since then, this burden has dramatically reduced due to improved living conditions, astounding achievements in science, and availability of antibiotics and vaccines. Scientific advances and large-scale campaigns by the World Health Organisation (WHO) have resulted in the eradication of smallpox, a viral disease responsible for 3-500 million death, near eradication of polio, and vaccination programs for diseases such as measles and rubella. However, the need for novel antivirals remains. Globally, there are still 3-5 million cases of severe illness caused by seasonal influenza annually, including 290-650,000 deaths. From 2010 to 2018 there was an impressive 40 % decrease in mortality due to human immunodeficiency virus (HIV), however there were still 1.7 million new cases of HIV and 770,000 deaths. Additionally, emerging viral infections present a constant threat, with recent outbreaks and pandemics including severe acute respiratory syndrome (SARS), Middle East respiratory syndrome (MERS), swine flu, Ebola, Zika and most recently coronavirus disease 2019 (COVID-19). There are very few antivirals available for these, and they pose a particular risk in the resource limited settings of developing countries. A complete understanding of the immune response to viral infection, as well as characterising the viruses themselves, is therefore essential (Holmes *et al*, 2017; Madhav *et al*, 2017; World Health Organisation).

The 21st century has also seen the emergence of the field of systems biology which emphasises the importance of taking a holistic approach to research, and integrating multiple sources of data to generate a comprehensive understanding of biological organisation. With the development of this field has come an array of 'omic' technologies, to study the genome, transcriptome, metabolome and proteome. Given the complex processes and interactions within the cell, being able to examine it as a whole is invaluable, but with these investigations comes ever larger datasets, and the need to compare, filter and analyse them in order to identify biologically interesting facets of the data (Kirschner, 2005; Horgan & Kenny, 2011).

The work presented in this thesis uses proteomic technologies to develop understanding of viral infection and the immune response. Antiviral restriction factors (ARFs) offer the first line of defence against viral infection. These are a diverse set of proteins, able to inhibit infection with different viruses at various stages of the life cycle (Duggal & Emerman, 2012).

Although the exact definition of an ARF remains debated (see Introduction 1.2.1) many criteria have been suggested. Here, proteomic technologies are applied to explore two characteristic traits of ARFs: stimulation by interferon (IFN), and downregulation by viral infection. IFN stimulation is considered largely in the context of the cell surface of primary monocytes and T cells, whilst viral infection focuses on vaccinia virus (VACV), as well as investigating previously published data on human cytomegalovirus (HCMV) and human immunodeficiency virus (HIV refers to HIV-1 unless otherwise stated) (Weekes *et al*, 2014; Matheson *et al*, 2015; Greenwood *et al*, 2016). From this proteomic screen, putative restriction factors are highlighted and further examined; in particular, endothelin converting enzyme (ECE1) is investigated as a candidate HCMV restriction factor due to being stimulated by IFN and targeted for degradation by HCMV.

1.2 Aim: Identify Candidate Antiviral Restriction Factors

1.2.1 Defining Antiviral Restriction Factors

Restriction factors act as the first line of defence against viral infection. There is a huge variety of ARFs, each acting in distinct cell types to protect against diverse viruses at different stages of infection. Consequently, there is not a single unambiguous definition for 'restriction factors' and which proteins this encompasses. Historically, the term was used when describing research on the Fv1 locus, which is implicated in the differing susceptibility of various strains of mice to murine leukemia virus (Pincus *et al*, 1971). It has since been adopted for use more broadly. For some, a restriction factor describes any protein with antiviral function, whilst other scientists apply a more stringent definition and have attempted to define a set of criteria. Duggal *et al* describes four main criteria for a restriction factor (Duggal & Emerman, 2012):

1. Host proteins whose main function is antiviral
2. Induced by IFN or viral infection
3. Antagonised by viral factors
4. Evolves under positive selection

These four broad characteristics describe common features of restriction factors. The proteins are usually IFN-stimulated, indicating their primary function is in defending against pathogens. Viral antagonists use several mechanisms to overcome restriction factors. These may include protein degradation, mislocalisation, or mimicry of the restriction factor substrate to prevent function. Some viruses however, have mutated the viral target in order to escape restriction. This mutation has led to an 'evolutionary arms race', with viruses and restriction factors under constant pressure to mutate. This results in an unusually high rate of non-synonymous mutations in the restriction factor, leaving genetic signatures of positive selection (Duggal & Emerman, 2012).

However, many restriction factors do not fulfil all of these criteria. For example, the IFN-induced transmembrane proteins (IFITMs) restrict multiple viruses but a direct viral antagonist has not been identified (Brass *et al*, 2009; Zhao *et al*, 2019). Probable serine incorporator (SERINC) 3 and SERINC5 are specifically antagonised, but do not show evidence of an 'evolutionary arms race' by being under positive selection (Murrell *et al*, 2016), and they are not IFN-stimulated (Rosa *et al*, 2015). Some researchers have coined additional terms such as 'resistance factors' to describe antiviral proteins for which a viral antagonist has not yet been identified (Doyle *et al*, 2015).

Although the exact definition of an ARF remains debated, in this thesis the term will refer to all proteins which limit viral infection. The discussed criteria provide a valuable basis for the identification of candidate restriction factors, and this investigation examines two of these: stimulation by IFN and downregulation by a viral factor. There is also a particular focus on the plasma membrane (PM), as proteins here are well situated to restrict viral entry or exit to the cell.

1.2.2 Known Restriction Factors

Cell Surface Restriction Factors

There are several important restriction factors which are known to act at cell membranes, preventing viral entry or exit. Perhaps one of the best known ARFs is the IFN-induced protein Tetherin (bone marrow stromal cell antigen 2, BST2), which restricts multiple enveloped viruses (le Tortorec *et al*, 2011). It was first discovered in the context of HIV infection, when a HIV viral protein U (Vpu) deletion virus was found to be retained within the cell due to Vpu being responsible for antagonising an antiviral factor. This antiviral factor was determined to be BST2 (Neil *et al*, 2008). Immunoelectron microscopy later showed BST2 physically tethering HIV particles to the cell surface and preventing viral exit from the cell (Hammonds *et al*, 2010; Fitzpatrick *et al*, 2010). More recently, roles for BST2 have been reported in immune sensing via activation of the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) pathway (Hotter *et al*, 2013). Several proposed mechanisms for the antagonism for BST2 by HIV suggest that Vpu and BST2 co-localise, and that Vpu reduces the level of BST2 at the cell surface. Sequestration inside the cell may be through impaired secretion, or internalisation. Vpu expression also leads to reduced overall levels of BST2, and this is prevented by proteasome inhibitors, indicating some degradation (Malim & Bieniasz, 2012).

IFITMs are other examples of membrane proteins that restrict a range of viruses, including influenza A virus, dengue virus and Ebola virus. IFITM1, IFITM2 and IFITM3 are the best studied, however the antiviral mechanism through which they act remains unclear. IFITM1 is predominantly expressed at the cell surface, whilst IFITM2 and IFITM3 are largely intracellular, with IFITM3 being localised to endosomes; however, their specific localisation depends on the cell type and expression level. It is thought that the IFITMs may alter membrane fluidity and restrict membrane fusion at various points during viral entry into the cell, either at the PM or in endosomes. This may prevent certain viruses from entering the cytosol, instead resulting in viral lysosomal degradation. A small number of enveloped viruses are resistant to the effects of IFITMs, such as Sendai virus which fuses at the cell surface. However, a viral antagonist has not yet been identified (Perreira *et al*, 2013; Smith *et al*, 2014).

SERINC3 and SERINC5 prevent viral fusion, restricting infection with HIV. The mechanism of this has not yet been determined, but may involve interaction with the viral envelope protein (Env) trimer, or modification of viral lipids to inhibit fusion. HIV negative factor (Nef) antagonises SERINC5 by removing it from the PM and sequestering it in endosomes (Ramirez *et al*, 2019; Gonzalez-Enriquez *et al*, 2017).

Other Restriction Factors

There are many other restriction factors, which are diverse in their ability to restrict various viruses at differing stages of infection. Some well characterised examples of ARFs that restrict multiple viruses are summarised in Table 1. As discussed, some act at the cell surface, whilst many others function within the cell. Sterile alpha motif domain and histidine-aspartic domain-containing protein 1 (SAMHD1) is expressed in macrophages and dendritic cells (DCs), and is another HIV restriction factor. Interestingly, rather than acting on a viral component, SAMHD1 hydrolyses the intracellular pool of the host's own deoxynucleotide triphosphates, reducing the level to below that required for the viral reverse transcriptase to function (Lahouassa *et al*, 2013). HIV-2 has evolved the accessory protein viral protein x (Vpx) which results in the ubiquitination and degradation of SAMHD1 (Hrecka *et al*, 2011). In contrast, HIV-1 does not have Vpx and this may play a role in its limited ability to infect monocyte-derived macrophages and DCs. SAMHD1 may restrict replication of other DNA viruses similarly, including herpes simplex virus (HSV) (Kim *et al*, 2013). The role of SAMHD1 is also important in regulating the innate immune response, with mutations found in patients suffering from the autoimmune disease Aicardi-Goutières syndrome (Rice *et al*, 2009).

Other key HIV restriction factors include tripartite motif-containing (TRIM) 5α protein, which acts via binding to the viral capsid. The mechanism has not been completely characterised, but is thought to involve destabilising the capsid and could be ubiquitin-dependent. Whilst rhesus TRIM5α is a potent restriction factor for HIV, the role of human TRIM5α in HIV infection is more complex and less potent (Ganser-Pornillos & Pornillos, 2019). The apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3 (APOBEC3) family causes cytidine deamination of the viral genome, leading to mutations and reduced infectivity.

APOBEC3 restricts HCMV infection in a similar manner. There are also a plethora of other restriction factors for HCMV (some of these are reviewed in (Biolatti *et al*, 2018)). IFI16 impairs viral DNA synthesis (Gariano *et al*, 2012) and is antagonised by the HCMV proteins UL83 and UL97. Viperin inhibits expression of late viral genes whilst HCMV UL37 actually co-opts viperin to enhance viral infection. The components of the nuclear domain-10 (ND10) complex restrict HCMV and are discussed further in section 1.4.2.

1.2.3 Therapeutic Potential of Restriction Factors

Identification of ARFs not only improves our understanding of a virus and the immune system, but could also provide targets for antivirals. Several ways in which the potential of ARFs can be harnessed for therapies has been explored in the context of HIV infection (Colomer-Lluch *et al*, 2018).

Many restriction factors are IFN-stimulated, therefore IFN treatment to upregulate ARFs may have therapeutic potential for HIV. IFN therapy with pegylated-IFNα2a or IFNα2b, usually in combination with Ribavirin, is already being used to treat infection with hepatitis C virus (HCV). Investigation of IFN treated patients co-infected with HCV and HIV identified suppression of the HIV viral load, correlating with IFN-induced upregulation of ARFs (Abdel-Mohsen *et al*, 2014). The use of IFNα2 is being investigated for treatment of HIV in clinical trials. However, prolonged IFN expression in chronic HIV infection has been observed to be detrimental, contributing to disease progression; blocking the IFN response has also been suggested to have therapeutic value (Wang *et al*, 2017). Additionally, IFN therapies have multiple side effects, and although they have been used for years, they have now been largely superseded by more modern alternatives. A more specific approach utilising ARFs may therefore be preferable.

Another approach could involve disruption of the interaction between an ARF and its viral antagonist, enabling the cells' own intrinsic immune system to respond specifically to the infection with limited side effects. A small molecule inhibitor of the HIV viral infectivity factor (Vif) was found to lead to increased incorporation of APOBEC3G into the virions (Nathans *et al*, 2008). Attempts have also been made to identify inhibitors of the BST2 antagonist, Vpu (Zhang *et al*, 2011). Challenges with this approach include

ensuring the inhibitor only affects the viral antagonist, particularly if it mimics a host protein, and not disrupting any interaction between the ARF and the virus required for restriction. Additionally, mutations may allow the virus to develop resistance.

Gene therapy is another avenue for exploiting ARFs in therapeutics. Unlike rhesus TRIM5α, human TRIM5α does not usually potently restrict HIV. However, mutations and fusions of TRIM5α that improve HIV capsid binding, can improve restriction (Yap *et al*, 2005; Neagu *et al*, 2009). Additionally, it does not have a viral antagonist, and was therefore recognised as a potential candidate for gene therapy. A rhesus/human chimeric TRIM5α was included in a combination lentiviral therapy which has shown promising results in a murine model and is in clinical trials (Anderson *et al*, 2009; Walker *et al*, 2012).

Table 1. Summary of characteristic of a selection of well characterised antiviral restriction factors

Information is adapted from several reviews (Duggal & Emerman, 2012; Kluge *et al*, 2015; Biolatti *et al*, 2018), with some additional examples and antagonists from the indicated sources. '-' indicates that this information was not available in published literature.

Restriction Factor	IFN Induced?	Viruses Targeted	Viral Lifecycle Stage Inhibited	Viral Antagonists	Under Positive Selection?	Reference
IFITM1-3	Yes	Orthomyxo-, flavi-, coronaviruses	Endosomal fusion or uncoating, membrane fluidity	None known	Some	(Duggal & Emerman, 2012; Kluge *et al*, 2015)
SERINC3,5	No	Retroviruses	Membrane fusion	Nef (HIV, SIV), glyco-gag (MLV)	No	(Kluge *et al*, 2015)
Fv1	No	Retroviruses (MLV)	Capsid uncoating	None known (though do see capsid mutations)	Yes	(Duggal & Emerman, 2012; Kluge *et al*, 2015)
TRIM5a and TRIM-CYP	Yes	Retroviruses	Capsid uncoating	None known (though do see capsid mutations)	Yes	(Duggal & Emerman, 2012; Kluge *et al*, 2015)
MxB	Yes	Retroviruses (HIV, SIV)	Uncoating, nuclear uptake and integration	None known (though do see capside mutations)	Yes	(Kluge *et al*, 2015)
MxA and Mx1	Yes	Orthomyxo-, paramyxo-, hepadna-, rhabdo-, alpha-, bunya-, toga-, picornaviruses	Nucleocapsid transport and another early life cycle step	None known	Yes	(Duggal & Emerman, 2012; Kluge *et al*, 2015)
APOBEC3 family	Some	Retro-, hepadna-, herpes-, papillomaviruses	(Reverse) transcription, causes hypermutation	Vif (lentiviruses), Bet (spumaviruses), Gag (gammaretroviruses)	APOBEC3DE, APOBEC3G, APOBEC3H	(Duggal & Emerman, 2012; Kluge *et al*, 2015)
SAMHD1	Yes	Retro-, pox-, herpes-, arteriviruses	(Reverse) transcription, hydrolyses cellular dNTPs	Vpx, Vpr (some SIVs), pUL97 (HCMV)	Yes	(Duggal & Emerman, 2012; Kluge *et al*, 2015; Businger *et al*, 2019)

Restriction Factor	IFN Induced?	Viruses Targeted	Viral Lifecycle Stage Inhibited	Viral Antagonists	Under Positive Selection?	Reference
IFI16	Yes	Herpesviruses	Viral transcription	UL97 and UL83 (HCMV)	Yes	(Kluge *et al*, 2015; Biolatti *et al*, 2018)
TRIM28	No	Retro-, herpesviruses	Transcription, induces latency	vPK (KSHV)	No	(Kluge *et al*, 2015)
ZAP	Yes	Retro-, filo-, alphaviruses	Viral protein translation	None known	Yes	(Duggal & Emerman, 2012; Kluge *et al*, 2015)
IFIT family	Yes	Flavi-, bunya-, rhabdo-, orthomyxo-, picorna-, corona-, poxviruses	Translation	methylation of viral RNA (VACV, SARS), masking of 5' end by Vpg (EMCV), hairpin at 5' end (VEEV), C9 (VACV)	Yes	(Kluge *et al*, 2015; Liu *et al*, 2019)
PKR	Yes	Poxviruses	Translation	K3L and E3L (VACV), TRS1 and IRS1 (HCMV) and many others	Yes	(Duggal & Emerman, 2012; Kluge *et al*, 2015)
SLFN11	Yes	Retroviruses	Translation	None known	Yes	(Kluge *et al*, 2015)
ND10 complex (Daxx, PML, Sp100)	PML and Sp100	HCMV	Inhibits early viral gene expression	IE1, UL82 (HCMV)	-	(Biolatti *et al*, 2018; Ashley *et al*, 2017)
HLTF	-	HCMV	Inhibits early viral gene expression	UL145 (HCMV)	-	(Nightingale *et al*, 2018)
Tetherin	Yes	Retro-, flavi-, herpes-, rhabdo-, paramyxo-, arenaviruses	Budding	Nef (some SIVs, HIV-1), Vpu (HIV-1), Env (HIV-2), glycoprotein (Ebola virus), K5 (KSHV)	Yes	(Duggal & Emerman, 2012; Kluge *et al*, 2015)
Viperin	Yes	Orthomyxo-, flavi-, herpes-, alpha-, paramyxoviruses	Budding	UL37 (HCMV)	Yes	(Duggal & Emerman, 2012; Kluge *et al*, 2015; Biolatti *et al*, 2018)

1.3 Screen 1: Interferon

The term IFN was first used in 1957 to describe a substance that interfered with viral infection (Isaacs & Lindenmann, 1957). IFN has since become known as a key component of the innate immune system, triggering a signalling pathway that culminates in the production of an array of IFN-stimulated genes (ISGs) with diverse functions involved in protecting against pathogens. The critical role of IFN in protecting against viruses *in vivo* has been known for some time, with IFNα/β receptor (IFNAR) knockout mice being more susceptible to a range of viral infections (Müller *et al*, 1994). The importance of IFN is further emphasised by the multitude of mechanisms employed by viruses to evade the IFN response, often requiring a substantial part of a limited viral genome; some of these are discussed further in section 1.4.

In this thesis, proteomic studies are used to investigate the IFN response at the cell surface, for the purpose of identifying novel candidate ARFs. Moreover, the dataset is also of value more broadly; a detailed characterisation of the IFN response in a variety of cell types is essential for a complete understanding of the immune response, may provide insights into IFN-related autoimmune conditions such as systemic lupus erythematosus (SLE), and is critical for the safe and effective use of IFN in therapeutics. IFN is currently being used for treatment of viral infections such as HCV, and is being developed for use in cancer therapies. However, IFN therapy is often accompanied by severe side effects. These may include flu-like symptoms, nausea, fatigue, psychiatric effects and a range of autoimmune diseases. These adverse effects can be dose-limiting or even prevent therapy. A deeper understanding of how IFN responses vary between cell types and patients is necessary to enable identification of patients who are more likely to benefit from treatment, or more targeted therapies (Sleijfer *et al*, 2005).

1.3.1 Types of Interferon

IFNs fall into three broad categories: type I (in humans these include 13 subtypes of IFNα, as well as β, ε, κ, ω), type II (IFNγ), and the most recently identified type III IFNs (IFNλ) (Schneider *et al*, 2014). Type I IFNs are the largest class and the most extensively studied.

They signal through the IFNAR, which is composed of the IFNAR1 and IFNAR2 subunits. The type I IFN subtypes have different binding affinities for the receptor and varying tissue expression patterns. IFNα and IFNβ are produced by most cells in response to viral infection and are the main type I IFN subtypes of interest. There is just one gene for IFNβ, and 13 genes encoding different subtypes of IFNα, all on chromosome 9. These IFNα subtypes all have a similar gene structure and a conserved protein sequence. However, they have differing biological activity and result in different patterns of ISGs being expressed. These differences may be explained by a combination of different affinities for the IFNAR and the stability of the complex formed with the receptor, cell-type specific responses, and variation dependent on the type of infecting virus that is inducing IFN expression (Gibbert *et al*, 2013).

IFNγ is the only type II IFN, and it acts via the IFNγ-receptor (IFNGR). These receptors are expressed on a range of cell types, so despite IFNγ being expressed largely by cells of the immune system, many cells can respond. Type III IFNs are the most recently discovered, and encompass the four types of IFNλ. There are four types of IFN-λ, though it is controversial as to how functional the most recently discovered IFNλ4 is (Hong MeeAe *et al*, 2016). They signal through distinct receptors to the other types of IFN. These receptors are expressed on epithelial cells and the type III IFNs are thought to restrict a subset of epitheliotropic viruses (Pott *et al*, 2011).

This project focusses on the effects of stimulation with IFNα2a. IFNαs have been well characterised and play a crucial role in the immune response to viruses. IFNα2 was one of the first IFNs to be cloned in the 1980s, and has since been the subject of much research; it is therefore one of the best characterised type I IFNs (Paul *et al*, 2015). For this reason, and due to its clinical relevance, IFNα2a is used in this investigation. It was approved in a modified form (pegylated-IFN) in 2002, by the US Food and Drug Administration, for use in treatment of HCV infection. It is also being investigated for use in treatment of hepatitis B virus infection as well as cancers, including metastatic malignant melanoma (Matthews & McCoy, 2004).

1.3.2 Type I interferon signalling pathways

Induction of IFN Expression

Following detection of pathogens via pattern recognition receptors (PRRs), an antiviral state is induced in the infected cell and neighbouring cells by production of type I IFNs. PRRs include toll-like receptors (TLRs), retinoic acid-inducible gene-I (RIG-I)-like receptors, and nucleotide-binding and oligomerization domain (NOD)-like receptors, with different PRRs capable of recognising different classes of pathogen associated molecular patterns (PAMPs). TLR4, for example, detects bacterial lipopolysaccharide (LPS) at the PM or in endosomes, whilst TLR3 detects double stranded (ds) RNA in endosomes. The different PRRs utilise diverse signalling cascades upon activation. RIG-I and melanoma differentiation-associated gene 5 (MDA5) for example, rely on the adaptor mitochondrial antiviral-signalling protein (MAVS), which activates TANK-binding kinase 1 (TBK1), leading to phosphorylation of IFN regulatory factor (IRF) 3. TLR3 and TLR4 use the toll/IL-1R resistance domain–containing adaptor-inducing IFNβ (TRIF) adaptor to activate TBK1 for IRF3 phosphorylation, whilst TLR7 and TLR9 are predominantly expressed in plasmacytoid DCs, and signal via myeloid differentiation factor 88 (MYD88) to activate the IRFs. Most of these diverse pathways culminate in the translocation of the transcription factor IRF3 (or sometimes IRF7) into the nucleus. This triggers an initial wave of transcription, where IFNβ and IFNα4 are produced, as well as IRF7. This then generates a positive feedback loop and a second wave of transcription, resulting in the production of other IFNα genes (McNab *et al*, 2015). As well as the nuclear translocation of IRF3 and IRF7, signalling is also via NF-κB and activator protein (AP)-1. These associate with IRFs and other proteins in the nucleus to form an 'enhancersome' that enables transcription of IFNβ (Agalioti *et al*, 2000; Honda *et al*, 2006) (Figure 1.1).

IFN Induced Signalling

The IFN produced can then trigger autocrine or paracrine signalling via the heterodimeric type IFNAR. This receptor is ubiquitously expressed and found on the surface of most cell types (De Weerd & Nguyen, 2012). Signalling is via a diverse set of pathways, but the canonical signalling pathway involves the activation of Janus kinase (JAK)-1 and

tyrosine kinase 2 (TYK2). These phosphorylate signal transducer and activator of transcription (STAT)-1 and STAT2, which heterodimerise, translocate to the nucleus and bind IRF9. Together, these constitute the ISG factor 3 (ISGF3) complex, which binds IFN-stimulated response elements (ISREs) and triggers transcription of ISGs. As well as the induction of ISGs by the STAT1/STAT2/IRF9 heterotrimer, the IFNs can signal through a range of other molecules. Signalling may occur through STAT1 homodimers, though this is more commonly associated with IFNγ, through other STAT proteins, or via the mitogen activated protein kinase (MAPK) pathway (Figure 1.1). This contributes to the diversity of pathways activated by IFN, enabling it to have a broad, and cell type dependent, effect on the innate and adaptive immune response (McNab *et al*, 2015).

Figure of IFN induction and signalling pathways removed for copyright reasons. Copyright holder is Springer Nature BV.

Figure 1.1 Type I IFN induction and signalling pathways

A variety of pattern recognition receptors are able to recognise infection and generate a type I IFN response. Once induced, IFN can act through a range of pathways to trigger the production of ISGs with antiviral functions. Figure adapted from (McNab *et al*, 2015).

1.3.3 Effects of Interferon

Interferon Stimulates Genes

IFN signalling leads to the transcription of ISGs. These encode a diverse set of proteins with a multitude of functions (Schneider *et al*, 2014). Some ISGs are involved regulating the IFN response (Ivashkiv & Donlin, 2014). This can be positive feedback, amplifying

the IFN response; many IRFs and PRRs are IFN-stimulated. STAT1 and IRF9 for example are IFN-stimulated, augmenting the response. Often, these proteins are present at quite low abundance, potentially to avoid aberrant signalling, however upon detection of a pathogen and IFN signalling they increase in abundance and enhance the sensitivity of the cell for detecting infection. Other ISGs have an opposing effect of desensitising the cell to IFN, for example suppressor of cytokine signalling (SOCS) proteins inhibit binding of STAT to JAK, and ubiquitin specific peptidase 18 (USP18) interacts with IFNAR2 to prevent binding of JAK. Finally, other ISGs have direct antiviral effects. These include many of the ARFs discussed previously, such as IFITMs, APOBECs, TRIMs and BST2 (see section 1.2).

Activation of the Immune Response

As well as the cell intrinsic effects discussed, IFNs are able to impact the innate and adaptive immune response, with diverse effects dependent on context. Type I IFNs can inhibit or promote the differentiation of precursor cell types into DCs, and may inhibit or activate DCs. Amongst other effects on DCs, IFNs enhance expression of co-stimulatory molecules and migration to lymph nodes, therefore promoting T cell activation. The effect of IFN stimulation on T cells is again complex and context dependent, altering survival, proliferation and cytokine production. The variable effects are thought to be due to differential activation of STATs and signalling pathways following activation of the IFNAR. Type I IFNs enhance survival and B cell help by CD4+ T cells in response to viral infection, but not bacterial infection. In West Nile virus infection, IFNs control differentiation of regulatory CD4+ T cells. In CD8+ T cells, type I IFNs can exert STAT1 dependent anti-proliferative effects, inhibiting growth. However, IFN can promote survival and clonal expansion of CD8+ T cells during infection; this may involve downregulation of STAT1. Type I IFNs are also involved in enhancing natural killer (NK) and B cell responses. Though again, this may be context and timing dependent. Some studies found that IFN signalling enhances the antibody response of B cells, whilst others have observed higher viral titres at late time point in infection in IFNAR deficient mice compared to controls (McNab *et al*, 2015).

1.3.4 IFN Responses in Leukocytes

Cell Type Variability

IFN responses are well documented as being context dependent. Different cell types have constitutive expression of varying components of the IFN signalling pathway, or different basal levels of these proteins. In some cells, priming with other cytokines is able to modify the IFN response; for example priming with IFNγ increases the concentration of the ISGF3 components, and in monocytes and DCs it induces expression of STAT4. There are differences in activation of STAT proteins between cell types, and varying induction of the STAT independent signalling pathways. Differential activation of STAT1, STAT3 and STAT5 was previously observed across T cells, B cells and monocytes by IFNβ. For example, increased activation of STAT3 and STAT5 was observed in B cells and CD4+ T cells compared to monocytes, and very little activation of STAT1 was seen in B cells (van Boxel-Dezaire *et al*, 2006, 2010). In NK cells, STAT4 is initially activated, leading to production of IFNγ. As levels of STAT1 increase, this becomes predominant, and IFNγ production is replaced by NK cell killing (Mack *et al*, 2011).

Primary Monocytes and CD4+ T cells

Given the variability in IFN response between cell types, it was necessary to focus this investigation on particular cells, and primary monocytes and CD4+ T cells were chosen. These leukocytes both originate from hematopoietic stem cells in the bone marrow, which differentiate into myeloid or lymphoid precursors. The myeloid precursor goes on to further differentiate into various cell types, including monocytes and neutrophils, whilst the lymphoid precursors differentiate into B cells, T cells and NK cells (Figure 1.2).

Monocytes have several roles in the immune system, including phagocytosis of pathogens, antigen presentation and production of cytokines. They can be classified into classical (CD14++, CD16-), intermediate (CD14++, CD16+) and non-classical (CD14+, CD16++) subsets, with ~85% of the population consisting of classical monocytes. Whilst all subsets differentiate into macrophages, classical monocytes provide the principal source of monocyte-derived DCs. They also demonstrate the greatest phagocytic activity

and secretion of cytokines (Ziegler-Heitbrock *et al*, 2010; Boyette *et al*, 2017; Ravenhill *et al*, 2020).

T cells are derived from the lymphoid precursor, maturing in the thymus into CD4+ helper T cells, or CD8+ cytotoxic T cells. These are activated by interaction with an antigen presenting cell (APC) expressing a co-stimulatory protein, and foreign antigen bound to major histocompatibility complex (MHC). Activation of CD8+ T cells leads to target cell killing, whilst naïve CD4+ T cells go on to differentiate into further subsets of effector T cells, which secrete cytokines and activate other innate and adaptive immune cells. Helper T cells activate B cells for antibody secretion, macrophages for phagocytosis, and naïve CD8+ T cells to become effector cells, as well as developing into memory T cells (Lanzavecchia & Sallusto, 2000; Alberts, 2015).

These cell types were chosen as they provide accessible primary cells with relevance in viral infection, and they respond to IFN, though there is little previous information on the IFN-induced effects at the cell surface. Analysis of primary cells is important in this context, as not all cultured cell lines express the same restriction factors. This is demonstrated by the HIV accessory proteins, which may be dispensable in cultured systems, but are important for *in vivo* replication (Malim & Emerman, 2008). The search for BST2 demonstrated the effect of expression of different ARFs between cell types, with Vpu being required for viral release in HeLa cells, but not in in 293T and HT1080 cells unless they had been stimulated with IFNα (Neil *et al*, 2008). Additionally, isolation from peripheral blood makes monocytes and T cells easily accessible, so it is feasible to investigate multiple donors.

Monocytes and CD4+ T cells play a critical role in the immune response, and are valuable cell types to examine for ARFs due to their relevance in viral infection. HCMV establishes a latent infection in monocytes, providing a major site of persistent infection in the host. Differentiation of monocytes into macrophages or DCs promotes reactivation (Taylor-Wiedeman *et al*, 1991; Reeves & Sinclair, 2013; Poole *et al*, 2015). HIV is able to infect both monocytes and CD4+ T cells, due to expression of CD4 and viral co-receptors. HIV predominantly infects CD4+ T cells, but monocytes may also play an important role in infection, possibly acting as a viral reservoir (Campbell *et al*, 2014).

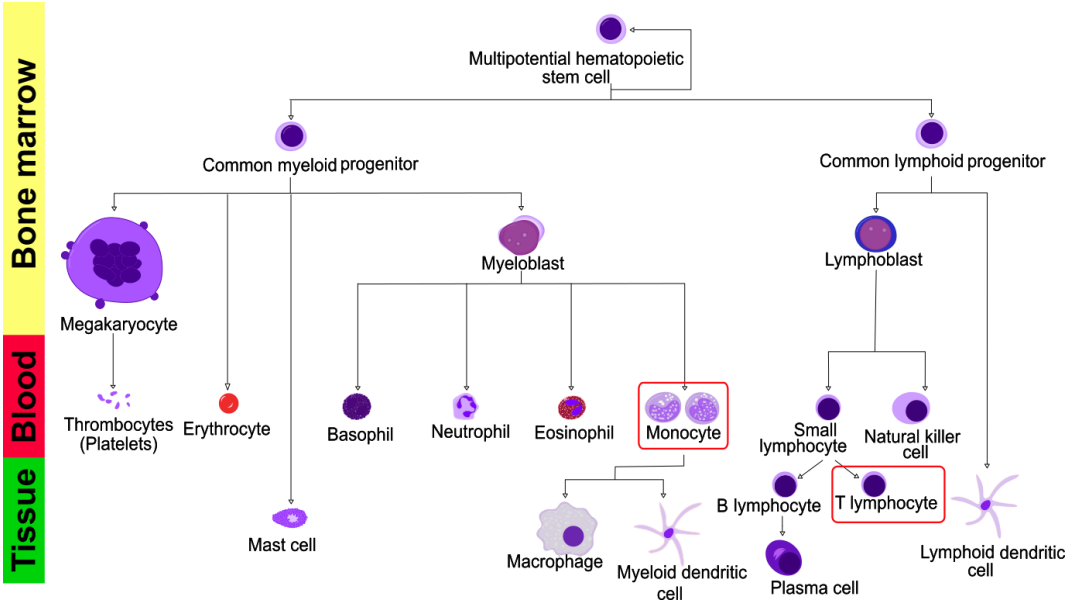


Figure 1.2 Blood lineages derived from hematopoietic stem cells

Multipotent stem cells differentiate into common myeloid or lymphoid precursors, and progenitor cells become increasingly specialised at each level of differentiation. Monocytes and T cells are derived from the myeloid and lymphoid lineages respectively. Figure is adapted from (Häggström, 2014), which was released under the Attribution-Share Alike 3.0 Unported license.

Published 'Omics Data on IFN Stimulation of Primary Monocytes and CD4+ T Cells

There is a wealth of previous 'omics data regarding IFN stimulation on a variety of cultured and primary cell lines, however the majority of this is transcriptomic. Much of this is deposited in the Interferome, a large database of IFN responsive genes, curated from various genome-wide microarray based studies (Rusinova *et al*, 2013). It includes data for type I, II and III IFNs and allows filtering based on the type of IFN and the cell types examined amongst other parameters. To date, it includes 40 experiments with type I IFNs, encompassing 107 datasets.

Searching for IFNα stimulation of primary monocytes yielded just one dataset in the Interferome (dataset 306) (Smiljanovic *et al*, 2012). This study used microarrays to compare expression profiles in patients with rheumatoid arthritis (RA) or SLE with healthy donors. They also looked at *in vitro* stimulation of primary monocytes pooled from a number of donors, with a range of cytokines (IFNα2a, IFNγ and tumour necrosis factor (TNF)-α) in order to generate profiles that could be matched to the disease

signatures. They determined that gene expression in SLE patients was predominantly driven by IFN, whilst TNF was dominant for RA, with disease dependent responses to each cytokine adding further complexity. The focus of the work was therefore on the disease aspects with little analysis of the IFN data, but it provides some data on IFNα stimulation of primary monocytes.

Another major investigation of IFNα stimulation of primary leukocytes at the transcriptomic level was by Schlaak *et al* (2002). In this study, four cell lines alongside primary PBMCs, T cells and DCs were stimulated with IFNα2a, before analysis of the expression of 150 known ISGs and genes of interest on a complementary DNA (cDNA) 'macroarray'. Only seven genes were induced across all cell types in all experiments, and multiple genes were only substantially induced in the primary cells. IRF7 was induced exclusively in the hematopoietic cells. Three donors were used for DCs, and five for T cells, allowing donor specific variation to be investigated. Multiple replicates of the experiment on cell lines generated similar results each time, suggesting a reproducible response to IFN, and a robust technique. However, much greater variation was observed amongst the primary cells from different donors. In the T cells, 45 genes were induced more than 2 fold in at least 1 donor, whilst just 10 genes were induced in all 5 donors. DCs also showed donor-specific responses, though with slightly less extreme variation between donors than in the T cells. This study was limited to only investigating a predetermined set of genes, predominantly decided based on previous studies of IFN stimulation of fibrosarcoma cells. The methodology therefore doesn't allow identification of novel ISGs, and even greater differences between the cell types may be expected in a broader study including more immune cell specific ISGs.

These transcriptomic studies provide valuable datasets, however the correlation between transcript and protein levels is often quite poor (Haider & Pal, 2013; Liu *et al*, 2016), and proteomic investigations are likely to reveal novel data, particularly given the ability to examine specific subcellular compartments such as the cell surface. One previous proteomic study utilised 2-dimensional gel electrophoresis (2DE) to investigate IFN stimulation of activated CD4+ T cells. The technique is known to underperform in the detection of membrane proteins and identified a very limited number of proteins; there were 11 'spots' of differentially expressed proteins between the IFN-stimulated and

unstimulated cells, corresponding to 7 proteins (Rosengren *et al*, 2005). The project presented here therefore substantially furthers these studies.

1.4 Screen 2: Systems Analyses of Viral Infection

As well as being stimulated by IFN, candidate restriction factors should also be downregulated by viral infection. In the work presented in this thesis, a new proteomic dataset has been generated investigating various aspect of VACV infection. This was considered alongside existing proteomic data regarding HCMV and HIV infection (Weekes *et al*, 2014; Matheson *et al*, 2015; Greenwood *et al*, 2016). These proteomic studies of viral infection provide a great deal of information on the ability of viruses to modulate host cells, as well as the kinetic profiles of expression of viral proteins, and have previously been used to identify novel facets of innate immunity and host-virus interactions. They have also been used here to identify candidate ARFs.

1.4.1 Vaccinia Virus

Background

VACV is a member of the *orthopox* genus of large dsDNA viruses (*Poxviridae* family, *Chordopoxvirinae* subfamily). It is genetically related to variola virus, the causative agent of smallpox. VACV infection was found to provide cross protection for smallpox, and was therefore used as an effective live vaccine, culminating in its eradication in 1980 (Fenner *et al*, 1988). It is unclear where VACV originated and it is believed to no longer exist in a natural host, but related viruses are seen circulating in cattle in Brazil (Moss *et al*, 2013). Despite the eradication of smallpox, research on VACV continues as a model for host-virus interactions. Additionally, it is being developed as a vector for vaccines, and as an oncolytic therapy (Verardi *et al*, 2012; Guo *et al*, 2019). A complete understanding of the virus is therefore critical to ensuring safety and efficacy.

VACV Life cycle

The VACV genome encodes approximately 216 proteins, and is conventionally divided into early (E) genes which are subdivided into E1.1 and E1.2, and post-replicative genes (PR). These PR genes are expressed following DNA replication, and are further divided into intermediate and late categories (Yang *et al*, 2010, 2011, 2015). The sequential expression of these is dependent on early genes encoding transcription factors for intermediate genes, which in turn encode transcription factors for the expression of late viral proteins. This expression cascade is also reliant on rapid turnover of mRNAs due to decapping by viral proteins. These decapping enzymes are additionally responsible for global shut down of host protein synthesis (Parrish *et al*, 2007; Parrish & Moss, 2007).

Virions enter the cell by fusion at the PM, or endocytosis. They are then transported on microtubules to cytoplasmic viral factories, where gene expression occurs. The virion therefore contains all the machinery required for transcription of the early viral proteins. The mRNAs produced are capped and polyadenylated. Early genes encode proteins required for modulation of the host immune response, DNA replication and transcription of intermediate genes. Following early gene expression, the viral core is disrupted in the proteasome dependent uncoating process, DNA replication occurs and intermediate genes are expressed. These encode transcription factors for late viral gene expression. The late genes encode virion structural proteins, and the early transcription factors which are packaged within the virion (Moss *et al*, 2013).

VACV initially assembles into an intracellular mature virus (IMV), which is the most abundant infectious form of the virus, and is released by cell lysis. Alternatively, the IMV may acquire an additional double membrane at the trans-golgi network or endosomes, forming an intracellular enveloped virus (IEV). The IEV is transported to the cell surface where it fuses with the PM, exposing the enveloped virion on the cell surface and forming a cell-associated enveloped virus (CEV). This may be propelled to surrounding cells by the formation of actin tails. Alternatively, the virus may be released as an extracellular enveloped virus (EEV) for long range spread of the virus (Smith & Law, 2004; Moss *et al*, 2013) (Figure 1.3).

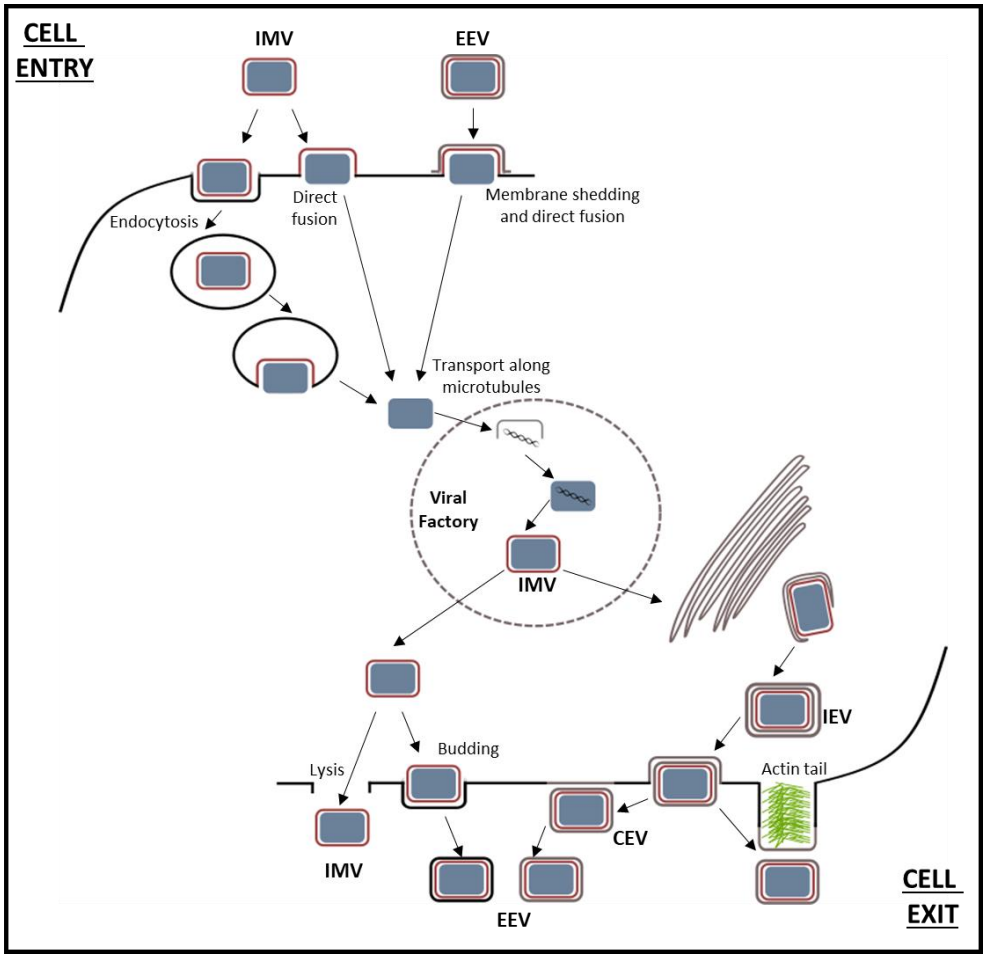


Figure 1.3 VACV life cycle

VACV IMVs or EEVs enter the cell by direct fusion or endocytosis, before being transported along microtubules to viral factories, where replication occurs. The IMV produced here may either be released by cell lysis, or acquire an additional membrane at the trans-golgi network or endosomes, forming an IEV. This is transported to the PM where it fuses and is exposed as a CEV, which may be released from the cell as an EEV. Figures is based on descriptions by (Smith & Law, 2004; Roberts & Smith, 2008).

Immunomodulation by VACV

A large proportion of the VACV genome is dedicated to subversion of the host innate immune response, blocking and evading the effects of complement, cytokines, chemokines, IFNs, apoptosis and NK cells (Smith *et al*, 2013a). Evasion of the type I IFN response is crucial to VACV, as replication in the cytoplasm makes VACV particularly susceptible to its effects. VACV therefore encodes multiple proteins to

inhibit the production and effects of type I IFN, in particular IFNβ, at all stages of the infection and the IFN response (Smith *et al*, 2018).

A key initial trigger for IFN production is detection by the host of viral infection, and as such, VACV has several mechanisms for inhibiting this step. VACV E3 for example binds the dsRNA which forms late during infection, limiting activation of PRRs (Chang *et al*, 1992), and C16 binds Ku, restricting the sensing of cytoplasmic DNA by the DNA-dependent protein kinase (DNA-PK) complex (Peters *et al*, 2013). A multitude of viral proteins then act to inhibit IRF3 and NF-κB signalling, to limit production of IFN. For example, A46 binds adaptor molecules such as MyD88, TRIF and translocating chain-associated membrane protein (TRAM) which interact with TLRs, preventing their activation of MAPK signalling, NF-κB and IRF3 (Smith *et al*, 2013a).

As some IFN is still likely to be produced by the cell, or by activated neighbouring cells, several viral proteins inhibit IFNs directly. Examples include B18, which is secreted from infected cells and binds IFN extracellularly to prevent it binding to the receptor (Colamonici *et al*, 1995; Symons *et al*, 1995). H1 dephosphorylates STAT proteins to block IFN signalling (Najarro *et al*, 2001) and C6 interacts with STAT2 to inhibit transcription of ISGs (Stuart *et al*, 2016), as well as inhibiting nuclear translocation of IRF3 (Unterholzner *et al*, 2011). Finally, VACV encodes proteins which antagonise ISGs. E3 binds ISG15 and suppresses its antiviral activity (Guerra *et al*, 2008), whilst C9 degrades the IFN-induced proteins with tetratricopeptide repeats (IFIT) proteins (Liu *et al*, 2019).

Systems Analyses of VACV Infection

A limited number of previous proteomic studies have been conducted which examined the modulation of the host proteome by VACV, often constrained by the technology available. Bartel *et al.* utilised 2DE followed by mass spectrometry to quantify 24 human proteins modulated by infection, and 3 viral proteins (Bartel *et al*, 2011). A further proteomic investigation studied human embryonic kidney (HEK293T) cells which were either mock treated, or infected for 20 hours in the presence or absence of cytosine arabinoside (AraC) (Chou *et al*, 2012). As AraC is a viral DNA replication inhibitor which prevents expression of PR genes, these were treated as early and late points during

infection respectively. Here, 136 viral proteins were quantified, and 3,798 host proteins. The overwhelming majority of proteins showing differential expression between the conditions were viral ones, with most host proteins not changing in abundance during infection.

Additionally, several studies have been conducted focussing on expression of the viral proteins. These were largely performed at the transcriptomic level, initially defining early and PR genes (Yang *et al*, 2010), and then further dissecting this class into the intermediate and late genes (Yang *et al*, 2011). The temporal expression of viral genes was also investigated at the proteomic level by Croft *et al*, performing two proteomic experiments at various time points up to 9.5 hours post infection (hpi). They quantified 101 viral proteins which were classified into 4 clusters. Of these, 47 proteins were quantified in both experiments and had concordant assignments to the 4 clusters (Croft *et al*, 2015).

1.4.2 Human Cytomegalovirus

Herpesviridae

The *Herpesviridae* family of viruses are enveloped, dsDNA viruses which generally establish a lytic infection that kills the cell upon release, as well as a latent infection for life-long persistence in the host. Historically, they were divided into alpha, beta and gamma herpesviruses based on biological characteristics, and more recently the classification was confirmed based on sequencing data. Broadly, *Alphaherpesvirinae* have a variable host range, relatively short replication cycle, and establish latent infections predominantly in neuronal cells. These includes HSV, and varicella-zoster virus (VZV), the causative agent of chickenpox. *Betaherpesvirinae*, such as HCMV, tend to have a more restricted host range and longer replication cycles, and establish latency in secretory glands, leukocytes, kidneys and other tissues. *Gammaherpesvirinae* include Epstein-Barr virus (EBV) and Kaposi sarcoma-associated herpesvirus (KSHV). These viruses are restricted to the family of their natural host, and usually to T or B lymphocytes, establishing latency in lymphoid tissues (Pellett & Roizman, 2013).

Impact of Human Cytomegalovirus

HCMV is a member of the *betaherpesvirinae* family. It is an ubiquitous pathogen, with a global seroprevalence of 83 %, with the highest seroprevalence in the WHO Eastern Mediterranean region (~90 %) and lowest in the WHO European region (~66 %) (Zuhair *et al*, 2019). For most individuals, transmission occurs during childhood, with the virus being acquired from contact with infected bodily secretions, and the infection is asymptomatic. However, reactivation during pregnancy may lead to transplacental transmission. Congenital HCMV infection can result in life-long sensorineural hearing loss, with ~10-15 % of babies infected showing symptoms, and others developing symptoms later on. Of the symptomatic infants, fatality is estimated to be between 4-30 % (Dollard *et al*, 2007). Initial infection with the virus during pregnancy increases the chance of transplacental transmission from ~1 % to around ~33 %, and is linked with severe disease (Kenneson & Cannon, 2007). Infection or reactivation is also damaging in immunocompromised individuals, particularly those undergoing organ transplants (Azevedo *et al*, 2015) or with acquired immunodeficiency disease syndrome (AIDS) (Emery, 2001).

Three drugs are currently available for treatment of symptomatic individuals, and pre-emptive therapy is sometimes employed where infection is the result of a medical intervention and can be predicted, such as organ transplants. These are all inhibitors of the viral polymerase ((val)Ganciclovir, Cidofovir and Foscarnet). There are a couple of other antivirals more commonly used for prophylaxis (Letermovir, Valacyclovir). However, these treatments have numerous side effects, such as leukopenia, renal dysfunction and ophthalmologic toxicity, as well as development of resistant viral infection. Additionally teratogenic effects have been observed in animals models so they are not recommended for use during pregnancy (Kotton, 2019; Malm & Engman, 2007). As HCMV is a leading cause of congenital disease, and due to the potential severity of infection, a vaccine, or improved antivirals, would therefore be highly desirable.

The Virus and its Life Cycle

HCMV has a the largest genome of the herpesviruses, at 236 kbp (Dolan *et al*, 2004), and coding for ~170 canonical proteins. Expression of these genes is temporally regulated,

and can be broadly divided into immediate early (IE), delayed early (DE) and late (L) viral genes (Wathen & Stinski, 1982), or into five classes (temporal profile (Tp) 1-5) by proteomic data (Weekes *et al*, 2014), and the replication cycle takes 48-72 h. The structure is characteristic of herpesviruses, with the dsDNA genome being encased in an icosahedral capsid, within an envelope acquired from the host cell membrane. HCMV tropism includes a wide range of cells such as fibroblasts, epithelial cells, endothelial cells, smooth muscle cells and macrophages (Sinzger *et al*, 1995).

HCMV can enter cells either by fusion of the viral envelope with the cell membrane (as during infection of fibroblasts) or following endocytosis (as in endothelial and epithelial cells). Microtubules then deliver the nucleocapsid to nuclear pore complexes where the viral genome enters the nucleus. The virus may enter a lytic life cycle. In this instance, IE genes regulate the transcription of host and viral genes. Around 14 - 24 hpi in fibroblasts, DE genes control viral DNA replication, followed by expression of L viral genes. The viral capsid assembles in the nucleus and then translocates to the cytoplasm. Throughout this process, it acquires a temporary envelope at the inner nuclear membrane, and then undergoes de-envelopment at the outer nuclear membrane. The nucleocapsid acquires tegument proteins at the cytoplasmic viral assembly compartment (AC) and undergoes secondary envelopment at the endoplasmic reticulum – golgi intermediate compartment. Finally, the virion is released from the cell by exocytosis (Fields *et al*, 2013; Beltran & Cristea, 2015) (Figure 1.4).

Alternatively, HCMV can establish a life-long latent infection. As with other herpesviruses, in latent infection, no new virions are produced, and it was thought that a smaller subset of viral genes are expressed (Poole & Sinclair, 2015). However, recent research found that gene expression during latency is similar to that of late lytic infection, but at lower levels, suggesting the differences between lytic and latent infection may be quantitative rather than qualitative (Shnayder *et al*, 2018). Latency occurs in myeloid cells, however on differentiation to macrophages or DCs this can result in reactivation of the virus. It is important to understand latency as it may provide new therapeutic avenues for clearing latent infection before reactivation occurs during transplant (Poole *et al*, 2014).

Diagram of HCMV life cycle removed for copyright reasons. Copyright holder is Taylor & Francis.

Figure 1.4 HCMV life cycle

CMV virions initiate infection by attaching to receptors at the cell surface and then fusing with the PM, or by endocytosis. The tegument proteins and capsid are released to the cytosol. Infectious HCMV particles enter the cell through receptor interactions, and the capsid and tegument proteins are delivered to the cytosol. The capsid travels to the nucleus along microtubules, where the viral genome is released and the gene expression cascade occurs (expression of IE, DE and then L viral genes). L gene expression triggers capsid assembly in the nucleus. Nuclear egress occurs, and the virus acquires tegument proteins in the cytosol. The virus is trafficked to AC where it acquires more tegument proteins and a viral envelope. Infectious particles are released by exocytosis. Figure reprinted from (Beltran & Cristea, 2015).

HCMV and the Immune Response

The ability of HCMV infection to remain asymptomatic in healthy individuals, whilst it can cause devastating disease in the immunocompromised, highlights the effectiveness of the immune response in controlling the infection. HCMV has a multitude of mechanisms for subverting the immune response.

HCMV infection is initially controlled in a number of ways by the intrinsic and innate immune response, including detection of the envelope glycoproteins B and H by TLR2, subsequent signalling through NF-κB and production of IFN (Boehme *et al*, 2006). Additionally, NK cells play a crucial role in control of HCMV infection (Biron *et al*, 1989; Gazit *et al*, 2004; Vivier *et al*, 2008). Several restriction factors are involved in the immune response (see section 1.2.2), including the subnuclear structure ND10. This is composed of promyelocytic leukaemia protein (PML), death domain-associated protein 6 (Daxx) and nuclear autoantigen Sp-100 (Sp100). Viral genomes are found in close proximity to the ND10 bodies, and knockdown of any of these proteins leads to enhanced replication of HCMV (Tavalai *et al*, 2008; Adler *et al*, 2011). Daxx limits transcription of IE genes by generating a repressive chromatin structure at the viral major immediate early promoter (Woodhall *et al*, 2006). Sp100 is an IFN-stimulated protein, and also limits viral transcription. HCMV evades this restriction by delivering the viral protein pp71 in the virion and this targets Daxx for proteasomal degradation, stimulating expression of IE genes (Saffert & Kalejta, 2006); HCMV IE1 downregulates Sp100 (Tavalai *et al*, 2011).

IE proteins also disrupt the IFN mediated immune response, with IE1 preventing STAT1, STAT2 and IRF9 associating to form ISGF3 (Paulus *et al*, 2006), and IE2 blocking the production of IFNβ (Taylor & Bresnahan, 2005). Other examples of HCMV proteins involved in immune evasion include US2, one of several HCMV proteins able to modulate MHC class I expression. It also downregulates integrins and thrombomodulin, and can act synergistically with UL141 to downregulate the NK cell ligand CD112 (Hsu *et al*, 2015). Alongside downregulating MHC class I, HCMV encodes UL18, an MHC class I homologue which binds the inhibitory leukocyte immunoglobulin-like receptor 1 (Lir-1) (Yang & Bjorkman, 2008).

Adaptive immunity is also crucial in controlling infection, as evidenced by HCMV reactivation in immunosuppressed transplant recipients (Azevedo *et al*, 2015). Presentation of HCMV antigens on APCs leads to a large HCMV-specific T cell response, similar in size to that seen with HIV and larger than for most other viruses, composed of virus-specific CD4+ and CD8+ T cells. The scale of the T cells response varies between individuals, but examination of memory T cells in seropositive individuals found ~10% of CD4+ and CD8+ cells were reactive to HCMV (Rosa & Diamond, 2012; Sylwester *et*

al, 2005). Adoptive transfer of HCMV specific CD8+ T cells reduced viremia in patients, emphasising the importance of the T cell response in controlling infection (Cobbold *et al*, 2005). B cells are also important in the adaptive immune response, with antibodies generated against tegument, envelope and non-structural proteins (Landini *et al*, 1988; Urban *et al*, 1996).

Proteomic studies of HCMV used in this thesis

A previous study used quantitative temporal viromics to analyse the proteomic changes observed in the host and virus during HCMV infection (Weekes *et al*, 2014). The data from Weekes *et al.* has been used here for comparison with VACV infection, and for the identification of ARFs. Fibroblasts were infected with the Merlin strain of HCMV, and either whole cell lysate (WCL) or PM enriched samples harvested at various time points throughout a single viral replication cycle. Each sample was digested and labelled with a different tandem mass tag (TMT), prior to analysis by triple-stage mass spectrometry (MS3). This gave the relative quantitation of each protein across the samples (proteomic techniques are discussed further in section 1.5. For each of the WCL and PM experiments, samples were harvested either at the time points described in Figure 1.5 (WCL2 and PM2) or in duplicate at mock, 24, 48 and 72 hpi (WCL1 and PM1). More than 8000 human proteins were quantified over the time-course, including ~1200 proteins with PM related gene ontology (GO) annotations, as well as 139 of the 171 canonical HCMV proteins.

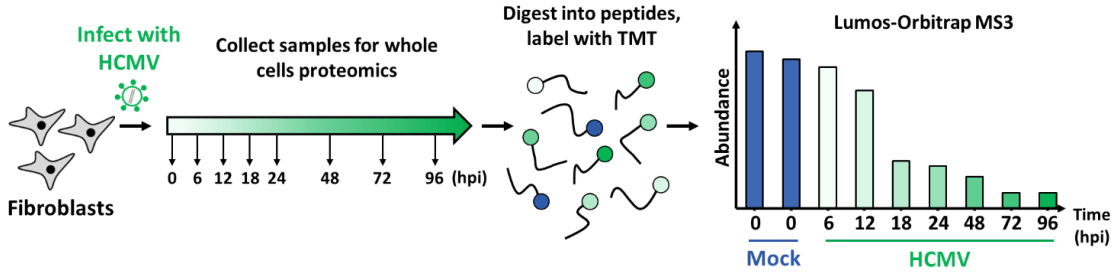


Figure 1.5 Schematic of HCMV time-course proteomics experiment

Fibroblasts were infected with the Merlin strain of HCMV, and samples harvested at various time points throughout infection for either WCL or PM analysis (Weekes *et al*, 2014). The proteins were then digested, labelled with TMT and subjected to mass spectrometry analysis. The time points displayed are those relating to the WCL2 and PM2 experiments.

Proteins downregulated by HCMV infection were identified in the data and used to predict those that might be important in infection. Examination of the downregulated proteins identified multiple protocadherins, some of which are putative NK cell ligands, as well as modulation of proteins in the IFN induction and signalling pathways, and ISGs. Additionally, analysis of the kinetics of expression of viral proteins defined five temporal classes of protein expression (Tp1-5).

This data demonstrates the power of multiplexed proteomics in large scale studies of infection. It also provides a comprehensive resource which is used in this thesis to aid identification of candidate ARFs. Several additional proteomic studies on HCMV have been conducted during my PhD in the Weekes lab. As they are not central to this project, they are not discussed further, however they focus on protein degradation during HCMV infection, leading to the identification of HLTF as a novel ARF (Nightingale *et al*, 2018), and the generation of a global interactome of HCMV proteins (Nobre *et al*, 2019).

1.4.3 HIV

Background

HIV was first discovered in the 1980s when individuals started presenting with unusual symptoms including lymphadenopathy, opportunistic infections, unusual cancers and a depletion of CD4+ T cells; this became recognised as AIDS. HIV was isolated in 1983 and found to be the causative agent of the disease. Several years later a second strain was discovered, and these have now been designated HIV-1 and HIV-2. The former is more dominant and pathogenic, and therefore the focus of much research and of the proteomic investigations discussed here. HIV is part of the lentivirus genera of *retroviridae*. These are a large family of viruses, which have an RNA genome that is reverse transcribed and integrated into the host chromosomal DNA upon infection. This then provides the template for production of viral RNAs for formation of new virions. The genome encodes 16 proteins, including structural proteins, viral enzymes, envelope proteins, regulatory proteins and accessory proteins.

It is estimated that in 2019, 37.9 million people globally were living with HIV. In 2010, just 7.7 million people had access to antiviral therapies, but by 2018 this had risen to 24.5 million. However, there were still 770,000 AIDS related deaths (end of 2018). It is estimated that in 2020, AIDS response in low and middle-income countries will cost $26.2 billion (UNAIDS, 2019). Therefore, despite improved access to therapy and a reducing death toll, HIV still poses a great social and economic burden. It has also been the subject of much ARF related research.

HIV life cycle

The HIV envelope glycoprotein 120 (gp120) binds CD4 to initiate cell entry. The main target of HIV is CD4+ T cells, but other cells expressing CD4 can be infected, including monocytes and DCs. Additional binding to a co-receptor, C-C motif chemokine receptor 5 (CCR5) or C-X-C motif chemokine receptor 4 (CXCR4), leads to exposure of the viral 'fusion peptide', part of glycoprotein 41 (gp41), which inserts into the cell membrane and causes membrane fusion. The capsid partially uncoats to allow reverse transcription of the viral RNA, and the linear DNA is integrated into the host chromosome. Following integration, the late phase of HIV replication occurs. This includes viral gene expression and the assembly of virus like particles at the PM. The particle buds through the PM and undergoes maturation, where the precursor Gag-Pol polyprotein is cleaved, resulting in the mature infectious virion (Engelman & Cherepanov, 2012; Maartens *et al*, 2014).

HIV and the immune response

The innate immune system is activated early during infection. The predominant pathways for this differ between infected cell types, but include detection of ssRNA by TLR7, and detection of DNA by cyclic GMP-AMP synthase (cGAS) in plasmacytoid DCs and monocyte-derived DCs respectively, whilst DNA sensing by IFI16 is important in CD4+ T cells. Even though they are not infected themselves, mucosal epithelial cells detect gp120 by cell surface TLR2 and TLR4, and secrete inflammatory cytokines. These signalling pathways lead to IFN induction, stimulating the production of ISGs, including numerous ARFs (see section 1.2). However, HIV has many mechanisms for downregulating the immune response, such as Vpx antagonising SAMHD1, and Vpu antagonising BST2. NK cells are also activated and are thought to have roles both in

killing infected cells and regulating adaptive immunity (Altfeld & Gale, 2015; Silvin & Manel, 2015; Bergantz *et al*, 2019).

There is a complex relationship between HIV infection and the pool of CD4+ T cells which it infects. Following an initial infection, there is an 'eclipse period' where no viral RNA is detected. Once the virus infected cells reach the draining lymph node and more CD4+ CCR5+ T cells are available for infection, peak viremia is reached. Following this, the viral load decreases to its 'set point'. There may be a role of CD8+ T cells in this control of viremia. The previously depleted CD4+ T cell levels return to normal, however the adaptive immune system is still activated. It is hypothesised that the activation of CD4+ T cells in HIV infection may actually worsen the pathogenesis, providing more targets for the virus to infect (McMichael *et al*, 2010; Mogensen TH *et al*, 2010).

Proteomic Studies of HIV Infection

HIV has also been studied using global proteomic techniques, and a couple of key TMT-based studies have been examined for comparison to other data in this thesis. In one study, CEM-T4 cells were infected with an Env-deficient, vesicular stomatitis virus glycoprotein G (VSVg)-pseudotyped HIV-1 virus for a PM proteomic time-course. TMT-based mass spectrometry was used to analyse samples at 0, 6, 24, 48 and 72 hpi. A total of 2320 proteins were quantified, including 804 proteins with PM related GO annotations. More than 100 proteins were downregulated during infection, and these were involved in cell adhesion, leukocyte activation and transmembrane transport (Matheson *et al*, 2015). An identical time-course was later performed, this time harvesting WCL samples. More than 6500 proteins were quantified, and many of the known ARFs were observed to be modulated by infection, as expected. Additional experiments considered the effects both on the WCL and the phosphoproteome of infection with a Vif deletion virus, and identified that depletion of protein phosphatase 2A (PP2A) during infection was Vif dependent (Greenwood *et al*, 2016). Most recently, a time-course was performed using primary CD4+ T cells. In this case, the cells were infected with a HIV reporter virus expressing a cell surface streptavidin binding tag, allowing infection at a multiplicity of infection (MOI) of one and separation of a pure population of infected cells. WCL samples were harvested from uninfected cells (either resting or activated), and then at 24 and 48 hpi, identifying additional proteins regulated

by HIV which had not been seen in T cell lines, such as AT-rich interactive domain-containing protein 5A (ARID5A) (Naamati *et al*, 2019).

1.4.4 Other TMT Based Quantitative Temporal Analyses of Viral Infection

Many other proteomic studies of viral infection have been performed, such as multiple time-courses of influenza infection (Lietzén *et al*, 2011; Kummer *et al*, 2014; Turnbull *et al*, 2016). Alongside those already discussed, other TMT-based proteomic time-courses include investigations of EBV, BK polyomavirus and HSV-1. Examination of EBV infection at both the whole cell and PM levels identified modulation of the cell cycle and innate and adaptive immune pathways. I performed a comparison of this data with the HCMV proteomic time-course, identifying common downregulation of a protocadherin, as well as multiple neuroligins; some of these were also modulated by KSHV (Ersing *et al*, 2017). BK polyomavirus is a small dsDNA virus. It expresses just seven known proteins and establishes life-long persistent infection. Proteomic investigation demonstrated the importance of the cell cycle in viral replication, and a lack of innate immune response (Caller *et al*, 2019). Analysis of HSV-1 infection quantified 6956 human proteins and 67 viral proteins over the course of infection. Multiple host proteins were downregulated, including golgi associated PDZ and coiled-coil motif containing protein (GOPC), a trafficking protein. The viral protein pUL56 was found to target GOPC for degradation, therefore modulating the abundance of signalling molecules at the cell surface (Soh *et al*, 2019). These studies demonstrate the power of quantitative proteomic investigations of viral infection, to study host-virus interactions.

1.5 Key Technology Used: Proteomics

Many molecular biology techniques are available to study the presence, abundance and interaction of proteins. Immunoblotting, immunofluorescence and flow cytometry are all essential in the lab, however, these methods are low throughput, investigating just a few proteins at a time, and are reliant on high quality reagents such as appropriate antibodies.

The term ‘proteomics’ was coined in 1995, following the first experiments in 1975 (Graves & Haystead, 2002). Since then, proteomic technologies have developed, and now enable an unbiased, high throughput analysis, generating quantitative data on thousands of proteins in a single experiment. The use of proteomic technologies is central to this thesis, having been applied to investigation of both IFN stimulation and viral infection; the value of proteomic technology in this field is evidenced by the findings in various viral infections discussed previously (see section 1.4).

1.5.1 Basic Methodology for Bottom-Up Proteomics

The basic principle of proteomics involves subjecting proteins to analysis by mass spectrometry, in order to identify them. In most cases (and in this thesis), a ‘bottom-up’ approach is used, whereby the extracted proteins are digested into peptides prior to mass spectrometry. This allows the analysis and identification of peptides, rather than intact proteins as is required for top-down proteomics, and is less challenging and more routinely used. The peptide sample is separated by liquid chromatography (LC) prior to analysis by mass spectrometry. In tandem mass spectrometry, the peptides are first ionised, converting them into the gas-phase, and are then analysed to determine their mass to charge (m/z) ratio in an MS1 scan. Selected precursor ions are then fragmented (often by collision induced dissociation, CID), and the m/z of these fragments is determined (ms/ms, or ms2), and used to create a spectrum (Zhang *et al*, 2013).

This process can be further enhanced by fractionation of the peptide sample prior to LC-mass spectrometry. Fractionation separates the mixture of peptide mixture into multiple samples of reduced complexity. This simplification of the sample reduces the co-elution of peptides, and allows a more comprehensive examination of the sample, with more peptides being quantified (Chan & Issaq, 2013).

Following the mass spectrometry run, the data is processed computationally. The raw data is converted to an appropriate format for analysis, prior to identification of the spectra. Peptide identification requires a theoretical database of peptides, generated by *in silico* digestion of the protein database. The intact mass from the MS1 scan can be used to identify peptides with a similar m/z, and the spectrum from the MS2 scan is

compared to the theoretical spectra for these peptides. This is done using searching tools such as Sequest (as used here), and Mascot. Peptide-spectrum matches require filtering to determine those that are most likely to have been correctly determined. A target-decoy strategy can be employed whereby 'decoy' peptides are included in the database, in order to estimate the false discovery rate according to how many spectra are matched to these decoy peptides. In practice, the most commonly employed approach to generate a decoy database is to reverse all peptide sequences. The filtered peptides then require assembly into the proteins from which they originated. This is also a complex process, as some peptides are redundant to multiple proteins. Assignment therefore uses an 'Occam's razor' approach, which aims to generate the smallest number of proteins possible to account for the observed peptides (Noor *et al*, 2020).

1.5.2 Tandem Mass Tags for Quantitative Multiplexed Proteomics

The basic proteomic methodology described has been extended to enable various additional information to be gained. In this thesis, quantitative proteomics has been performed using TMT (Thompson *et al*, 2003; McAlister *et al*, 2012). At the time of performing these experiments, there were 11 different tags, but this has since been extended to 16 tags (Thermo), each of which can be used to label a different sample, allowing them to be multiplexed. TMTs are isobaric tags which contain an amine reactive group that can link to lysine residues or peptide N-termini, and a cleavable reporter. Each of the cleavable reporters has a different mass due to the incorporation of heavy nitrogen and carbon atoms. An additional round of tandem mass spectrometry (ms/ms/ms or ms3) selects the most abundant ions from the MS2 spectrum, and uses high energy CID (HCD) to dissociate the cleavable reporter from the isolated peptides; the signal from the reporter is proportional to the abundance of the labelled peptide (Ting *et al*, 2011; McAlister *et al*, 2014). The m/z of the reporter is used to determine which sample it came from. The relative abundance of a given protein in all samples is therefore determined.

Advantages of this approach are that multiplexing allows precise relative quantitation, and also reduces the running time required on the mass spectrometer. Additionally, unlike the commonly used metabolic labelling technique, stable isotope labelling with amino

acids in cell culture (SILAC), the samples are labelled following digestion, rather than requiring cells to be grown in specific media. This enables investigation of primary cells.

Quantitative multiplexed proteomics enables the comparison of any set of conditions. Here, it has been used to investigate IFN stimulation of multiple donors, and to examine various time points during infection. Other uses include comparison of infection of wild type (WT) and mutant viruses to help decipher the function of the viral proteins (Nightingale *et al*, 2018), identifying biomarkers (Whitwell *et al*, 2020), or other aspects of drug discovery (Amiri-Dashatan *et al*, 2018).

Additionally, the production of the sample for labelling can be modified to consider different aspects of the proteome. One example of this is enriching for PM proteins, so these can be specifically examined, as discussed below. Alternatively, post translational modifications can be considered by enriching for phosphoproteins (Swaney & Villén, 2016), or other modifications. Spatial proteomics uses different speeds for centrifugation of the cell lysate in order to separate out the nuclear, organellar and cytoplasmic fractions, each of which can be labelled with a different TMT, in order to determine where in the cell different proteins are localised (Itzhak *et al*, 2017).

1.5.3 Plasma Membrane Profiling

PM proteins are of low abundance, hydrophobic and often underrepresented in proteomic samples. Enrichment for these proteins prior to digestion, labelling and analysis by mass spectrometry enables a more detailed investigation of the cell surface, and a variety of methods have been developed for this purpose; these are reviewed in (Li *et al*, 2020) and briefly summarised here. The methods can be broadly categorised into approaches that separate PM proteins based on biophysical characteristics, those that use interaction of the proteins with lectins or antibodies, or by chemically coupling a tag to an exposed protein and using affinity purification to extract the bound proteins. There are some additional methods utilising metabolic labelling or enzymes, however metabolic labelling is restricted to cultured cells, and the enzyme based approaches are less well developed.

Separating proteins on biophysical characteristics could be based on size, charge or hydrophobicity. Size and density based separation for example, has utilised centrifugation, though this allows relatively low resolution, and often requires a large amount of starting material. The second category of methods, affinity based enrichment for PM proteins using antibodies, only allows enrichment of known PM proteins, and is further limited by the requirement for high specificity antibodies. Affinity to lectins allows a more unbiased approach, particularly if multiple lectins which bind different glycan structures are combined, however the coverage achieved is still limited. Finally, chemical labelling of amines, carboxyl groups or glycan chains has been commonly utilised. These approaches generally involve a reagent with a reactive residue for binding to the protein, a linker, and a tag, often biotin.

Multiple methods were previously compared, recording the number of proteins identified, and percentage of these which were annotated as being at the PM (Weekes *et al*, 2010). Some approaches resulted in an increased proportion of intracellular proteins being identified, due to cell permeability, whilst others led to high levels of streptavidin contamination. Biotinylation with amino-oxy biotin, and a pull-down with high capacity streptavidin beads, led to the greatest number and purity of PM proteins. PM profiling was therefore developed using this approach (Weekes *et al*, 2012, 2014). It involves the oxidation of sialic acid residues on cell surface glycoproteins into aldehydes, and subsequent biotinylation of these with amino-oxy biotin; this reaction is catalysed by aniline. The high affinity interaction between biotin and streptavidin is utilised to enrich for these biotinylated PM proteins (Figure 1.6). In addition to experimentally enriching for cell surface proteins, the proteins quantified can be filtered to identify those with annotations suggesting they may be present at the PM; this is done using GO terms (Ashburner *et al*, 2000; The Gene Ontology Consortium, 2019) related to the PM. This accounts for the majority of proteins quantified if the sample is not fractionated, and is reduced to ~60% in fractionated data, as more proteins are quantified and therefore this includes more abundant intracellular contaminants. However, it is worth noting that this filtering approach may miss any poorly annotated PM proteins.

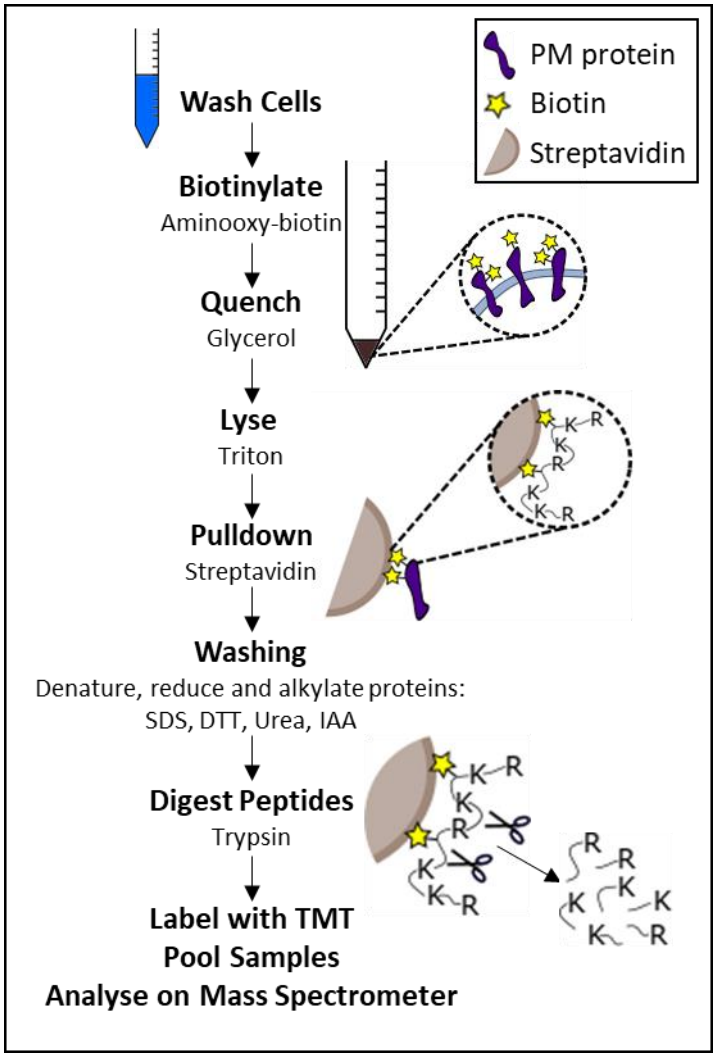


Figure 1.6 Plasma membrane profiling

PM profiling is used to examine cell surface proteins by mass spectrometry. The sialylated PM proteins are first biotinylated and then pulled-down using streptavidin beads, prior to on bead digestion of the proteins by trypsin. Peptides from each samples are labelled with TMT, pooled, and analysed by mass spectrometry. Further details can be found in the methods (chapter 2.9).

1.6 Aims of this Thesis

The aims of this thesis were to use unbiased proteomic screens to investigate two characteristic features of ARFs: stimulation by IFN, and downregulation by viral infection, and to identify putative restriction factors on the basis of these characteristics.

Aims:

1) **Characterise the effects of IFNα stimulation at the surface of primary monocytes and T cells:** The first screen used PM profiling to characterise in depth the cell surface proteome of primary monocytes and CD4+ T cells, determine how IFN stimulation modulates this proteome, and examine the variation in protein expression between different donors. This data is presented in chapter 3.
2) **Examine proteomic changes induced upon infection with VACV:** The second screen used whole-cell based proteomics to characterise VACV-infected cells, surveying all proteomic changes induced by the virus. This allows identification of enriched groups of modulated proteins that may be biologically important. Further proteomic screens were conducted to examine which proteins were specifically targeted for downregulation by the virus, and in the investigation of predicted host-pathogen interactions. This data is presented in chapter 4.
3) **Use data on IFN stimulation and viral downregulation to identify candidate ARFs:** The data from the screens presented here is examined alongside previously published proteomic screens on other viral infections, in order to identify novel putative restriction factors. This led to the identification of Endothelin Converting Enzyme 1 (ECE1), which is then investigated in the context of HCMV infection. This is presented in chapter 5.

CHAPTER 2: MATERIALS AND METHODS

2.1 Ethics Statements

Ethical approval for the use of blood samples (as required for the proteomic screen investigating IFN induced changes at the surface of primary leukocytes) was granted by the University of Cambridge Human Biology Research Ethics Committee (HBREC.2016.011) and written informed consent was obtained from volunteers prior to blood donations. Leukocyte enriched blood samples were obtained from the National Health Service Blood and Transplant service (NHSBT, Cambridge, UK). These were from anonymised donors, and composed of cells remaining following clinical use of blood donations.

2.2 Cell Culture

2.2.1 Cell lines

All cells were cultured at 37 ºC, 5 % CO_2.

Cultured Monocyte and T Cell Lines for IFN Screen

THP-1 (human monocyte) cells were used for proteomic studies of IFN stimulation, and for validation. Jurkats and SUPT1s (human T cells) were used for validation of primary T cell data. They were all maintained in RPMI-1640 medium (R8758, Sigma) supplemented with 10 % volume per volume (v/v) foetal bovine serum (FBS, Sigma). The THP-1s were kindly provided by Professor Paul Lehner (Department of Medicine, University of Cambridge), and the Jurkats and SUPT1s were from Dr Nick Matheson (Department of Medicine, University of Cambridge). For proteomic and immunoblot experiments involving IFN stimulation, THP-1, Jurkats or SUPT1s were stimulated with 1000 arbitrary units (AU)/ml IFNα2a (Reagent Proteins, BCA-309), in their regular culture media.

Cancer Cell Lines for ECE1 Antibody Validation

Three different cancer cell lines were used to produce lysates for testing of ECE1 antibodies. DLD1 and SW48 cells (both human colorectal adenocarcinoma) were kindly provided by Dr Matthew Hoare (Cancer Research UK, Cambridge). MDA-MB-231 (epithelial) cells were provided by Professor Paul Lehner; these cells were transduced with CRISPR-associated protein 9 (Cas9) due to previous use, though not relevant to these investigations. All cells were maintained in RPMI-1640 supplemented with 10 % FBS and 1 % penicillin / streptomycin (pen/strep) (Sigma).

Other Cell Lines

Human foetal foreskin fibroblasts (HFFF) were immortalised with human telomerase (HFFF-TERT) as described in (McSharry *et al*, 2001), and kindly provided by Dr Peter Tomasec (Cardiff University). These were used for the VACV infection time-course and the HCMV restriction assays. These cells have the same morphology and growth characteristics as HFFFs, and are permissive to infection with the HCMV Merlin strain. HEK-293Ts were used for production of lentivirus to make various HFFF-TERT cell lines; these were kindly supplied by Professor Paul Lehner. Both HEK-293Ts and HFFF-TERTs were maintained in Dulbecco's Modified Eagle Medium (DMEM, D6429, Sigma) supplemented with 10 % v/v FBS.

2.2.2 Cell Culture for Maintenance

Suspension cells were split as required, by taking an aliquot of the cells and diluting into fresh media at a ratio between 1:3 and 1:10. Adherent cells were washed once with phosphate buffered saline (PBS, Sigma), and incubated at 37 ºC with trypsin (Sigma) until all cells had detached from the flask. Cells were then resuspended in fresh media to split at a ratio between 1:2 and 1:10 as required.

2.3 Primary Cell Enrichment and Culture

Protocols are described as they were used for the proteomic screen investigating the effects of IFN stimulation at the surface of primary leukocytes. Methods were scaled accordingly for optimisation and validation experiments.

2.3.1 CD4+ T cells

Isolation of PBMCs from Peripheral Blood

Cells were isolated from peripheral blood for the enrichment of CD4+ T cells. 50 ml blood was aspirated with 4 ml 4 % citrate (4 % weight per volume (w/v) trisodium citrate dehydrate dissolved in H_2O). The citrated blood was diluted 1:2 in PBS/citrate (4 % citrate diluted 1:10 in PBS to a final concentration of 0.4 %), and 25 ml slowly layered onto 15 ml Ficoll (GE Healthcare) for density gradient centrifugation (x 4 tubes). The 4 tubes were weighed to ensure they were within 0.5 g of each other and centrifuged at 800 g, 20 mins with the brakes off. The layer containing the buffy coat was aspirated into new falcon tubes (x 2 tubes), made up to 50 ml with PBS/citrate and mixed by inverting. Peripheral blood mononuclear cells (PBMCs) were centrifuged (800 g, 10 mins), the supernatant poured off and the pellets from the two tubes resuspended and combined. This was again made up to 50 ml in PBS/citrate, and washed by centrifugation (400 g, 10 mins). Two further washes were performed (400 g, 10 mins).

Enrichment of CD4+ T Cells by Negative Selection

Following isolation of the PBMCs, the cells were counted and $5x10^5$ retained for analysis by flow cytometry (FCM). The remaining cells (~$7x10^7$) were enriched for CD4+ T cells using the 'Dynabeads Untouched Human CD4 T Cells' kit (Invitrogen, 11346D) according to the manufacturers instructions. Approximately $1.5x10^7$ cells were obtained from the negative selection. Again, $5x10^5$ cells were kept to confirm enrichment by FCM, and the remainder cultured overnight with or without IFNα2a.

Overnight Culture and IFN stimulation

For the proteomic screen, cells were centrifuged (400 g, 5 mins) and resuspended in 10 ml RPMI-1640 medium supplemented with 10 % human AB serum (Sigma, H4522). 5 ml was added to each of two wells in a 6-well plate, and one well stimulated with 1000 AU/ml IFNα2a. The cells were maintained overnight at 37 ºC, 5 % CO_2 prior to enrichment of PM proteins for mass spectrometry (MS) analysis.

2.3.2 CD14+ and Pan- Monocytes

Isolation of PBMCs from NHSBT leukocyte cone

PBMCs were isolated from leukocyte cones obtained from the NHSBT, and enriched for CD14+ monocyte or pan-monocyte populations. Blood was transferred from the cone into a falcon tube and ~40 ml PBS/citrate used to wash through the cone. The resulting 50 ml blood was further diluted 1:4 PBS/citrate, and 25 ml citrated blood layered slowly onto 13 ml Ficoll (x 8 tubes). The 8 tubes were weighed to ensure they were within 0.5 g of each other and centrifuged at 800 g, 20 mins with the brakes off. The buffy coats from the 8 tubes were aspirated and transferred into 6 falcon tubes, each made up to 50 ml with PBS/citrate and centrifuged at 600 g, 8 mins. The supernatant was poured off, the pellets resuspended and combined into 4 tubes, and centrifuged at 400g, 8 mins. This was repeated twice more, combining the pellets into two tubes and finally a single tube.

Enrichment of CD14+ and Pan- Monocytes by Negative Selection

Following isolation of the PBMCs, the cells were counted and $5x10^5$ retained for analysis by FCM. Approximately $6x10^8$–$1x10^9$ cells were obtained from the leukocyte cone, and of these $2x10^8$ were enriched for CD14+ monocytes (more for the pan monocytes as additional experiments were performed alongside this proteomic screen). Enrichment for CD14+ and pan- monocytes was by negative selection, using the 'Monocyte Isolation Kit II' (Miltenyi Biotec, 130-091-153) and 'Pan Monocyte Isolation Kit' (Miltenyi Biotec, 130-096-537) respectively, according to the manufacturer's instructions. Enrichment

resulted in ~6x10^7 CD14+ monocytes. Of these, 5x10^5 were retained for validation of enrichment by FCM.

Overnight Culture and IFN stimulation

Of the cells obtained from enrichment, 3x10^7 cells were resuspended in 6 ml X-Vivo 15 serum-free hematopoietic cell medium (Lonza, BE02-060F). This was split into two U-bottom tubes for culturing overnight. One was left unstimulated and the other was stimulated with IFNα2a (1600 AU/ml for CD14+ monocytes, 1000 AU/ml for pan-monocytes). The cells were maintained overnight at 37 ºC, 5 % CO_2.

2.4 Generation of Cell Lines

All cell lines were produced in either HFFF-TERTs (ECE1 assays) or THP-1s (transmembrane protein 123 (TMEM123) antibody tests), as described. The generation of constructs and production of cell lines is detailed below for each type of cell line each of the different types of cell lines.

2.4.1 Overexpression Cell Line

Production of ECE1 Overexpression Plasmid and Cell Line

Cells overexpressing ECE1 were produced using the gateway cloning system (Invitrogen). In this system an insert is transferred from an entry vector into an expression (or 'destination') vector via recombination at attL and attR sites. ECE1 (isoform b) in the pENTR223 entry vector was purchased from Harvard plasmids (HsCD00399893) and transferred into the expression vector pHAGE-P_{SFFV} by an 'LR' reaction. The insert here lacked a stop codon, so produced a FLAG-HA tagged version of ECE1.

For the LR reaction, the entry vector (pENTR223) was diluted 1:5, and 1 µl (~30 ng) was incubated with 0.2 µl of the destination vector (pHAGE-P_{SFFV}), 0.4 µl TE and 0.4 µl LR clonase II overnight at room temperature. The reaction was terminated by treatment with

0.2 µl proteinase K at 37 ºC for 10 mins. The LR reaction mix was transformed into competent cells as described in section 2.4.5. Insertion of the ECE1 sequence into the destination vector was confirmed by sequencing using the SFFVp forward primer and pHAGE-Rev reverse primer, as well as an internal primer specific to ECE1 (Table 2.3). Following confirmation of the insert, HFFF-TERTs were transduced with the pHAGE-P_{SFFV}-ECE1 plasmid as described in section 2.4.7.

Other Overexpression Cell Lines

Daxx and Sp100 overexpressing HFFF-TERTs, as well as cells transduced with the control gateway plasmid (GAW) were kindly provided by Dr Luis Nobre and Dr Katie Nightingale. In the cases of Daxx and Sp100, the plasmids were generated by a polymerase chain reaction (PCR) from either cDNA or purchased plasmids. Primers were used for the PCR reaction that incorporated attB sites, and a gateway BP reaction could then be used to insert the gene into a donor vector (pDONR223), to form an attL containing entry vector. This could then be used for an LR reaction as described above. Dr Nightingale also provided me with the GAW and Daxx plasmids for generation of the HFFF-TERT cell lines used in the final restriction/plaque assay. The HCMV-US2 overexpressing HFFF-TERT cell line was made by Dr Luis Nobre and kindly provided to me for this project.

2.4.2 shRNA Knockdown Cell Lines

Production of ECE1 and TMEM123 shRNA Knockdown Plasmids and Cell Lines

Short hairpin RNA (shRNA) sequences to target the *ECE1* or *TMEM123* genes were designed and validated by Sigma (Table 2.1), with the exception of ECE1_sh3, which was adapted from a published siRNA sequence (Lambert *et al*, 2008). The target sequence was inserted into a template to produce a pair of partially complementary oligonucleotides, which included BamH1 and EcoR1 restriction sites for cloning, and a stem loop to form the necessary structure. Oligonucleotide pairs were ordered from Sigma (Appendix Table 1). These were cloned into the pHR-SIREN vector (a kind gift from Professor Paul Lehner) by digestion and ligation.

The pHR-SIREN vector (2 µg) was digested with BamH1 and EcoR1 restriction enzymes for 2 h at 37 ºC, before dephosphorylating at 30 mins at 37 ºC to prevent self-ligation. This was then heat inactivated for 5 mins at 70 ºC. The digested and dephosphorylated vector was column purified using QIAquick PCR purification kit (Qiagen), following the manufacturer's instructions.

Table 2.1 Target sequences for shRNAs

Name	Target sequence
Control_sh1	GTTATAGGCTCGCAAAAGG
Control_sh2	GGCATATAACTATTTAGGTAT
Control_sh3	CGTGATCTTCACCGACAAGAT
Sp100_sh1	CGCTAGGAAGCCAACAAACAA
TMEM123_sh1	CCACACAACTCCAGTGCTAAC
TMEM123_sh2	GGGATGGTCTCAACAAATATG
ECE1_sh1	GCAGTTCCAGACCTCTACTTT
ECE1_sh2	GCCGATGAGAAGTTCATGGAA
ECE1_sh3	CTTCCACAGCCCCCGGAGT
ECE1_sh4	GCCTTAAACTTTGGTGGCATA
ECE1_sh5	TGTCTATGTCAGTGCCGATTC
ECE1_sh6	GATCAATGAATCCGAGCCTAT

To produce the insert, the forward and reverse oligonucleotides were first diluted to 100 µM. 1 µl of each of these was combined with 1 µl 10x T4 ligase buffer (New England Biolabs), 0.5 µl T4 polynucleotide kinase and made up to 10 µl total with H_2O. Phosphorylation was for 30 mins at 37 ºC, and the oligonucleotides were then annealed for 10 mins at 95 ºC before slowly cooling to room temperature.

The insert was diluted 1:10 for ligation, and 0.5 µl of this was combined with 2 µl (~50ng) of the dephosphorylated vector, 1 µl T4 ligase buffer, 0.5 µl T4 ligase (New England Biolabs) and made up to a total of 10 µl with H_2O. The insert was ligated into the vector at 4 ºC overnight.

5 µl of the ligation product was transformed into competent cells and amplified as described in section 2.4.5. The shRNA expression vector was sequenced using the U6P primer confirming presence of the insert (Table 2.3), and used to transduce either HFFF-TERTs or THP-1s as described in section 2.4.7.

Other shRNA Knockdown Cell Lines

Control shRNA and Sp100 shRNA cell lines in HFFF-TERTs were kindly provided by Dr Luis Nobre and Dr Katie Nightingale (Weekes lab, Department of Medicine, University of Cambridge). The plasmids were produced and the cells transduced in a similar manner to that described above, with the target sequences given in Table 2.1. They also provided the plasmids for generation of control shRNA THP-1 cell lines, and the HFFF-TERTs used in the final restriction and plaque assays.

2.4.3 CRISPR Cell Lines

Polyclonal CRISPR Populations

Knockout cells were produced using the clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9 systems as previously described (Sanjana *et al*, 2014; Shalem *et al*, 2014). Five guide sequences specific to *ECE1*, as well as guides for the restriction factor Sp100 and non-targeting control guides were designed based on sequences from the Sabatini single guide RNA library (Table 2.2) (Wang *et al*, 2015). Pairs of oligonucleotides for each guide RNA (gRNA) were ordered from Sigma (Appendix Table 1). The CRISPR cell lines were produced in collaboration with Dr Katie Nightingale and Dr Ben Ravenhill (Weekes lab, Department of Medicine, University of Cambridge).

The pairs of oligonucleotides were phosphorylated and annealed as described in section 2.4.2 and diluted 1:250; 2 µl of this was used for a digestion-ligation reaction. The oligonucleotide duplex was incubated with 100 ng of the pKLV expression vector (pKLV-U6gRNA(BbsI)-PGKpuro2ABFP, Addgene Plasmid #50946, (Koike-Yusa *et al*, 2014)), 1 µl 10x Tango buffer, 1 µl Dithiothreitol (DTT), 1 µl adenosine triphosphate, 0.5 µl FastDigestBbsI (Fermentas), 0.25 µl T4 ligase and made up to 10 µl with H_20. This

was incubated in a thermocycler for 5 mins at 37 ºC followed by 5 mins at 23 ºC; this cycle was repeated 6 times. The product was amplified by transformation into competent cells as described in 2.4.5. Insertion of the correct guide sequence in the pKLV vector was confirmed by sequencing with the U6P primer (Table 2.3). HFFF-TERTs stably expressing Cas9 (pSpCas9(BB)-2A-GFP (PX458), Addgene Plasmid #48138; cells kindly provided by Dr Katie Nightingale) were transduced with the gRNA as described in section 2.4.7, and subjected to puromycin selection for two days, to produce polyclonal CRISPR knockout populations, used for an initial HCMV restriction assay.

In order to avoid integration of the gRNA, a second transduction was performed in the presence of 2 μM raltegravir. Cells were maintained in media containing raltegravir for one week following selection. Immunoblot was used to confirm knockdown of the gene of interest in these polyclonal populations, and cell lines expressing the two guides that resulted in the best knockdown were chosen for single cell cloning.

Table 2.2 CRISPR guide sequences

Name	Target sequence
Control g1	GTACAGCTAAGTTAAACTCG
Control g2	GATGTCCGTTGTAGTCCTCG
gSp100_1	GCAGCCTGTCATCTACACCC
gSp100_2	TGAGATGGGGAACCCGAAGG
gECE1_1	GGAAACCCGAAAATCAGCCA
gECE1_2	GAACTGGGTGAAGAAGAACG
gECE1_3	CATGGAGCTCAAGATGGAGC
gECE1_4	GAGCTCCATGGACCCCACAG
gECE1_5	CTGGGGAAGCTGCTGGGCGG

Single Cell Cloning to Generate Monoclonal CRISPR Populations

The chosen polyclonal cell populations were seeded at 0.5 cells/well in a 96-well plate in DMEM (10 % FBS, 1 % pen/strep). As the cells grew, the media was replenished with a 1:1 mixture of fresh DMEM (FBS, pen/strep) and conditioned media. Conditioned media

was produced by filtering the media from a sub-confluent flask of WT HFFF-TERTs through a 0.22 μm filter. The cells were progressively transferred to a 24-well and then 6-well plate, before testing for a knockout monoclonal population by immunoblot.

2.4.4 shRNA Knockdown and Overexpression in Two Colour System

For the two colour HCMV restriction assay system, shRNA or overexpression plasmids were transduced into HFFF-TERTs already expressing either mCherry (controls) or iRFP (test gene). These mCherry and iRFP cells were kindly provided by Alice Fletcher-Etherington (Weekes Lab, Department of Medicine, University of Cambridge). Alice produced these by replacing the puromycin resistance gene in the pHAGE-P_{SFFV} vector with blasticidin resistance, and then cloning in either mCherry or near-infrared fluorescent protein (iRFP) using the gateway system. Following transduction, the cells were selected using 10 μg/ml blasticidin (TOKU-E). The mCherry or iRFP expressing HFFF-TERTs could then be additionally transduced with either the pHR-SIREN vector expressing an shRNA or the original puromycin resistant pHAGE-P_{SFFV} plasmid overexpressing a gene of interest, and selected again using hygromycin or puromycin respectively. The shRNA and overexpressing plasmids were as described previously.

2.4.5 Transformation

The entire LR reaction product or 5 μl of the ligated plasmid mix was added to 20 μl of Alpha-Select Silver Efficiency or Bronze Efficiency Competent E. coli cells (Bioline). This was incubated on ice for 15 mins, followed by heat shock for 45 seconds at 42 ºC. 200 μl lysogeny broth (LB) was added and this was incubated in a shaker for 1 h at 37 ºC, or immediately plated out onto LB agar plates containing ampicillin. Plates were inverted and incubated overnight at 37 ºC. Colonies were picked and added to 4 ml LB containing ampicillin, and incubated in a 37 ºC shaker overnight. Plasmids were purified from 2 ml of the overnight culture using the QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's instructions.

2.4.6 Confirmation of Insert

Following purification of the plasmid, insertion of the correct sequence in the lentiviral expression vector was confirmed by sequencing using the primers described in Table 2.3.

Table 2.3 Sequencing primers to confirm inserts in lentiviral expression vectors

Primer	Sequence	Vector	Use
U6P	GGGCAGGAAGAG GGCCTAT	pHR-SIREN, pKLV	shRNA knockdown, CRISPR
SFFVp	CGCGCCAGTCCTCCGATTG	pHAGE-P_{SSFV}	Overexpression
pHAGE-Rev	GCTTCGGCCAGTAACGTTA	pHAGE-P_{SSFV}	Overexpression
ECE1_int	AACGAGAAGGTGCTGACCGG	N/A	ECE1 overexpression

2.4.7 Stable Cell Line Production by Lentiviral Transduction

Transfection of HEK293Ts

HEK-293T cells were seeded at $1.5x10^5$ cells/well in a 12-well plate, in 1 ml culture media; 24 h after plating, cells were 70-80 % confluent and ready to transfect. Packaging plasmids used for transfection were either a four helper plasmid mix (VSVG, TAT1B, MGPM2, HCMV-Rev1B) or VSV-g / pCMV (using 200 ng VSV-g and 400 ng pCMV.DR8.91, a kind gift from Professor Paul Lehner). The helper plasmids were mixed with 500 ng lentiviral expression plasmid and diluted in 100 µl of pre-warmed OptiMEM (Gibco), with 3 µl TransIT-293 transfection reagent (Mirus). This mixture was incubated for 30 mins at room temperature before being added dropwise onto the cells. At 24 h the media was carefully replaced with fresh DMEM.

Transduction of HFFF-TERTs

HFFF-TERTs (WT or expressing mCherry or iRFP) were seeded at $1.5x10^5$ cells/well in a 12-well plate at 24 h after transfection of the HEK-293Ts. At 48 h after transfection, the lentiviral supernatant was passed through a 0.22 µm filter onto the fibroblasts.

Transduction of THP-1 Cells

THP-1s were seeded at $2x10^6$ cells/well in a 24-well plate, in 0.5 ml culture media. At 48 h after transfection, the lentiviral supernatant from the HEK-293Ts was passed through a 0.22 μm filter onto the THP-1 cells, and the media pipetted up and down several times. Cells were rocked in an incubator (37 ºC, 5 % CO_2) for 3.5 h. The cells were then pipetted up and down 8-10 times, and transferred into a 6-well plate containing an additional 2 ml media.

Selection

Cells were placed in selection using either 1 ug/ml puromycin (Acros Organics) or 50 ug/ml Hygromycin (TOKU-E) at 48–72 h after transduction, and maintained in selection media for between 2 days and 2 weeks as required (Table 2.4).

Table 2.4 Plasmids for production of cell lines

The vectors used to transduce various cell types are given. The cell lines were all produced in either WT HFFF-TERTs (wt), HFFF-TERTs already transduced with mCherry of iRFP for use in the two colour restriction assays (two colour), HFFF-TERTS already transduced with Cas9 for production of CRISPR cells (Cas9), or THP-1 cells. For the helper plasmids, '4 plasmid' refers to VSVG, TAT1B, MGPM2, HCMV-Rev1B, and VSV-g/pCMV indicates 200 ng VSV-g and 400 ng pCMV.DR8.91 were used. Either 1 ug/ml puromycin (Acros Organics) or 50 ug/ml Hygromycin (TOKU-E) was applied for selection.

Cell Types	Vector	Parent Cell	Helper Plasmids	Selection
shRNA knockdown	pHR-SIREN	HFFF (wt) HFFF (two colour) THP-1	4 plasmid VSV-g/pCMV VSV-g/pCMV	Hygromycin
Overexpression	pHAGE-P_{SFFV}	HFFF (wt) HFFF (two colour)	4 plasmid VSV-g/pCMV	Puromycin
CRISPR	pKLV	HFFF (Cas9)	VSV-g/pCMV	Puromycin

2.5 Molecular Biology Techniques

2.5.1 Flow Cytometry

Cell Surface Flow Cytometry

FCM was used to confirm enrichment and validate selected hits. Cells were washed in FCM buffer (PBS and 2 % FBS), prior to blocking in Human Trustain (Biolegend, 1:20 in FCM buffer) for 10-20 mins at room temperature. All staining was for 30 mins at 4 ºC in the dark, using primary antibodies diluted in FCM buffer as indicated in Table 2.5. All samples were then washed in FCM buffer (400 g, 5 mins), and if required a secondary antibody was applied (1 h, 4 ºC, in the dark) diluted in FCM buffer as described in

Table 2.6, before washing again. Samples were fixed in 4 % paraformaldehyde (10 mins, room temperature, Biolegend), and resuspended in PBS for analysis.

Intracellular Flow Cytometry

Cells were washed in FCM buffer and then fixed in 4 % paraformaldehyde for 15 mins at room temperature. The samples was then centrifuged, the supernatant poured off and cells resuspended in MeOH prior to incubation on ice for 15 mins to permeabilise. Cells were washed in FCM buffer and then blocked and stained as described for cell surface FCM, omitting the final fixation step.

Flow Cytometry Data Acquisition and Analysis

Samples were analysed on the Becton-Dickinson FACSCalibur flow cytometer (confirmation of enrichment of primary cells, ECE1 and TMEM123 antibody testing) or the Becton Dickinson LSR Fortessa (validation of other IFN stimulated proteins by FCM). Data was analysed using FlowJo vX software.

Table 2.5 Table of primary antibodies used for flow cytometry experiments

Target	Conjugated Fluorophore	Company	Catalogue Number	Species	Clone	Dilution
CD14	PE	Biolegend	301850	Mouse	M5E2	1:20
CD4	APC	Biolegend	317415	Mouse	OKT4	1:20
BST2	PE	Biolegend	348405	Mouse	RS38E	1:20
CD69	PE	Biolegend	310905	Mouse	FN50	1:20
CD38	PE	Biolegend	356603	Mouse	HB-7	1:20
CD40	PE	Biolegend	334307	Mouse	5C3	1:20
Siglec1	APC	Biolegend	346007	Mouse	7-239	1:20
CD274	PE	Biolegend	329705	Mouse	29E.2A3	1:20
SLAMF7	-	Santa Cruz	sc-53577	Mouse	162.1	1:50
ECE1	-	R&D Systems	MAB17841	Rat	303913	1:100
ECE1	-	Invitrogen	PA5-24563	Rabbit	-	1:50
ECE1	-	Abcam	ab71829	Rabbit	-	1:50
TMEM123	-	Abcam	ab81423	Rabbit	-	1:50
TMEM123	-	Santa Cruz	SC-377295	Mouse	G-2	1:50
TMEM123	-	Novus	NB100-56371	Rabbit	-	1:50
TMEM123	AF647	R&D Systems	FAB3010R	Mouse	297617	1:12.5

Table 2.6 Table of secondary antibodies used for flow cytometry experiments

Target	Conjugated Fluorophore	Company	Catalogue Number	Species	Dilution
anti-Mouse	AF647	Invitrogen	A21236	Goat	1:1000
anti-Rat	AF647	CST	4418S	Goat	1:1000
anti-Rabbit	AF647	Invitrogen	A31573	Donkey	1:500

2.5.2 Immunoblotting

Except where otherwise stated, the following method was used for immunoblotting. Cells were lysed in radioimmunoprecipitation assay (RIPA) buffer (Cell Signalling Technology) containing cOmplete Mini Protease Inhibitor Cocktail (Roche; 1 tablet per 10 ml RIPA) for 15 minutes at 4 ºC, and supernatants were then clarified by centrifugation at 14,000 g for 10 mins. The Pierce bicinchoninic acid (BCA) protein assay kit (ThermoFisher) was used to measure protein concentration according to the manufacturer's instructions. Samples were denatured and reduced with 6× protein loading dye (375 mM Tris pH 6.8, 12 % SDS, 30 % glycerol, 0.6 M DTT, 0.06 % bromophenol blue) for 5 mins at 95 ºC. 50 µg of protein was separated by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) using Mini-Protean TGX precast gels (Bio-Rad), then transferred to polyvinylidene difluoride (PVDF) membranes (0.45 µm pore) using the TransBlot SD Semi-Dry Transfer System (Bio-Rad). Transfer was for 45 mins at 20 v. The membrane was blocked with 5 % milk dissolved in tris-buffered saline with 0.2 % tween (TBST) before probing overnight at 4 ºC with the appropriate primary antibodies (described in Table 2.7). The following day, the membrane was washed 3 times in TBST, for 5 mins each time. It was then rocked for 1 h at room temperature with the secondary antibodies: IRDye 680RD goat anti-mouse (LI-COR, 925-68070), IRDye 800CW goat anti-rabbit (LI-COR, 925-32211), or IRDye 800CW goat anti-rat (LI-COR 926-32219), all diluted 1:10,000 in TBST. After a further 3 x 5 min washes in TBST, fluorescent signals were detected using a LI-COR Odyssey CLx, and images were processed using Image Studio Lite (LI-COR).

The protocol was modified for testing of antibodies for TMEM123 as indicated in the figure legends. In some cases, lysis was in 2 % SDS (20 % SDS (Sigma) diluted in PBS), with Benozonase (1:50). Lysis proceeded for 30 mins in a 37 ºC shaker, before sonicating for 5 mins and clarifying at 14,000 rpm for 10 mins at room temperature. Additionally, samples were sometimes mixed with loading dye and then heated at either 95 ºC for 5 mins, 60 ºC for 20 mins or 37 ºC for 30 mins, as indicated, and the stated amounts of proteins were run on the gel. Where indicated, transfer to the PVDF membrane was using the Trans-Blot wet transfer system (Bio-Rad), at 100 v for 1 h 15 mins.

Table 2.7 Antibodies used for immunoblots

Target	Company	Catalogue Number	Species	Dilution
GAPDH	R&D Systems	MAB5718	Mouse	1:10,000
ECE1	Abcam	ab71829	Rabbit	1:1000
ECE1	R&D	MAB17841-SP	Rat	1:500
ECE1	Proteintech	N/A (was still in testing)	Rabbit	1:500
ECE1	Invitrogen	PA5-24563	Rabbit	1:1000
TMEM123	Abcam	ab81423	Rabbit	1:500
TMEM123	Santa Cruz	SC-377295	Mouse	1:500
TMEM123	Novus Biologicals	NB100-56371	Rabbit	1:250
Sp100	GeneTex	GTX131570	Rabbit	1:2000
Daxx	BioRad	MCA2143	Mouse	1:1000
Calnexin	LifeSpan Bio	LS-B6881	Rabbit	1:10,000

2.5.3 RT-qPCR

Synthesis of cDNA

Reverse transcription-quantitative PCR (RT-qPCR) was used for validation of proteomic data. Total RNA was extracted from cells using the RNeasy Mini Kit (Qiagen), following the manufacturers instructions. Contaminating DNA was removed from 1 µg RNA by combining 5 µl 10x Turbo DNase buffer (Invitrogen), 1 µl DNase and making up to 50 µl with H_20. This was incubated at 37 ºC for 20-30 mins, and then 5 µl of DNase Inactivation Reagent added. After 5 mins at room temperature, the DNase treated RNA was centrifuged at 10,000 g for 1.5 mins, and the RNA transferred to a new tube for reverse transcription.

cDNA was synthesised from 100 ng of the DNase treated RNA, using the GoScript Reverse Transcription Kit (Promega), with 1 µl oligo(dT) primer and made up to 5µl with water. This was heated for 5 mins at 70 ºC, and then chilled on ice for 10 mins. The reverse transcription mix (4 µl GoScript 5x reaction buffer, 2 µl 25mM MgCl2, 1 µl PCR nucleotide mix, 0.5 µl recombinant RNasin ribonuclease inhibitor, 1 µl GoScript reverse transcriptase and 6.5 µl water) was added, and annealed for 5 mins at 25 ºC prior to

extension for 1 h at 42 ºC. The reverse transcriptase was inactivated by heating for 15 mins at 70 ºC.

qPCR Analysis

The cDNA was diluted 1:10 and analysed using the TaqMan quantitative PCR (qPCR) system. Each sample was analysed in triplicate for both the gene of interest and the internal control glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*). Control samples were included where either water or RNA prior to reverse transcription was used instead of the sample. A mastermix was made up including 10 µl TaqMan gene expression master mix (Applied Biosystems), 1 µl of the relevant TaqMan probe (ThermoFisher) for either *ECE1* (Hs01043735_m1), *TMEM123* (Hs00920881_m1) or the internal control *GAPDH* (Hs02786624_g1), and 7 µl H2O, for each well required. 18 µl of this mix was added to each well in the 96-well plate (Applied Biosystems), and 2 µl of the relevant sample added to the well. Analysis was performed on the 7500 Fast and 7500 real-time PCR systems (Applied Biosystems). The PCR program started with activation at 95 ºC for 2 min, followed by 40 cycles of denaturation at 95 ºC for 5 s and annealing/extension at 60 ºC for 30 s.

2.6 Viruses

The HCMV with green fluorescent protein (GFP) fused to UL36 used in the restriction assays was as described in Nightingale *et al.* (2018). It is derived from the Merlin strain of the virus, which is the designated HCMV reference sequence. This strain was previously cloned into a bacterial artificial chromosome (BAC), which was found to differ to the original clinical isolate of the virus by point mutations in RL13 and UL128. These mutations are acquired rapidly during passaging, and enhance replication in fibroblasts (Stanton *et al*, 2010). The mutation in UL128 has since been corrected, however as expression of this impairs growth of the virus, a tetracycline (tet)-operator was inserted upstream of the gene to allow expansion. The BAC clone was further modified for expression of GFP from a porcine teschovirus-1 2A (P2A) self-cleaving peptide following the UL36 open reading frame. UL36-GFP HCMV stocks were

prepared by Dr Katie Nightingale and Alice Fletcher-Etherington, and kindly provided to me for the restriction assays. Briefly, HFFF-Tet cells were infected at low MOI, enabling tet repression of UL128 for expansion, and the virally infected supernatant was collected and centrifuged, first to pellet cell-free virus, and then to remove debris. For the plaque assay, AD169-GFP (RCMV288) was used (McSharry *et al*, 2001), and this was prepared by Dr Katie Nightingale.

2.7 HCMV Restriction Assays

2.7.1 shRNA, Overexpression and CRISPR Cell Restriction Assays

In order to examine whether a particular gene had an impact on infection, knockdown (shRNA and CRISPR) and overexpressing HFFF-TERT cell lines were produced, in this case for ECE1. Control cells were transduced with non-targeting shRNAs, CRISPR guides or a GAW plasmid, and the known restriction factors Sp100 and Daxx were used as positive controls. The cell lines were generated as described in section 2.4.

Cells were seeded at $1.5x10^5$ cells/well in a 24-well plate, with each of the control, Sp100 and/or Daxx, and ECE1 cell lines seeded in triplicate for each MOI to be examined (with the exception of the first two shRNA experiments which were performed in duplicate). At 24 h, cells were washed once with PBS and then infected with UL36-GFP HCMV at a desired low MOI (0.003 – 0.1) in 150 µl of serum free DMEM. Infection proceeded at 37 ºC for 2 h on a rocker, before the infection media was removed and replaced with DMEM supplemented with 10 % FCS.

At 24 hpi the samples were harvested for analysis. The media was removed and cells washed once in PBS. They were then incubated with 0.2 µl trypsin (Sigma) until detached, then 0.2 µl of DMEM and 0.2 µl of 4 % paraformaldehyde fixation buffer (Biolegend) were added. The cells were transferred to eppendorfs, and after 10 mins centrifuged at 500 g for 5 mins. Cells were resuspended in 300 µl PBS and analysed on the Becton-Dickinson FACSCalibur flow cytometer, in order to measure the percent of GFP positive, and therefore infected, cells. Data was analysed using FlowJo vX.

2.7.2 Two Colour Restriction Assay

HFFF-TERTs-mCherry cells expressing either the GAW overexpression control or a control shRNA were mixed 1:1 with HFFF-TERTs-iRFP cells expressing the relevant overexpression plasmid or shRNA for testing (ECE1, Sp100 or Daxx). The mixed population of cells was seeded at $1.5x10^5$ cells/well in triplicate for each MOI to be examined. The infection and harvesting of the restriction assay was performed as described for other restriction assays (section 2.7.1). The % GFP+ cells in both the mCherry and iRFP populations in each well was measured using the Becton Dickinson LSR Fortessa and analysed using FlowJo vX. The data was analysed by taking the ratio in a given well of:

$$\frac{\%\ \text{GFP+ cells in the iRFP population}}{\%\ \text{GFP+ cells in the mCherry population}}$$

This was then averaged across the triplicate samples for each cell type.

2.7.3 Plaque Assay

HFFF-TERT cell lines (control, knockdown or overexpression, as appropriate) were seeded at $1.5x10^5$ cells/well in a 12-well plate, in triplicate for each cell type and for each dilution of the virus. After 24 h they were they were infected with AD169-GFP at an MOI of 0.002, 0.006, 0.02 and 0.06, in serum free DMEM, and infections were in 400 µl/well. Infection proceeded for 2 h on a rocker at 37 ºC. After this time, the inoculum was removed and replaced with 3 ml of 1:1 Avicel (FMC BioPolymer, 2% w/v in H_2O) and 2x DMEM (125 ml H_2O, 50 ml 10x minimum essential media (MEM, Gibco), 50 ml FBS, 15 ml sodium bicarbonate (Gibco), 10 ml pen/strep, 5 ml glutamine (Sigma)). After two weeks at 37 ºC, with care not to move the cells at all during this time, the Avicel/DMEM mix was removed and the cells washed carefully with PBS three times. The cells were fixed with 0.5 ml 4 % paraformaldehyde in each well for 15 mins. This was replaced with PBS for storing the cells at 4 ºC, prior to counting the plaques.

2.8 Additional Methods from Chapter 4 on VACV

The work presented in 'Chapter 4: Vaccinia Virus Screen' was produced in collaboration with Dr Jonas Albarnaz and Dr Yongxu Lu in the laboratory of Professor Geoffrey L. Smith (Department of Pathology, University of Cambridge). Viral infections for the proteomic time course, and whole cell lysate (WCL) harvesting were performed by them, as well as the immunoblots for validation and the restriction assays with VACV and HSV-1. All further processing of samples for proteomics, and data analysis, was performed by myself. Brief methods are given here for experiments performed by members of the Smith Lab, and more details can be found in (Soday *et al*, 2019).

2.8.1 Cell Culture

HFFF-TERTs were obtained from the Weekes lab and maintained as described previously (section 2.2). Histone deacetylase 5 (HDAC5) knockout and overexpression cells were produced in U2OS (human osteosarcoma cells), HeLa and HEK-293T cells. VACV was titrated onto BS-C-1 cells (African green monkey, kidney epithelial cells). All were maintained in DMEM supplemented with FBS (10 % v/v) and pen/strep, at 37 ºC, 5 % CO_2. VACV was propagated in RK_{13} cells (rabbit kidney epithelial cells), which were maintained MEM supplemented with 10 % FBS and 1-2 % pen/strep.

2.8.2 Production of Cell Lines

HDAC5 knockout cell lines were produced by CRISPR/Cas9 in HeLa and HEK-293T cells. Cells were transfected with CRISPR guides designed to target exon 3 and 4 of *HDAC5* (Exon 3: AGGTCGGGAACCATCCTTGG, Exon 4: CCAGCCCTGTGGAGCTACGG). The polyclonal cell populations were single cell cloned. Two HDAC5 knockout clones were generated from HeLa parental cells with guides targeting either exon 4 (H5KO1) or exon 3 (H5KO2), and two HDAC5 knockout clones were generated from HEK-293T cells with both guides targeting exon4 (H5KO3 and H5KO4). Overexpression cell lines were produced in U2OS cells by lentiviral transduction, resulting in doxycycline inducible overexpression of HDAC5-Flag.

2.8.3 Viruses

For the proteomic screen of VACV infection and for validation immunoblots, the WT VACV western reserve (WR) strain was used. For the investigation of the viral C6 protein, a derivative of the mutant lacking the *C6L* gene was used (Unterholzner *et al*, 2011), and for restriction assays a virus with the capsid protein A5 fused to GFP was employed (Carter *et al*, 2003). The HSV-1 virus used in the restriction assay in this chapter was the s17 strain, with GFP fused to viral protein 26, and this was provided to the Smith lab by Prashant Desai (Hollinshead *et al*, 2012).

2.8.4 Infection

Proteomic Experiments

HFFF-TERTs were seeded at $1x10^6$ cells in a 25-cm^2 flask and infected at an MOI of 5. The time course was performed in biological triplicate, with mock and VACV infected samples harvested at the time points described. Where indicated, 40 μg/ml AraC was added to the cells at the time of infection. For the MG132 proteomic data, the infection was allowed to proceed for 2 h before the inocula was removed and replaced with medium with or without 10 μM MG132. The infected samples ± MG132 were both included in the experiment in triplicate, alongside triplicate infection with the C6 deletion mutant; mock infection ± MG132 was a single replicate in order to fit within the 11-plex.

Immunoblot

HFFF-TERTs were seeded at $1x10^6$ cells in a 6-well plate and infected at an MOI of 5.

Viral Replication Assays

For viral replication assays, $2x10^6$ WT or knock out cells were seeded in a 6-well plate and infected at an MOI of 0.001 with A5-GFP VACV. After two days, the infected cells and the supernatant were collected and titrated on BS-C-1 cells. For HSV-1 assays, supernatants were collected at three days post infection and titrated on U2OS cells.

2.8.5 Immunoblots

Immunoblots in chapter 4 (Quantitative Temporal Analysis of Vaccinia Virus Infection), were performed by Dr Jonas Albarnaz, in a similar manner to that described in section 2.5.2. In this case, lysis was performed in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 10 % glycerol, 0.5 % Triton X-100, 0.1 % NP-40, supplemented with protease and phosphatase inhibitor cocktails (Roche), followed by clarifying by centrifugation at 17,000 g for 15 mins at 4 ºC. Samples were reduced with 100 mM DTT in SDS-gel loading buffer for 5 min at 100 ºC, before separation on a gel and transfer using the Bio-Rad Trans-Blot Turbo transfer system. Primary antibodies used are as described in Table 2.8. Secondary antibodies used were IRDye 680RD-conjugated goat anti-rabbit IgG (cat. no. 926-68071), IRDye 680LT-conjugated goat anti-mouse IgG (926-68020), IRDye 800CW-conjugated goat anti-rabbit IgG (926-32211), IRDye 800CW-conjugated goat anti-mouse IgG (926-32210), IRDye 680LT-conjugated goat anti-rat IgG (926-68029), and the membrane was imaged using an Odyssey infrared imager (LI-COR Biosciences). For further details see (Soday *et al*, 2019).

Table 2.8 Antibodies used for immunoblots in chapter 4 (vaccinia virus screen)

Immunoblots performed by Dr Jonas Albarnaz or Dr Yongxu Lu (Smith Lab, Department of Pathology, University of Cambridge).

Target	Company	Catalogue Number	Species	Dilution
COL1A2	Santa Cruz	sc-376350	Mouse	1:500
COL6A2	Santa Cruz	sc-374566	Mouse	1:500
IFIT1	Thermo	PA3-848	Rabbit	1:1000
IFITM1-3	Santa Cruz	sc-374026	Mouse	1:500
TRIM5	Santa Cruz	sc-373864	Mouse	1:500
TUBULIN	Serotec	MCA77G	Rat	1:10,000
HDAC5	Santa Cruz	sc-133225	Mouse	1:500
HDAC1	Santa Cruz	sc-81598	Mouse	1:500
FLAG	Sigma	F3165	Mouse	1:1000
D8	Described in (Parkinson & Smith, 1994)		Mouse	1:1000
C6	Described in (Unterholzner *et al*, 2011)		Rabbit	1:1000

2.9 Proteomics

Multiplexed proteomic screens examined IFN stimulation and VACV infection (Table 2.9). For the IFN screens, this included samples of THP-1 cells ± IFN, as well as primary cells from five donors for monocytes ± IFN, and two donors of pan-monocytes ± IFN. In the case of CD4+ T cells, data was initially acquired on samples from three donors ± IFN, and this was later extended to five donors (experiments summarised in Table 2.9).

For the VACV screens, three identical time courses of WT VACV infection were performed, including three mock time-points, six infection time-points and one 6 h mock or infection with AraC treatment. An initial 10-plex time-course experiment was used for optimisation of time-points; this was performed similarly to the triplicate time-courses described in the methods here, but the final data was not extensively analysed and so it is not discussed further. An additional 11-plex investigating the effects of MG132 treatment, and deletion of the viral protein C6 was then conducted, which included triplicate WT infection, WT+MG132, and infection with a C6 deletion mutant. There were single replicates for mock ± MG132 (experiments summarised in Table 2.9).

In all cases, samples were initially examined as a 3 h (sometimes 1 h) single run, and some samples were then fractionated into simplified mixtures of proteins, and analysed in multiple 3 h runs to quantify additional proteins.

Table 2.9 Proteomic screens performed, and MS runs analysed in final dataset

Project	Experiment	Samples in Multiplex	Number of Single Shots	Number of Fractions
IFN (Plasma membrane proteomics, Chapter 3)	THP-1	2 plex	2	0
	Primary monocytes	10 plex	3	6
	Primary CD4+ T cells	6 plex	1	0
	Primary CD4+ T cells	10 plex	1	6
	Pan Monocytes	4 plex	2	0
VACV (Whole cell proteomics, Chapter 4)	Time course 1	11 plex	0	12
	Time course 2	11 plex	1	12
	Time course 3	11 plex	1	12
	MG132 and C6 mutant	11 plex	1	12

2.9.1 Protein Extraction

Plasma Membrane Protein Enrichment (IFN Screens)

PM profiling was used to enrich PM proteins in THP-1s and primary leukocytes, prior to analysis by mass spectrometry (MS) (data in chapter 3); this was performed as previously described (Weekes *et al*, 2010, 2012). All spins were at 4 ºC, and PBS containing calcium and magnesium (D8662, Sigma) was used for the biotinylation and enrichment; all reagents were kept ice cold.

The suspension cells were transferred to a 15 ml falcon tube and centrifuged to collect (400 g, 5 mins), then washed twice in cold PBS. Sialic acid residues were oxidised with sodium meta-periodate and biotinylated with aminooxy-biotin, with the reaction catalysed by Aniline. This was performed by removing the supernatant from the washed cells, wrapping the falcon tubes in foil and adding 3 ml of biotinylation mix to each (see below). The falcon tubes were rocked at for 30 mins at 4 ºC in the dark. The reaction was then quenched with 600 µl 5x glycerol solution (1M glycerol (Sigma) diluted 1:2000 in PBS), mixed by inversion and centrifuged (400g, 5 mins). The supernatant was removed and the cells washed twice in PBS, before being resuspended in 1 ml 1 % triton lysis buffer (see below) and transferred to an Eppendorf. Lysis proceeded on ice for 30 mins, and then the tube was centrifuged at 13,000 g for 5 mins and the supernatant transferred. Two further spins were performed, retaining the supernatant with a minimal amount of debris. The final supernatant was snap frozen in liquid nitrogen prior to pull-down and digestion.

Once multiple samples of biotinylated glycoproteins had been collected, they were enriched with high affinity streptavidin agarose beads (Thermo). The following method is described for 10 samples, and volumes were adjusted according to the number of samples. Streptavidin agarose beads were resuspended and 500 µl and added to a Poly-Prep column (Bio-Rad) on a vacuum manifold. The beads were washed 4 times with 800 µl cold 1 % triton lysis buffer, before resuspending in 1200 µl lysis buffer. 95 µl beads were added to each lysate, and they were incubated on a rotor for 75 mins at 4 ºC. Each sample was resuspended by pipetting several times and then transferred to an individual polyprep column on the vacuum manifold. Several rounds of washing were then

performed. In each case, the required volume was added to each polyprep column using a repeat pipettor, the wash reagent removed under vacuum and then this was repeated the required number of times. Cells were first washed with 4x 1 ml lysis buffer, and then 4x 1 ml 0.5 % SDS in PBS (diluted from 20 % SDS, Invitrogen) before incubation for 20 mins in 500 µl 0.5 % SDS/PBS/DTT (add 550 µl 1M DTT (Sigma) to 5ml 0.5 % SDS/PBS) in order to denature and reduce proteins. The samples were subsequently washed with 8x 1.5 ml urea solution (see below), once with 10 ml urea solution, and then alkylated by incubating for 20 mins in the dark with 500 µl urea solution /iodoacetamide (IAA) (5ml urea solution with 500 µl 500 mM IAA dissolved in urea solution). A further 2x 1.5 ml washes were performed with urea solution, 3x 1.5 ml washes with ultra pure H_2O (JT Baker), and finally one 10 ml wash with H_2O. The samples were transferred in 2x 300 µl H_2O into a screw cap column (Pierce). The column was centrifuged to remove the H_2O at 2000 g for 1 min, and 40 µl trypsin / HEPES solution (one vial of Promega trypsin dissolved in 1 ml 200 mM hepes pH 8.5) added for on-bead digestion. After 3 h in a 37 ºC shaker, spin columns were placed in fresh eppendorfs and centrifuged at 2000 g for 1 min. The tryptic peptides were stored at -80 ºC prior to TMT labelling.

Reagents for PMP:

Biotinylation mix - for two samples: 25-40 mg sodium periodate (Thermo) was dissolved in ice-cold PBS pH 6.7 (weight in mg x 0.47 ml, to make 10 mM solution), keeping on ice and agitating periodically. In a foil covered tube 6 ml PBS at pH 6.7, 6.7 µl aminooxy-biotin (Biotium) and 6.1 µl aniline (Acros Organics) were combined, and mixed immediately. 660 µl of the sodium periodate was then added to this.

500 mM IAA – The weight of IAA (Sigma) in mg was multiplied by 10.8 µl and the IAA was diluted in this volume of lysis buffer, urea solution or HEPES as indicated.

1 % Triton Lysis Buffer - For 50 ml lysis buffer, 500 µl 1M Tris-HCL (Sigma), 25 ml 2 % Triton (Thermo, one vial 10 % Trition X-100 added to 40 ml ultrapure H_2O), 1.5 ml 5M NaCL (Sigma) and 23 ml ultra pure H_2O were combined. 100 µl 500 mM IAA (diluted in lysis buffer), was added to 10 ml lysis buffer with one protease inhibitor tablet.

Urea Solution – 90 g urea was dissolved in 25 ml 1M Tris-HCL (Sigma) and 156.5 ml ultra pure H_2O (JT Baker).

Whole Cell Protein Extraction (VACV Screens)

For the proteomic studies of VACV infection (chapter 4), WCL protein extractions were performed by Dr Jonas Albarnaz. Briefly, cells were washed twice with PBS and then scraped in 250 µl guanidine lysis buffer (200mM HEPES pH 8.4 is produced from a 1M HEPES stock (Sigma), and 3 ml of this is added to 9 ml 8 M guanidine-HCL (Pierce), to make 6M guanidine / 50 mM HEPES pH 8.5 lysis buffer). Cells were collected in eppendorfs, centrifuged briefly and sonicated prior to centrifugation at 21,000 g for 10 mins to remove debris. The supernatant was transferred, centrifuged again and the second supernatant snap frozen in liquid nitrogen. These samples were then given to me for further processing.

A portion of the lysate (50 µl) was digested for proteomic analysis, and the rest retained for immunoblot analysis where required. Proteins were reduced and then alkylated by incubating for 20 mins with 2.5 µl 100 mM DTT and then for an additional 20 mins in the dark with 1.5 µl 10x IAA (dissolved in 200 mM HEPES). Excess IAA was quenched by incubating for 15 mins with a further 2.5 µl 100 mM DTT and samples were diluted in 143.5 µl 200 mM HEPES to achieve a final concertation of 1.5 M guanidine. Samples were then digested in 2 µl endoproteinase LysC (Wako) for 3 h at room temperature. They were then further diluted in 398 µl HEPES to a final concentration of 0.5 M guanidine, prior to adding 50 µl trypsin (Pierce), and incubating in a 37 ºC shaker overnight. The next day, the reaction was quenched with 65 µl 50 % formic acid (FA, Sigma, diluted to 50 % in H_2O) to achieve a concentration of 5 % FA, and centrifuged at 21,000 g for 10 mins to remove undigested protein.

Peptides were then subjected to C18 solid-phase extraction (Sep-Pak, Waters). A 50 mg SepPak was activated with 2x 1 ml 100 % acetonitrile (AcN, Merck) followed by 1 ml 70 % AcN / 1 % FA. The column was washed with 3x 1 ml 1 % FA. Samples were diluted with 300 µl 1 % FA, loaded onto the SepPak and allowed to pass through slowly. The SepPak was washed again with 3x 1 ml 1 % FA prior to elution of the peptides in 2x 350 µl 70 % AcN / 1 % FA. Samples were dried completely by vacuum centrifugation in a

speedvac and resuspended in 75 µl 200 mM HEPES pH 8.5. The amount of peptides was quantified using a micro BCA protein assay kit (Pierce) following the manufacturer's instructions.

2.9.2 Peptide Labelling with Tandem Mass Tags

TMT reagents (0.8 mg, Thermo) were dissolved in 43 µl anhydrous AcN (Acros Organics) Each tag was aliquoted and stored at -80 ºC. The general protocol for TMT labelling is given here, and more details are below for particular experiments; details of sample labelling are given in Table 2.10 and Table 2.11. The peptide samples were diluted into HEPES and AcN such that the final concentration with the TMT label was 30 % AcN (v/v). Labelling proceeded for 1 h at room temperature, before quenching by adding 5 % hydroxylamine (Sigma) to a final concentration of 0.5 %. The required amount of sample was combined, 1:10 vol 50 % FA added, and diluted in 1 % FA. This was then desalted using a StageTip, or SepPak prior to analysis by MS or fractionation. In general, a small amount of the sample was analysed on the mass spectrometer as single run to confirm labelling efficiency and that there was an equal amount of protein in each channel, prior to fractionation.

Plasma Membrane Samples (IFN Screens)

No BCA was performed for quantitation, as the amount of protein is lower in PM enriched samples and is below the limit of detection. For the THP-1 and pan-monocyte samples, 7 µl tryptic peptides was combined with 4 µl HEPES and 5 µl TMT to make a final concentration of ~30 % AcN. Following 1 h of labelling, the reaction was quenched with 5 % hydroxylamine. The required samples were combined in a 1:1 (THP-1) or 1:1:1:1 (pan-monocytes) ratio and 1:10 volume 50 % FA added. The combined samples were dried to < 40 µl, diluted in 200 µl 1 % FA and subjected to desalting by StageTip (Rappsilber *et al*, 2007). The StageTip was made by punching 2x 3 layers of C18 material and compacting it in a 200 µl pipette tip. The tip was conditioned by applying 50 µl methanol and centrifuging at 1800 g for ~1 min (or until all the liquid had gone through). The StageTip was rinsed with 50 µl 70 % AcN/1 % FA followed by 50 µl 1 % FA. The

~200 µl peptide was loaded and the liquid spun through over several minutes. The C18 with the peptide bound was again rinsed with 2x 50 µl 1 % FA, and then the peptide eluted in 50 µl 70 % AcN/1 % FA. The desalted sample was dried completely, resuspended in 10 µl MS buffer (86 µl H2O, 10 µl 50 % FA, 4 µl 100 % AcN), and analysed on the mass spectrometer, initially for a 1 h single run, and then a 3 h run. These samples were not fractionated, they were only examined as single runs.

For the CD14+ monocytes, 28 µl each sample was combined with 9 µl AcN and 3 µl TMT. Following labelling, 2 µl of each sample was quenched and combined. This was desalted using a StageTip (as above), dried down and resuspended in MS buffer for single shot analysis on the mass spectrometer. Analysis of the single shot showed the percent of peptides labelled with a TMT was 89 %, slightly lower than expected. As sufficient labelling is required to ensure the peptides can be assigned to the correct sample, greater than 90 % labelling is usually desired for PM samples, so the 38 µl of remaining samples was added to 3 µl TMT and again labelled for 1 h, then quenched with 5 % hydroxylamine. 3 µl of each sample was combined and processed as previously, and single shot analysis confirmed improved labelling. Similar amounts of protein were quantified with each label, so there was not an excessive amount of electronic normalisation required. 25 % of the remaining sample was combined for a final single shot, and 75 % was combined in a 1:1:1:1:1:1:1:1:1:1 ratio and subjected to C18 solid phase extraction (SepPak, as above), prior to solid-phase cation exchange (SCX) fractionation.

For the CD4+ T cells, a small amount of the samples from three donors were initially labelled and combined in a 6-plex for single shot analysis (protocol as described for THP-1s and pan-monocytes). When samples were collected from two additional donors, all samples were labelled and a small amount of each combined in equal volumes. Single shot analysis was used to adjust for equal loading, and the remaining sample combined 1:1:1:1:1:1:0.8:0.8, subjected to C18 solid-phase extraction with a SepPak (as above) and SCX fractionation.

Table 2.10 Sample labelling for PM proteomics samples

Samples were labelled with TMT as indicated in the table. Dx represents donor x.

TMT	Thp-1	CD14+ Monocytes	CD4+ T cells (6 plex)	CD4+ T cells (10 plex)	Pan Monocytes
126	-	D1 + IFN	D6 + IFN	D6 + IFN	-
127N	-	D1	D6	D6	-
127C	-	D2 + IFN	-	D7 + IFN	D11 + IFN
128N	-	D2	D7 + IFN	D7	-
128C	-	D3 + IFN	-	D8 + IFN	D11
129N	-	D3	D7	D8	-
129C	Thp-1 + IFN	D4 + IFN	-	D9 + IFN	D12 + IFN
130N	-	D4	D8 + IFN	D9	-
130C	Thp-1	D5 + IFN	-	D10 + IFN	D12
131N	-	D5	D8	D10	-

Whole Cell Protein Samples (VACV Screen)

50 µg of protein was labelled, and made up to 35 µl with HEPES. 12 µl AcN was added, and 3 µl TMT. Following labelling for 1 h, 1 µl of each sample (~1 µg) was added to 180 µl 1 % FA and 1 µl 50 % FA. The combined sample was desalted with a StageTip (as above), dried down in a speedvac and resuspended in 10 µl MS buffer for a single shot analysis. This showed the amount of protein in each sample of the WCL1 time course was approximately equal, apart from sample one, in which ~20 % less total signal was detected. Samples were therefore combined in a 1.2:1:1:1:1:1:1:1:1:1:1:1 ratio to avoid excessive digital normalisation. For this reason, the single shot data was not later analysed with the fractionated data for this sample. For time-courses WCL2 and WCL3, and the MG132 / C6 experiment, near equal amounts of protein where present with each label. Therefore, the remainder of the samples were quenched with 5 µl 5 % hydroxylamine and combined in a 1:1:1:1:1:1:1:1:1:1 ratio; the single shot and fractionation data were later analysed together for these samples. The combined sample was dried down to < 100 µl, diluted in 1.5 ml 1 % FA and subjected to C18 solid-phase extraction with a SepPak (as above). The eluted peptide was then dried down completely prior to offline high pH reversed-phase (HpRp) fractionation.

Table 2.11 Sample labelling for WCL proteomic samples

TMT label	WT VACV Time Course 1, 2 and 3	wtVACV / vDC6 / MG132
126	0h Mock	Mock
127N	2h VACV	Mock + MG132
127C	4h VACV	wtVACV 1
128N	6h Mock	wtVACV 2
128C	6h Mock + AraC	wtVACV 3
129N	6h VACV	VACV ΔC6 1
129C	6h VACV + AraC	VACV ΔC6 2
130N	8h VACV	VACV ΔC6 3
130C	12h VACV	wtVACV + MG132 1
131N	18h Mock	wtVACV + MG132 2
131C	18h VACV	wtVACV + MG132 3

2.9.3 Fractionation

Offline Tip-Based SCX Fractionation for Plasma Membrane Samples (IFN Screen)

SCX fractionation was based on a previous protocol (Dephoure & Gygi, 2011), in a modified form for small amounts of peptide. A fritted tip was made by taking a p10 filter pipette tip, removing the filter and cutting it in half, then inserting this into the end of a p200 pipette tip. 10 mg of resin (polysulfoethyl A bulk material, Nest Group Inc) in 150 µl 100 % AcN was loaded onto the tip under vacuum. SCX buffer A and B were made as follows:

Buffer A – 4 ml H_2O, 168 µl 250 mM KH_2PO_4 pH 2.65 (final concentration of ~7mM), 10 % H_3PO_4 (as required to reach a final pH of 2.6-3 as confirmed on a pH strip), and 1.8 ml AcN (final concentration of 30 %).

Buffer B – 1.9 ml H2O, 2.1 ml 1M KCl (final concentration of 350 mM), 168 µl 250 mM KH_2PO_4 pH 2.65 (final concentration of ~7mM), 10 % H_3PO_4 (as required to reach a final pH of 2.6-3 as confirmed on a pH strip), and 1.8 ml AcN (final concentration of 30 %).

The SCX material was conditioned with 2x 400 µl buffer A, then 400 µl buffer B and finally 4x 400 µl buffer A. The dried peptides were resuspended in 400 µl buffer A and applied to the tip under vacuum, with a flow rate of ~150 µl/min. It was then washed with 150 µl buffer A. A series of 6x 150 µl washes were then performed with increasing concentrations of K+, starting at 10 mM (10 µl buffer B into 340 µl buffer A), through 25, 40, 60, 90 to 150 mM KCl (150 µl buffer B into 200 µl buffer A). The eluate from each was collected in a separate eppendorf, and the 6 fractions dried completely by vacuum centrifugation. They were then resuspended in 200 µl 1 % FA, desalted using a StageTip (as above), dried again and resuspended in MS buffer for analysis.

Offline HpRp Fractionation for Whole Cell Lysate Samples

TMT-labelled tryptic peptides from WCL samples were subjected to HpRP fractionation; this was performed by Dr Colin Davies (Weekes Lab, Department of Medicine, University of Cambridge) using an Ultimate 3000 RSLC UHPLC system (Thermo Fisher Scientific) equipped with a 2.1 mm internal diameter (ID) x 25 cm long, 1.7 µm particle Kinetix Evo C18 column (Phenomenex). The mobile phase consisted of A: 3 % AcN, B: AcN and C: 200 mM ammonium formate pH 10. Isocratic conditions were 90 % A / 10 % C, and C was maintained at 10 % throughout the gradient elution. Separations were conducted at 45 ºC. Samples were loaded at 200 µl/min for 5 min. The flow rate was then increased to 400 µl/min over 5 min, after which the gradient elution proceed as follows: 0-19 % B over 10 min, 19-34 % B over 14.25 min, 34-50 % B over 8.75 min, followed by a 10 min wash at 90 % B. UV absorbance was monitored at 280 nm and 15 s fractions were collected into 96-well microplates using the integrated fraction collector. Fractions were recombined orthogonally in a checkerboard fashion, combining alternate wells from each column of the plate into a single fraction, and commencing combination of adjacent fractions in alternating rows (ie for fraction 1 combine column 1, row A, C, E etc, for fraction 2 combine column 2, row B, D, F etc). This yielded 12 fractions. The intermediate remaining wells were combined in the same manner to yield a second set of 12 fractions. Wells prior to the start or after the stop of the elution of peptide-rich fractions, as identified from the UV trace, were excluded. The first set of 12 fractions was dried by vacuum centrifugation and resuspended in 10 µl MS solvent prior to MS analysis.

2.9.4 LC-MS3

MS data was acquired using an Orbitrap Fusion for all PM enriched experiments apart from the two initial monocyte single runs and the 10-plex CD4 single run, where an Orbitrap Fusion Lumos mass spectrometer was used instead (Thermo Fisher Scientific, San Jose, CA). The Orbitrap Fusion Lumos was also used for all the WCL proteomic studies. Operation of the mass spectrometer was performed by Dr James Williamson (Department of Medicine, University of Cambridge) in the case of the Orbitrap Fusion, and Dr Robin Antrobus (Cambridge Institute for Medical Research, University of Cambridge) for the Orbitrap Fusion Lumos.

In both cases, an Ultimate 3000 RSLC nano UHPLC equipped with a 300 µm ID x 5 mm Acclaim PepMap µ-Precolumn (Thermo Fisher Scientific) and a 75 µm ID x 50 cm 2.1 µm particle Acclaim PepMap RSLC analytical column was used.

Orbitrap Fusion Experiments

Samples were loaded at 10 µl/min for 5 mins in loading solvent (0.1 % TFA) before beginning the analytical gradient which used solvents A (0.1 % FA) and B (80 % AcN + 0.1 % FA). All separations were carried out at 55 ºC. The following gradient was used: 3-5.6 % B over 4 minutes, 5.6-32 % B over 162 minutes, followed by a 5 minute wash at 80 % B and a 5 minute wash at 90 % B and equilibration at 3 % B for 5 minutes. For the analysis of the TMT samples, a MultiNotch MS3-based method was used (Ting *et al*, 2011; McAlister *et al*, 2014) with the following settings: MS1: 400-1400 thomson (Th), quadrupole isolation, 120,000 resolution, $2x10^5$ automatic gain control (AGC) target, 50 ms maximum injection time, ions injected for all parallisable time. MS2: Quadrupole isolation at an isolation width of m/z 0.7, CID fragmentation (normalised collision energy (NCE) 30) with ion trap scanning out in rapid mode from m/z 120, $1x10^4$ AGC target, 70 ms maximum injection time, ions accumulated for all parallisable time in centroid mode. MS3: in synchronous precursor selection mode the top 10 MS2 ions were selected for HCD fragmentation (NCE 65) and scanned in the Orbitrap at 50,000 resolution with an AGC target of $5x10^4$ and a maximum accumulation time of 150 ms, ions were not accumulated for all parallelisable time. The entire MS/MS/MS cycle had a target time of

3 s. Dynamic exclusion was set to +/- 10 ppm for 90 s. MS2 fragmentation was trigged on precursors $5x10^3$ counts and above.

Orbitrap Fusion Lumos Experiments

Samples were loaded at 5 ul/min for 5 mins in loading solvent (0.1 % FA) before beginning the analytical gradient which used solvents A (0.1 % FA) and B (80 % AcN + 0.1 % FA). All separations were carried out at 55 ºC. The following gradient was used: 3-7 % B over 3 minutes, 7-37 % B over 173 minutes, followed by a 4 minute wash at 95 % B and equilibration at 3 % B for 15 minutes. For the analysis of the TMT samples, a MultiNotch MS3-based method was used (Ting *et al*, 2011; McAlister *et al*, 2014) with the following settings: MS1: 380-1500 Th, 120,000 resolution, $2x10^5$ AGC target, 50 ms maximum injection time. MS2: Quadrupole isolation at an isolation width of m/z 0.7, CID fragmentation (NCE 35) with ion trap scanning in turbo mode from m/z 120, $1.5x10^4$ AGC target, 120 ms maximum injection time. MS3: In synchronous precursor selection mode the top 6 MS2 ions were selected for HCD fragmentation (NCE 65) and scanned in the Orbitrap at 60,000 resolution with an AGC target of $1x10^5$ and a maximum accumulation time of 150 ms. Ions were not accumulated for all parallelisable time. The entire MS/MS/MS cycle had a target time of 3 s. Dynamic exclusion was set to +/- 10 ppm for 70 s. MS2 fragmentation was trigged on precursors $5x10^3$ counts and above.

2.9.5 Data Processing

Processing of Raw Proteomics Data

Mass spectra were processed using 'MassPike' a Sequest-based software pipeline for quantitative proteomics, through a collaborative arrangement with Professor Steven Gygi's laboratory (Harvard Medical School). Within the MassPike platform is software for conversion of the spectra into mzXML format using an extractor modified from Thermo Fisher's RAW File Reader library (version 4.0.26).

The human UniProt database (26th January 2017) was combined with a database of common contaminants such as porcine trypsin and endoproteinase LysC. For analysis of

the VACV infection samples, this was additionally combined with the VACV strain WR UniProt database (23rd February 2017). The combined database was concatenated with a reverse database composed of all protein sequences in reversed order. Searches were performed using a 20 parts per million (ppm) precursor ion tolerance (Haas *et al*, 2006) and product ion tolerance of 0.03 Th. Static modifications were included for TMT on lysine residues and peptide N termini (229.162932 Da) and carbamidomethylation of cysteine residues (57.02146 Da), while oxidation of methionine residues (15.99492 Da) was set as a variable modification.

A target-decoy strategy was employed to control the fraction of erroneous protein identifications (Elias & Gygi, 2007, 2010); this involves using intentionally incorrect 'decoy' sequences (in this case the reversed protein sequences) in the database, in order to assess the number of matches to these and estimate a false discovery rate (FDR). Filtering of peptide spectral matches was to a FDR of 1 % at the peptide level, and subsequent filtering ensured an FDR of 1 % at the protein level (Kim *et al*, 2011a; Wu *et al*, 2011). Filtering was performed using a linear discriminator analysis (Huttlin *et al*, 2010), distinguishing correct from incorrect peptide identifications using an approach similar to the Percolator algorithm (Kall *et al*, 2007), though utilising a distinct machine learning algorithm. These approaches set filters taking into account multiple parameters simultaneously; these include XCorr, ΔCn, missed cleavages, peptide length, charge state, and precursor mass accuracy. Protein assembly was guided by principles of parsimony, generating the smallest set of proteins necessary to account for all observed peptides.

MassPike software quantified proteins by summing TMT reporter ion counts across all matching peptide spectral matches for the given protein (McAlister *et al*, 2012, 2014). A minimum of one unique or razor peptide per protein was used for quantitation. Briefly, a 0.003 Th window around the theoretical m/z of each reporter ion (126, 127n, 127c, 128n, 128c, 129n, 129c, 130n, 130c, 131n and 131c for 11-plex experiments) was scanned for reporter ions, and the maximum intensity nearest to the theoretical m/z was used. The number of TMT reporter ions in each MS3 spectrum is directly proportional to the signal-to-noise (S:N) ratio for each ion, and is the primary determinant of the quality of quantitation (Makarov & Denisov, 2009). Conservatively, each individual peptide used in quantitation was therefore required to contribute enough TMT reporter ions (at least ~1250 per spectrum) to ensure sufficient quality and a representative picture of relative

protein abundance (McAlister *et al*, 2012). In order to minimise peptide co-isolation, an isolation specificity of 0.5 was employed (Ting *et al*, 2011). Peptide-spectral matches with no MS3 spectra, or poor quality MS3 spectra (more than 9 TMT channels missing and/or a combined S:N ratio of less than 25 for each used channel up to a maximum of 250 ie 100 total for a 4 plex, 150 for a 6 plex, 250 for a 10 or 11 plex) were excluded from quantitation. All peptides that met the stated criteria for a given parent protein were then summed. In effect, this weights the contribution of each peptide to the total S:N of the protein, according to the quantitation of the TMT reporter ions for that peptide. Protein level quantitation values were exported to excel for further analysis.

At this stage, proteins identified by matches to the reverse hits and common contaminants databases were removed. The data was then normalised assuming an equal total amount of protein in every channel. For this, a ratio was calculated between the total S:N in a given channel, and the total S:N in the first channel used; this ratio was used to scale all S:N values in that channel, providing a normalised S:N. In all further investigations, S:N refers to this normalised S:N.

2.9.6 Data Availability

Unprocessed peptide data files for the VACV proteomic screen are available at doi:10.17632/wxk9gnw22r.1. These files include details of peptide sequence, redundancy, protein assignment, raw unprocessed TMT reporter intensities and isolation specificity. The MS proteomics data have been deposited to the ProteomeXchange Consortium (http://www.proteomexchange.org/) via the PRIDE partner repository with the dataset identifier PXD012785. The IFN dataset will be made similarly available following publication.

2.10 Data Analysis and Statistics

2.10.1 Chapter 3: IFN Screen

The normalised S:N values, generated as describe above, were used for all further analysis, presentation and generation of plots. Additionally, as it is challenging to confidently assign peptides to a specific human leukocyte antigen (HLA) alleles, and to account for different alleles being expressed in the different donors, the S:N values were summed to give a single value for HLA-A, HLA-B, HLA-C and HLA-DRB1. Additionally, all classical HLAs were excluded from the investigation of protein abundance.

In the dot plot of IFN induced changes at the surface of THP-1s, p-values were estimated based on significance B (Cox & Mann, 2008), calculated using Perseus (Tyanova *et al*, 2016). Significance B assumes a normal distribution of fold changes (FCs) observed, but allows a different standard deviation (SD) for up- and down- regulated proteins. Additionally, proteins are grouped into bins according to their S:N; proteins with a higher S:N are more accurately quantified and therefore a smaller FC may still be significant.

Abundance values for proteins were calculated based on the 'iBaq' methodology, adapted from the original description (Schwanhäusser *et al*, 2011). The maximum MS1 precursor intensity for each peptide was determined, and a summed MS1 precursor intensity for each protein across all matching peptides was calculated. This summed intensity was divided by the number of theoretical tryptic peptides for that protein between 7 and 30 amino acid residues in length to give an estimated iBAQ value. To determine the abundance of a protein at the surface of unstimulated cells, the summed intensity was adjusted in proportion to the S:N values: (Donor 1 + Donor 2 + Donor 3 + Donor 4 + Donor 5 unstimulated) / ∑(all donors with/without IFN). Colour coding for the comparison of the abundance of proteins at the surface of monocytes and T cells was based upon calculating the mean and SD of the log transformed ratios for all proteins quantified, and defining cut off ratios at one, two or three SDs away from the mean. For donor-to-donor comparison, the intensity was adjusted in proportion to the unstimulated sample for that donor (eg Donor 1 unstimulated / ∑(all donors + / - IFN).

When examining IFN induced FC, in order to avoid artificial inflation of changes due to poor quantitation of a protein in one channel for a given donor, the data for a donor where either the stimulated or unstimulated channel contributed less than 2 % of the total S:N across the 10-plex for that proteins was removed. This was only applied to the primary monocytes and T cells, where data for five donors was available. Only proteins with data for three or more donors were examined.

For identifying IFN regulated proteins, upregulated proteins met the 'sensitive' criteria by having an average FC > 1 SD away from the mean (calculated as described above for the abundance ratio), and a FC > 1 in all donors where quantified. The SD away from the mean FC determined the colour of the average FC bar in the graphs. The filter for being stringently upregulated employed an additional criteria of a Benjamini-Hochberg (BH)-corrected p-value < 0.05, calculated using a paired, two-tailed student's t-test on log transformed data. Downregulated proteins were assessed identically. This was all calculated using excel.

R^2 values for donor-to-donor comparisons were calculated using excel.

2.10.2 Chapter 4: Vaccinia Virus Screen

For analysis of the triplicate time course dataset, the S:N in each channel was analysed as a fraction of the total observed S:N for that protein across the 11 plex. Considering the fractional TMT signal rather than the absolute normalised intensity effectively corrected for difference between the number of peptides quantified for each protein and between replicates.

For analysis of proteomic changes in the host, proteins were defined as 'sensitively' upregulated if they had a FC > 2 at any infection time point compared to the 18 h mock, and to be stringently upregulated they were also required to have a BH-corrected p-value < 0.05 at that time point. The same criteria were applied to downregulation. A two-tailed Student's t-test was used to estimate p-values for proteins quantified in all three replicates, comparing each time point with the 18 h mock, and these were BH-corrected. These p-values are also displayed in the dotplot and on graphs of expression profiles. Expression

profiles display the mean abundance across all replicate time-courses where the protein was quantified, with error bars denoting the standard error of the mean (SEM). Hierarchical clustering was performed using Cluster 3.0 on the average of the three time-courses, with the FC at each time point compared to the average of the three mock samples.

For clustering of viral proteins, the AraC channels were removed, and the data normalised to the maximum signal for each protein. XLStat (Addinsoft) was used for k-means clustering, and each cluster was then subjected to hierarchical clustering using cluster 3.0.

The 11 plex incorporating the MG132 and C6 deletion mutant data encompassed triplicate samples for WT infection, WT + MG132, and infection with a C6 deletion mutant. P-values were estimated using a Students two-tailed t-tests on the normalised data comparing either WT + MG132 or the C6 deletion mutant to the triplicate WT infection, and these were BH-corrected.

Viral replication assays were performed and analysed by Dr Yongxu Lu (Smith lab). Experiments were performed in triplicate, and the data displayed is the mean ± SEM. P-values were estimated using a two-tailed t-test.

Pathway analysis for identification of enriched categories of proteins was performed using the Database for Annotation and Visualisation and Integrated Discovery (DAVID) version 6.8 (Huang *et al*, 2009a, 2009b). The set of proteins indicated in the text was searched against a background of all human proteins quantified in that experiment.

2.10.3 Chapter 5: Candidate Antiviral Restriction Factors

Overlap between IFN stimulated proteins and virally downregulated proteins was performed using the sensitive criteria for IFN stimulation described above, and viral downregulation from multiple datasets as described in the text.

For all ‘single colour’ HCMV restriction assays, data is presented as the mean ± SEM, if triplicate data is available. If the data is duplicate, error bars indicate the range and this

is specified in the figure legend. P-values were estimated using a two-tailed Student's t-test comparing the percent infection in the three test samples to the three control samples.

For two colour restriction assays, the restriction ratio is calculated as described in the text. Bar plots display the mean restriction ratio ± SEM where triplicate data is available. P-values were estimated using a paired two-tailed Students t-test, with paired samples representing the percent infection of the iRFP cells compared to the mCherry cells in each of the three replicate wells.

2.11 Software Used

Table 2.12 Table of software used

Software	Source	Use
MassPike	Prof Steven Gygi, Harvard	Processing of raw MS data
Excel	Microsoft	Data analysis, graph generation and t-tests
XLStat	Addinsoft	K-means clustering, Fischer's exact test
Cluster 3.0	Stanford University & University of Tokyo	Hierarchical clustering
Java Treeview	Stanford University	Visualising clustering results
Perseus	Max Planck Institute (Tyanova *et al*, 2016)	Significance B for dot plots
DAVID	(Huang *et al*, 2009a, 2009b)	Pathway enrichment analysis
R	(R Core Team, 2018)	Beeswarm plot of VACV proteins
Inkscape 0.92.4	https://www.inkscape.org	Drawing of some schematics
Powerpoint	Microsoft	Production of figures
Word	Microsoft	Writing of thesis
FlowJo vX	Becton Dickinson	Analysis of flow cytometry data
Image Studio Lite	LiCor	Analysis of immunoblots

CHAPTER 3: INTERFERON SCREEN

The Effect of Interferon at the Surface of Primary Leukocytes

3.1 Introduction and Aims

3.1.1 Introduction

IFN Stimulation of Primary Leukocytes

Several previous 'omics studies characterising the effects of IFNα on leukocytes have been conducted. However, many are at the transcriptomic level, or were constrained by the available proteomic technology at the time (Rusinova *et al*, 2013; Smiljanovic *et al*, 2012; Schlaak *et al*, 2002) (see Introduction 1.3.4). In this chapter, a proteomic screen investigating the effects of IFNα2a at the surface of primary leukocytes in multiple donors is described. This was intended both for the characterisation of the IFN response in these cell types, an important component of innate immunity, and for examination alongside data on viral infection to identify candidate ARFs (see Chapter 5).

The investigation of primary cells was particularly important for this research. The proteomes of primary cells differ from their cell culture equivalents, and the HIV accessory proteins (Vif, Vpr, Vpu and Nef) are necessary for maintaining a high viral load in vivo, whilst being dispensable for replication in cultured cell lines. This suggests a disparity in the cells responses to infection (Kestler 3rd *et al*, 1991; Malim & Emerman, 2008). Analysis of primary cells was previously challenging, as metabolic labelling requires the cells to be maintained in culture for some time prior to analysis. The use of labelling with chemical tags, such as TMTs, subsequent to cell lysis now makes quantitative comparative analysis in these cell types straightforward. Additionally, previous research observed vast donor-to-donor variation across IFN responses in primary T cells (Schlaak *et al*, 2002) (see Introduction 1.3.4). It was therefore important to perform these experiments in primary cells from multiple donors. Again, the use of TMTs for multiplexing samples was ideal for this purpose.

Being at the interface between cells and their environment, PM proteins are well situated to be components of signalling pathways, to interact with pathogens, or to act as ARFs. PM proteins are pharmacologically important, acting as the targets of many drugs, and are often under-represented in whole cell proteomic experiments (Savas *et al*, 2011; Kar *et al*, 2017). This investigation therefore uses PM profiling proteomics to focus on the effects of IFN at the cell surface.

The analysis presented here therefore substantially expands on previous investigations as it provides an unbiased exploration of the effects of IFN, without needing to pre-determine the genes of interest as in microarrays. Additionally, analysis was at the proteomic rather than transcriptomic level, focused specifically on cell surface effects, and considered primary cells across multiple donors.

Cell Surface Proteome of Primary Monocytes and T cells

As well as investigating the effects of IFN, this approach can also be used to investigate the cell surface proteome of the leukocytes. Previous investigation of the surface of primary T cells in the Cell Surface Protein Atlas identified 270 proteins in the CD4+ CD25- subset of cells (Bausch-Fluck *et al*, 2015). Another 'omic investigation of the naïve T cell surface identified 173 proteins using MS, and this was extended to 229 proteins using flow cytometry, though this was not quantitative (Graessel *et al*, 2015). This was further extended using bioinformatic analysis of transcriptomic data to identify proteins that may be localised at the cell surface. The investigation here quantified substantially more proteins than these previous studies, and can furthermore be used to compare the surface proteome of the two primary cell populations.

3.1.2 Summary

In this chapter, a proteomic screen exploring the effects of IFNα2a stimulation on the cell surface of primary monocytes and CD4+ T cells is presented, examining five donors for each cell type. The cells were stimulated overnight with IFNα2a, samples were enriched for PM proteins, and then subjected to TMT-based MS3 analysis. Ten-plex TMT allowed multiplexing of the unstimulated and IFN-stimulated samples from each of the five

donors, for analysis in a single experiment. This resulted in the quantification of 607 annotated PM proteins from monocytes and 485 PM proteins from T cells, as well as more limited data on the effects of IFN on THP-1s (a cultured monocyte cell line), and on pan-monocyte populations extracted from two additional donors. This provides an interesting dataset for the examination of the cell surface proteomes, as well as IFN-induced changes, and variation between donors. Transmembrane protein 123 (TMEM123) was identified as being IFN-stimulated in all cell types and donors examined, and some attempts were made to validate this observation.

3.1.3 Aims

1) Use a proteomic screen to investigate the effects of IFNα2a at the surface of primary leukocytes in:
    a. CD14+ Monocytes
    b. CD4+ T cells
2) Analyse the data to characterise the cell surface proteome of these leukocytes.
3) Analyse the data to identify IFN-induced changes at the cell surface of these primary leukocytes.
4) Consider donor-to-donor variation within the data.
5) Validate IFN stimulation of selected hits.
6) Overlap this data with corresponding data identifying virally-induced downregulation of cell surface proteins to identify candidate ARFs (chapter 5).

3.2 Validation of the Workflow for the IFN Proteomic Screen

3.2.1 Experimental Outline

To investigate the effects of IFNα at the surface of primary leukocytes, multiplexed quantitative proteomics using TMT-based mass spectrometry (MS) was employed (see

Introduction 1.5.2). At the time of performing this experiment, multiplexing up to 10 samples was possible (though this has now been extended to 16-plex), therefore primary leukocytes were isolated from five donors for each cell type, enabling both IFN-stimulated and unstimulated samples from each donor to be analysed together in a single experiment. Primary CD4+ T cells and primary monocytes were examined. To obtain these, peripheral blood mononuclear cells (PBMCs) were extracted from leukocyte cones from the NHS blood and transplant service (NHSBT) for monocytes, and from peripheral blood for T cells. Leukocyte cones provide a far greater number of PBMCs than a 50 ml peripheral blood donation, and were therefore used for extraction of monocytes, as these are usually present at relatively low abundance in the blood. The desired cells types were enriched from the PBMCs using negative isolation methods. They were then cultured overnight in the presence or absence of IFNα2a, before biotinylation of the cell surface glycoproteins and subsequent enrichment using streptavidin beads (see Introduction 1.5.3). These PM proteins were digested with trypsin, the 10 samples labelled with TMTs and subjected to MS3 analysis. In this way, relative quantitation of PM proteins in each sample was obtained (Figure 3.1).

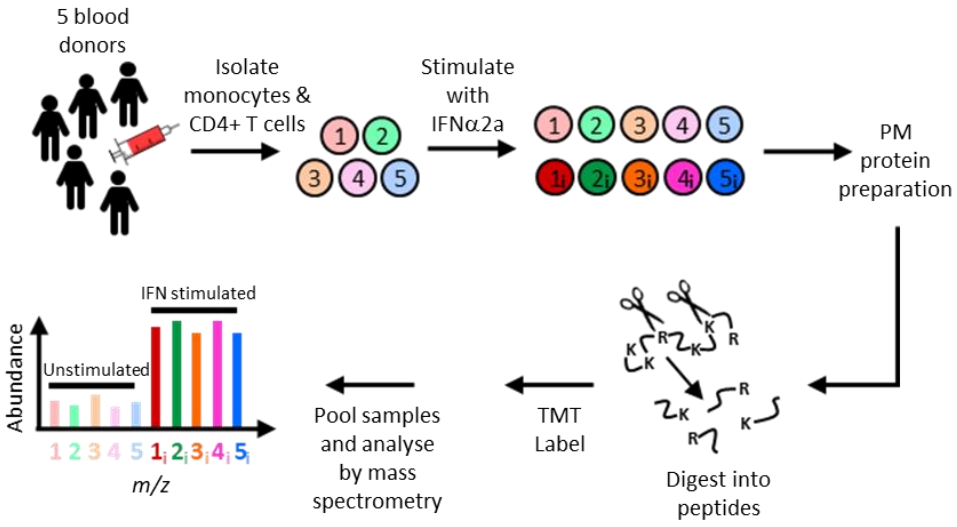


Figure 3.1 Experimental workflow for PM proteomic analysis of primary leukocytes

PBMCs were extracted from either leukocyte cones (monocytes) or peripheral blood (T cells) from five donors for each cell type. Primary monocytes or T cells were enriched from the PBMCs by negative isolation, and then cultured overnight in the presence of absence of IFN. PM proteins were enriched by biotinylation of cell surface glycoproteins and a pull-down with streptavidin beads. Proteins were digested into peptides using trypsin, labelled with TMT, and pooled for MS analysis.

3.2.2 Validation of the Workflow on THP-1s

The protocol was initially trialled on THP-1s, a cultured monocyte cell line. In most experiments presented, the MS sample was first analysed as a 'single shot' (one or three hour run on the mass spectrometer), which was examined to ensure adequate labelling with the TMTs, and roughly equal amounts of protein in each channel, to avoid excessive electronic normalisation. Following this, most samples were fractionated, separating the peptides into multiple samples, each of which were subjected to a separate MS run, enabling deeper coverage. In the case of the THP-1 sample, only the single shot was performed as this provided enough data to quantify positive controls and validate the approach. This quantified 720 proteins; 570 of these had a GO annotation of 'plasma membrane', 'cell surface', 'extracellular' or 'short GO', and were therefore considered as annotated PM proteins (Table 3.2). 31 PM proteins were upregulated more than 1.5 fold, including several positive controls, giving confidence in this approach (Figure 3.2).

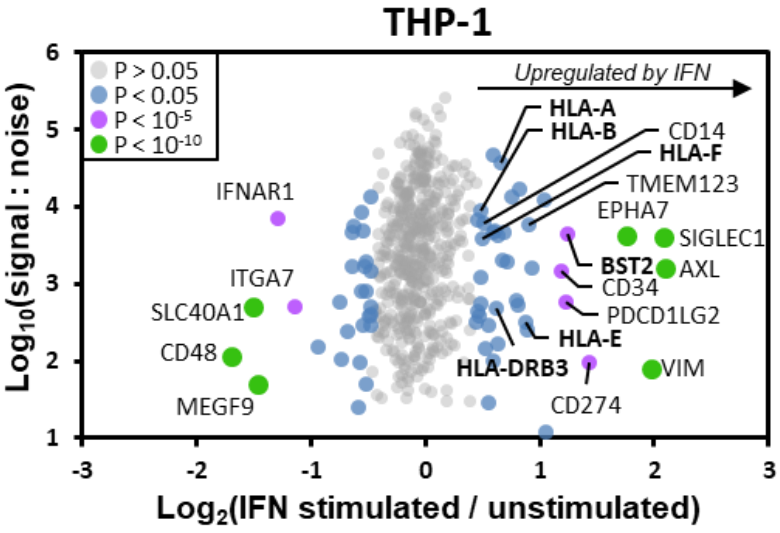


Figure 3.2 Cell surface changes in THP-1 cells upon IFN stimulation

Scatter plot of IFN induced changes in the 570 annotated PM proteins quantified in Thp-1 cells. Benjamini-Hochberg (BH)-corrected significance B was used to estimate p-values. The positive controls HLA and BST2 are highlighted in bold.

Multiple HLAs were upregulated, along with the IFN-stimulated HIV restricted factor, BST2 (Neil *et al*, 2008). AXL, a receptor tyrosine kinase, was also upregulated; this was previously reported to be stimulated by IFNα in macrophages, acting to suppress TNFα production (Sharif *et al*, 2006). A smaller subset of 13 proteins were downregulated 1.5 fold by IFNα, including the IFNAR1 component of the IFN receptor. Proteolytic downregulation of IFNAR1 after signalling provides negative regulation of the IFN response (Zheng *et al*, 2011).

3.3 Enrichment of Primary Leukocytes

3.3.1 Optimisation of Primary Monocyte Enrichment

Having validated some technical aspects of the screen on THP-1s, it was also important to optimise enrichment of the primary cells. Two kits for enrichment of primary monocytes were trialled, examining both the yield and purity. A sufficient number of cells were required in order to obtain enough material for MS analysis following the enrichment for PM proteins. PBMCs were enriched from leukocyte cones obtained from the NHSBT by density gradient centrifugation over Ficoll, and then subjected to negative selection for CD14+ monocytes using either the Dynabeads Untouched Human Monocytes kit or the MACS Monocyte Isolation Kit II. Both protocols involved incubation of the PBMCs with an antibody cocktail designed to bind cells other than CD14+ monocytes, which could then be removed using magnetic beads. The Dynabeads protocol employs a magnet to remove previously added magnetic beads, whilst MACS utilises a column of beads through which the cells are passed. This leaves just the desired population of cells. The MACS kit yielded improved viability (Figure 3.3A, B) and purity (Figure 3.3C, D) of cells compared to the Dynabeads kit, and was therefore used for all further experiments.

An additional level of optimisation was then performed on the MACS enrichment protocol. The MACS negative selection kit uses biotin conjugated antibodies to label cells, which are then removed by passing through a column of microbeads which have anti-biotin monoclonal antibodies conjugated to them. The initial protocol tested for MACS enrichment involved passing the cells through the MACS 'LS' column of

microbeads. An additional trial was performed with the inclusion of an 'LD' column. The LD column is intended to be more stringent, removing even weakly bound cells, due to having a slower flow rate. Only a small improvement in purity (~0.5 %) was observed following this additional step, however this was accompanied by a substantial reduction in yield of > 50 %, and it was therefore decided to omit this for further experiments.

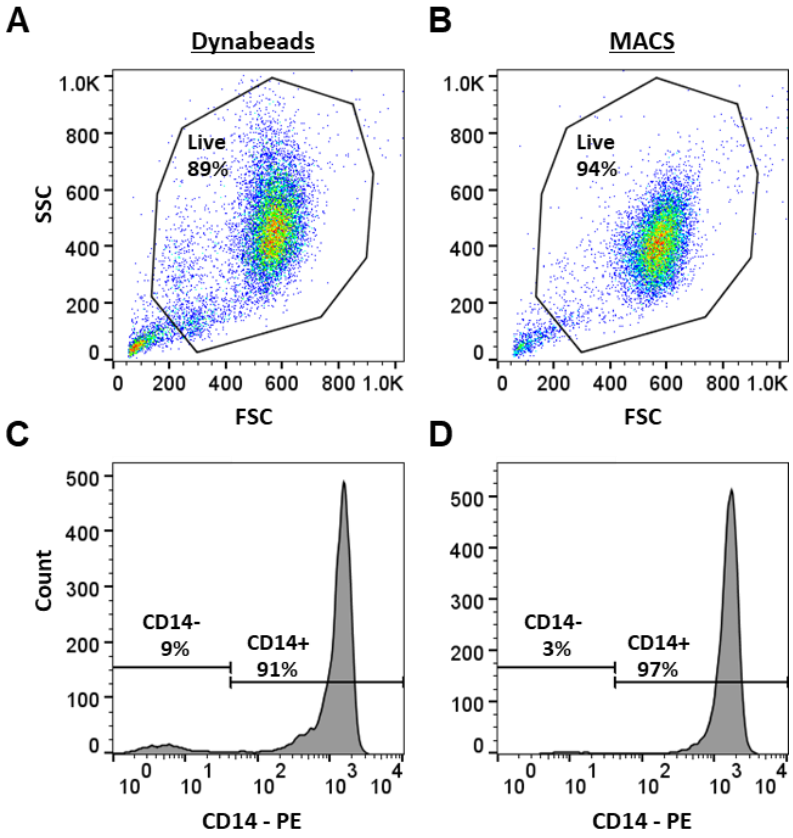


Figure 3.3 Comparison of methods for enrichment of primary CD14+ Monocytes by flow cytometry

(A) Viability of CD14+ monocytes following enrichment using the Dynabeads Untouched Human Monocytes kit was ~ 89 %.

(B) Viability of CD14+ monocytes following enrichment using the MACS Monocyte Isolation Kit II was ~ 94 %.

(C) Purity of CD14+ cells was ~91 % using the Dynabeads kit

(D) Purity of CD14+ cells was ~97 % using the MACS kit

3.3.2 Purity of Enrichment

PBMCs were isolated from NHSBT leukocyte cones, and enriched for primary monocytes using the MACS Monocyte Isolation Kit II, as described. For primary CD4+ T cells, PBMCs were isolated from 50 ml peripheral blood, and T cells enriched using the Dynabeads Untouched Human CD4 T Cells kit. Purity was confirmed by cell surface flow cytometry with antibodies detecting CD14 (Figure 3.4A, B) or CD4 (Figure 3.4C, D) for monocytes and T cells respectively, prior to overnight IFN stimulation. In all cases, a purity of ≥90 % was confirmed (Table 3.1).

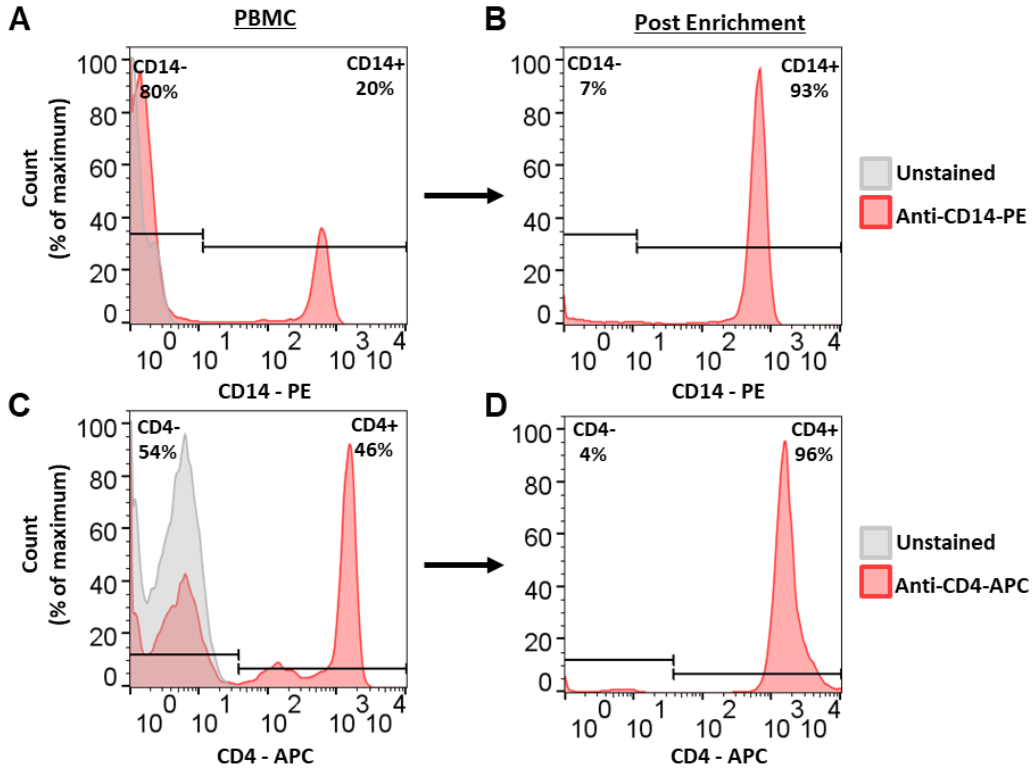


Figure 3.4 Analysis of purity of primary monocytes and CD4+ T cells

Purity of the primary cells was examined by cell surface flow cytometry:

(A) Analysis of CD14 expression in PBMCs from donor 3 prior to enrichment.

(B) Analysis of CD14 expression in monocyte enriched population from donor 3.

(C) Analysis of CD4 expression in PBMCs from donor 8 prior to enrichment.

(D) Analysis of CD4 expression in T cells enriched population from donor 8.

Table 3.1 Pre- and post- enrichment cell purity

Cell purity assessed by flow cytometry for CD14 (monocytes, left) or CD4 (T cells, right).

Monocyte donors	Pre-enrichment % CD14+	Post-enrichment % CD14+	CD4+ T cell donors	Pre-enrichment % CD4+	Post-enrichment % CD4+
Donor 1	18	91	Donor 6	30	96
Donor 2	26	93	Donor 7	45	91
Donor 3	20	93	Donor 8	46	96
Donor 4	22	92	Donor 9	20	96
Donor 5	28	90	Donor 10	25	97

3.4 Mass Spectrometry Analysis

Raw MS files were processed, and contaminating proteins and matches to the reverse database were removed. Prior to analysis, the data was digitally normalised to ensure equal loading in all channels. This was done by scaling the signal:noise (S:N) values so that the total intensity was the same in each channel. Additionally, as multiple donors had been analysed, a variety of HLA alleles were quantified; identification of the true MHC allele a peptide is derived from by MS is challenging, as many peptides are redundant to multiple alleles. The S:N values were therefore summed to give a single quantitation for HLA-A, HLA-B, HLA-C and HLA-DRB1. In the monocytes, 607 PM proteins were quantified, and 485 in the T cells (Table 3.2). In both cases PM annotated proteins comprised ~60 % of the proteins quantified.

Table 3.2 Number of proteins quantified in each cell type

The 'total number' column gives all proteins quantified in the given cell type (excluding reverse hits and contaminants), whilst 'annotated PM proteins' were filtered according to GO annotations as described in section 3.2.2.

Cell type	Total number of proteins quantified	Number of annotated PM proteins quantified
Thp-1	720	570
Pan monocyte	716	582
CD14+ monocyte	989	607
CD4+ T cells	798	485

3.5 Quantitative Analysis of the Cell Surface Proteome

In order to investigate the composition of the cell surface proteome, a method based on 'intensity based absolute quantification' (iBAQ, (Schwanhäusser *et al*, 2011)) was used. The maximum precursor intensity for all peptides contributing to a given protein were summed, before dividing by the theoretical number of tryptic peptides between 7-30 amino acids in length for that protein. This therefore generated a measure of abundance of a protein, by comparing its observed signal to what would be expected based on its size. As this abundance is given in terms of the whole 10-plex, the value was scaled according to the signal detected in the five unstimulated samples compared to the total signal for that protein, to give an estimate of the proteins expression at the surface of unstimulated cells. Classical class I and II MHCs were excluded from all abundance analysis to eliminate bias introduced by differentially expressed alleles between donors.

3.5.1 Monocyte and T cell Surface Proteomes

Just 17 proteins were found to contribute more than 1 % each to the cell surface proteome of monocytes, with the summed contribution from all 17 totalling ~75 % (Figure 3.5A). DAVID enrichment analysis compared these proteins to the background of all the PM proteins quantified in each cell type, and revealed enrichment of a single cluster, with terms including 'cell adhesion molecules' and 'leukocyte migration'. Proteins contributing to this cluster included integrin alpha-L (ITGAL) and integrin beta-2 (ITGB2), which form the leukocyte function-associated antigen-1 (LFA-1), and some of its ligands, intracellular adhesion molecule (ICAM)-1 and ICAM3. Furthermore, the four most abundant proteins composed half of the proteome: CD44, solute carrier family 2, facilitated glucose transporter member 3 (SLC2A3), leukosialin (SPN) and basigin (BSG). In CD4+ T cells, 22 proteins contributed more than 1 % of the cell surface proteome, with the summed contribution from all 22 totalling 67 % (Figure 3.5B). No enriched classes of proteins were identified here, but SPN and CD44 were again some of the most abundant proteins.

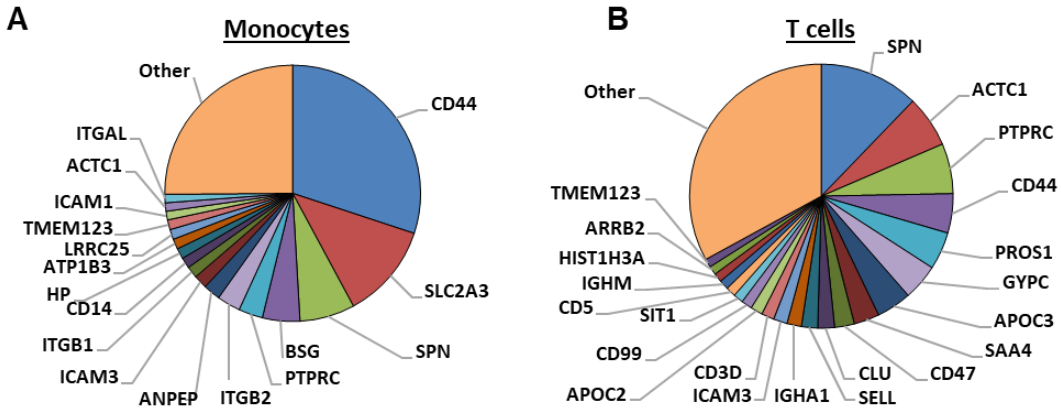


Figure 3.5 Cell surface proteome of primary monocytes and T cells

Abundance of proteins was calculated using an approach based on iBAQ methodology (Schwanhäusser *et al*, 2011), and pie charts display the relative contribution of individual proteins to the surface of the cells. Proteins contributing <1 % of the surface proteome are grouped into 'other'. Classical class I and II MHCs were excluded from all abundance analysis to eliminate bias introduced by differentially expressed alleles between donors.

(A) Primary monocyte surface proteome.

(B) Primary T cell surface proteome.

3.5.2 Comparison of Surface Proteomes

Comparison of the 281 PM proteins quantified in both monocytes and CD4+ T cells (excluding classical HLAs) showed a positive correlation in protein abundance, with CD44 and SPN being particularly abundant in both cell types (Figure 3.6). Other proteins were expressed to a greater extent on just one cell type, for example the monocyte marker CD14 was 36 fold more abundant on monocytes (ratio calculated based on percent contribution to surface proteome of each cell type), and signal regulatory protein alpha (SIRPA), a myeloid membrane receptor which interacts with CD47, was 32 fold more abundant on monocytes. SLC2A3 (also known as glucose transporter (GLUT)-3), was previously reported as being 8.4 times more predominant in monocytes compared to lymphocytes, whilst GLUT1 (SLC2A1) was more predominant in lymphocytes (Fu *et al*, 2004), in keeping with the data presented here. In T cells, CD5 was present at a greater abundance.

This comparison only includes proteins quantified in both the monocyte and T cell dataset. There were some proteins, such as leucine rich repeat-containing 25 (LRRC25)

and C-type lectin domain family 12 member A (CLEC12A), which are known to be predominantly expressed in monocytes, and were only detected in the monocyte samples (Liu *et al*, 2018; Marshall *et al*, 2004). However, absence of a protein in the dataset of a particular cell type does not necessarily indicate it was not present. It may simply not have been detected by chance, particularly if the protein was present at low abundance.

Another caveat of this data is that the populations of cells were not completely pure, so contaminants from cell types other than those studied might have affected the proteins quantified, though these were likely to appear at low abundance. Additionally, if the population was naturally heterogeneous, for example the CD4+ T cells encompass multiple subsets (Okada *et al*, 2008; Caza & Landas, 2015), profiles from these subsets will be averaged out. Some proteins present in just one subset may therefore appear at very low overall abundance, or a protein that was highly abundant on just a subset of cells (either T cells or contaminants), may appear at high abundance.

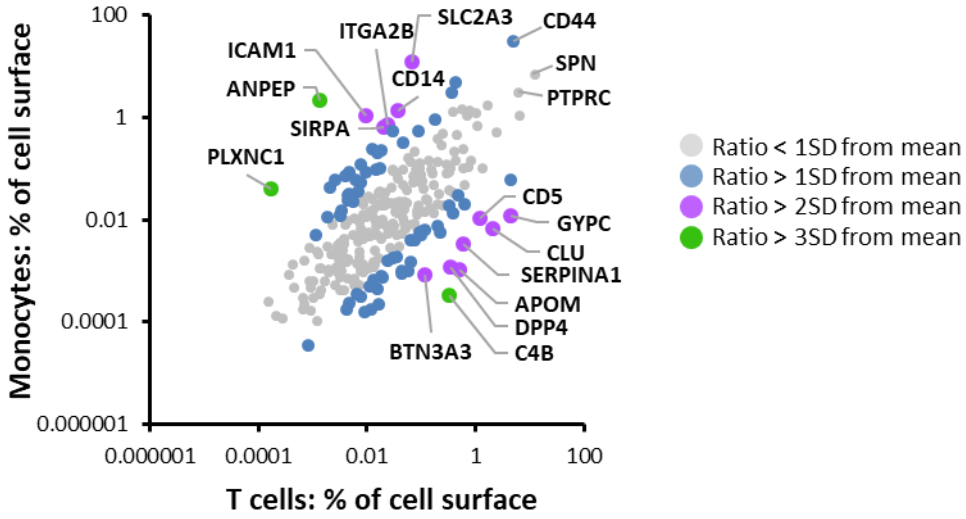


Figure 3.6 Comparison of monocyte and T cell surface proteomes

Comparison of the contribution of different proteins to the cell surface of primary monocytes and T cells. 281 proteins were quantified in both cell types. Classical class I and II MHCs were excluded from all abundance analysis to eliminate bias introduced by differentially expressed alleles between donors.

3.5.3 Variation between Donors

Using multiplexed proteomics also enabled the comparison of protein abundance across the five donors to assess variability between individuals. In this case, the abundance values were scaled according to the intensity in the unstimulated sample from each of the five donors separately, compared to the total signal. Both monocytes and T cells revealed remarkably invariant cell surface composition, and a strong positive correlation was observed for all pairwise comparisons (Figure 3.7).

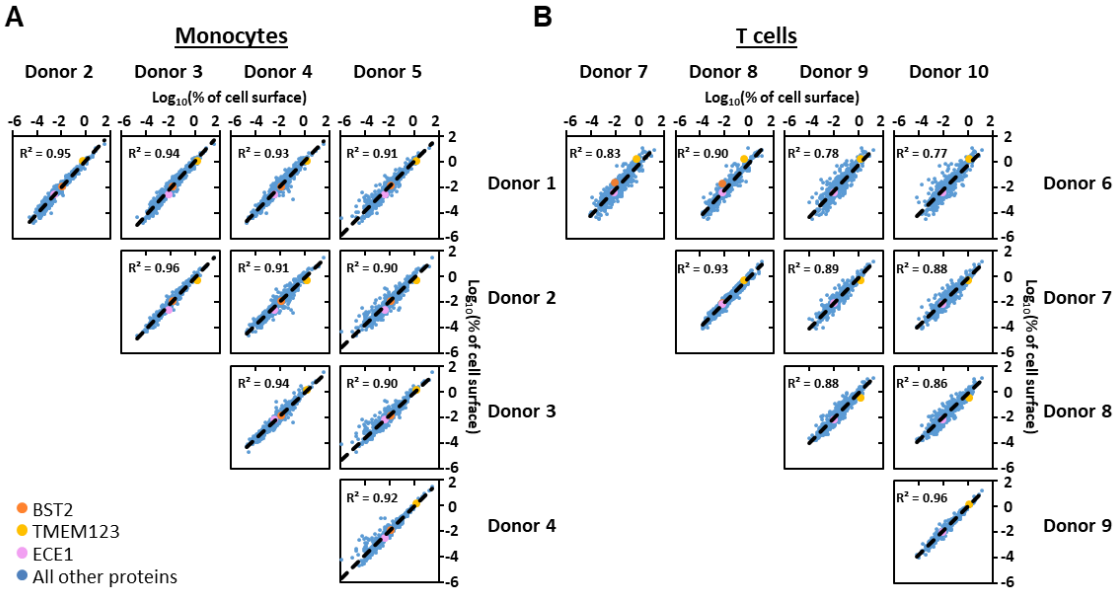


Figure 3.7 Variation of cell surface proteomes between donors

Pairwise comparison of the relative contribution of each protein to the cell surface in an individual donor. The iBAQ values for each protein was scaled according to the unstimulated S:N for each donor, and the % contribution to the cell surface proteome calculated. Classical class I and II MHCs were excluded from this analysis.

(A) Comparison of cell surface proteomes between monocyte donors.

(B) Comparison of cell surface proteomes between T cell donors.

3.6 The Effects of IFN at the Surface of Primary Leukocytes

Characterising the IFN response at the PM of these primary cells is important both in terms of developing our knowledge of IFN, crucial for its use in therapies, and to understand IFN related autoimmune diseases, as well as for identification of novel ARFs. To this end, the IFN induced changes in primary monocyte and T cells were examined.

The data was filtered to only include donors for a given protein if it was likely to be accurately quantified in both the stimulated and unstimulated channel. In some instances, the proteins were detected at the level of noise in one channel for a donor, giving a low intensity value. This could result in an artificially inflated FC. In order to avoid this, if a channel included less than 2 % of the total signal for that protein, that donor's data was removed for the given protein. As some proteins were now quantified in fewer donors, this dataset was then further filtered to remove any proteins that did not have data for at least three donors. This resulted in a final dataset of 606 proteins in monocytes and 482 in the T cells.

3.6.1 Monocytes

CD14+ Monocytes

Initial criteria considered the average IFN induced FC across all donors where a protein was quantified. The average FCs were taken for all proteins quantified in the monocytes, this data was log transformed and the mean calculated, giving the mean overall FC. Proteins with an average FC more than 1 SD away from this mean were considered upregulated. In the primary monocytes, 57 proteins fulfilled this criteria (Figure 3.8A, Table 3.3).

It was clear from visual inspection of this that in some instances, a particularly high FC in one or two donors meant a high enough average to pass the filtering in proteins which were not consistently IFN-stimulated. Although donor-to-donor variation was expected, and proteins which were not induced by IFN in all five donors might still have been of interest, the most robust hits should be identified by filtering for consistency. A second

criteria was therefore imposed, requiring the proteins to be IFN-stimulated to any extent (FC > 1) in all donors where quantified (Figure 3.8B). Of the proteins with an IFN induced FC > 1 SD away from the mean, 72 % were consistently upregulated in all donors. This is considered as the 'sensitive' criteria for IFN modulation, and used for all further analysis, however a more 'stringent' list could be obtained by considering only those proteins that also have a BH-corrected p-value < 0.05. Equivalent criteria were applied to identify downregulated proteins (Figure 3.8C), and identified 55 proteins that were consistently downregulated (Table 3.3).

Table 3.3 Proteins modulated by IFN in primary monocytes and T cells

The number of up and down regulated proteins are given for primary monocytes and CD4+ T cells, according to the stated criteria. Following application of the filtering described at the start of section 3.6, 606 proteins were examined for monocytes, and 482 for T cells.

	CD14+ Monocytes		**CD4+ T cells**	
	Upregulated	**Downregulated**	**Upregulated**	**Downregulated**
Average FC > 1 SD from mean	57	65	42	45
Sensitive criteria (average FC > 1 SD from mean, and FC > 1 in all donors)	41	55	13	14
Stringent Criteria (Sensitive criteria, with BH-corrected p-value < 0.05)	21	30	0	0

The sensitive criteria for upregulated proteins included several well characterised IFN-stimulated proteins, such as BST2, multiple HLAs and ISG15 (Figure 3.8). ISG15 is a known IFN-stimulated, ubiquitin-like protein involved in modulating the immune response, and inhibiting viral replication through a range of mechanisms (reviewed by Perng & Lenschow, 2018). It can be conjugated to target proteins by 'ISGylation', which may result in modulation of protein abundance, cellular localisation or formation of protein complexes. In the case of the influenza A virus non-structural protein 1 (NS1) protein, ISGylation disrupts its nuclear localisation and interaction with host proteins, limiting its ability to interfere with the antiviral response. ISGylation of human

papillomavirus (HPV) and HCMV proteins also restricts viral infection, whilst ISGylation of host proteins required for viral egress restricts HIV infection. ISG15 can also modulate the host immune response, for example unconjugated ISG15 stimulates IFNγ production, and the IFN response is further modulated by interaction with STAT1 and USP18 (Perng & Lenschow, 2018).

Sialic acid binding Ig-like lectin 1 (SIGLEC1) and CD69 were also identified by the sensitive criteria for IFN stimulation. SIGLEC1 provides negative feedback, controlling production of IFN. This is via recruitment of the E3 ubiquitin ligase TRIM27, resulting in degradation of TBK1 (Zheng *et al*, 2015). CD69 is known to act downstream of type I IFNs in order to inhibit egress of lymphocytes, so they are retained in the responding lymphoid organ during an immune response (Shiow *et al*, 2006).

Although monocytes are usually differentiated into DCs *in vitro* using granulocyte-macrophage colony-stimulating factor (GM-CSF) and interleukin (IL)-4/TNFα, in a previous study, monocytes stimulated with GM-CSF and IFNα also showed characteristics of DCs. This included increased expression of CD11c (integrin alpha-X, ITGAX), CD86, HLA-DR and C-C chemokine receptor type 7 (CCR7), all of which were observed in this data. The TNFα derived DCs expressed higher levels of genes involved in phagocytosis and adhesion, whilst the IFN derived DCs expressed greater levels of genes associated with migration (Santini *et al*, 2000; Korthals *et al*, 2007).

Upregulation of CD86 and CD83, DC maturation markers, was also observed in a previous study where monocytes were stimulated by IFNα2a only (Gerlini *et al*, 2008), and this was seen in the data presented here. In this study, they also observed an increase in adhesion molecules. The IFNα2a stimulated cells did not acquire a dendritic morphology or as potent an ability to induce naïve T cell proliferation as GM-CSF/IL-4 induced DCs. However, the cells were able to present antigen and activate memory T cells. Gerlini *et al.* suggested that the CD83+ CD14+ non-dendritic APCs may have a role in stimulating a memory immune response. Interestingly, they could not identify any CD83+ CD14+ monocytes in PBMCs from patients undergoing IFNα therapy, but they could find them in varicella skin lesion, where recruitment of plasmacytoid DCs results in a high level of type I IFN. A more comprehensive analysis of these markers and

examination of the monocyte morphology following IFN stimulation may therefore be interesting to investigate whether there is any differentiation of the cells.

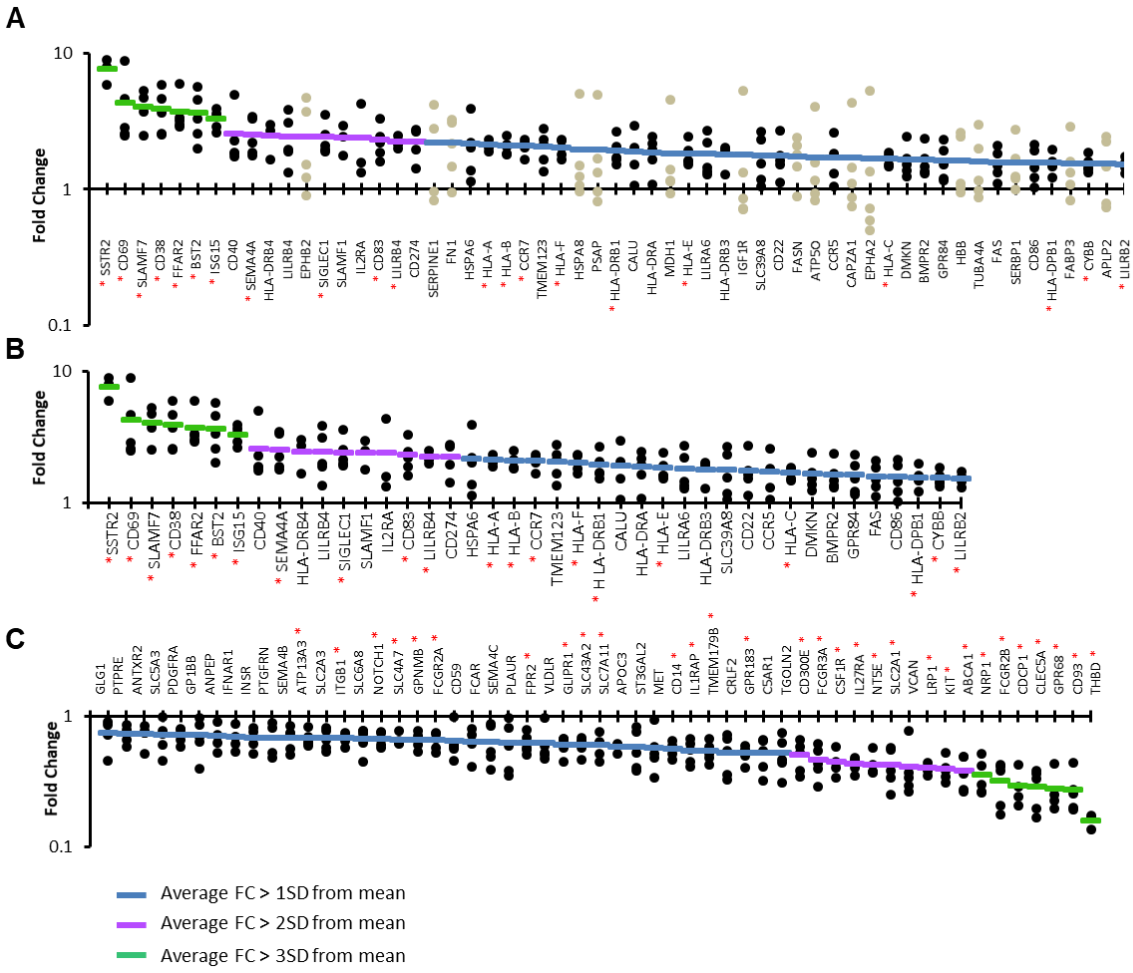


Figure 3.8 IFN-induced changes at the surface of primary monocytes

(A) FC is shown for all proteins with an average IFN induced FC > 1SD away from the mean in primary monocytes. Dots display the FC in each donor for the given protein, and the line is average FC. Proteins which contributed less than 2 % of the overall S:N in either the IFN-stimulated or unstimulated channel have been removed for the given donor. Those that were not upregulated in all donors are shown in grey. P-values were estimated using a BH-corrected paired two-tailed t-test on log transformed data; *p<0.05.

(B) FC of proteins upregulated by IFN stimulation in primary monocytes. As in (A) but with additional filtering to only show proteins consistently upregulated in all donors. P-values were estimated using a BH-corrected paired two-tailed t-test on log transformed data; *p<0.05.

(C) FC of proteins consistently downregulated by IFN in primary monocytes, filtered as in (A) and (B). P-values were estimated using a BH-corrected paired two-tailed t-test on log transformed data; *p<0.05.

THP-1s

Although the THP-1 sample (section 3.2.2) was not fractionated, so fewer proteins were quantified, an initial examination of this with the primary monocyte data could provide an interesting comparison between a primary cell type and its cultured equivalent. Previous research has shown that THP-1 cells express a lower level of CD14 than primary monocytes, and therefore show a reduced response to LPS, as measured by the production of IL-8. Overexpression of CD14 in the THP-1s increased the level of IL-8 production in response to LPS, though did not entirely rescue it. Additionally, whereas heparin enhances the LPS induced IL8 release in primary monocytes, it did not have the same effect in THP-1s (Bosshart & Heinzelmann, 2016). This suggests that there are further differences in this pathway between the primary and cultured cells, and this may extend to other cellular responses.

There were 357 proteins quantified in both the THP-1 and primary monocyte datasets (Figure 3.9). Of these, 11 proteins were commonly upregulated (SIGLEC1, interleukin 2 receptor subunit alpha (IL2RA), CD274, BST2, TMEM123, FAS and multiple HLAs), and 2 commonly downregulated (IFNAR1 and proto-oncogene KIT). Several proteins showed changes in either THP-1s or primary monocytes, but not the other, though in some cases this may be due to the threshold applied (sensitive criteria for primary monocytes, 1.5 FC for THP-1s), or the requirement for consistency between the donors. There are however, many more proteins modulated in the primary cells. CD59, a complement regulatory protein involved in inhibiting the membrane attack complex (Kim & Song, 2006), was the only protein to show conflicting regulation, being downregulated in the primary monocytes and upregulated in the THP-1s.

Repetition of the THP-1 proteomic experiment presented here and perhaps more primary cell donors would be required, along with fractionation of the THP-1 sample, for a more thorough and meaningful comparison. Additionally, it would be interesting to examine the primary cells alongside the cultured cells in the same multiplex experiment, to allow a more direct comparison of the relative abundance of different proteins.

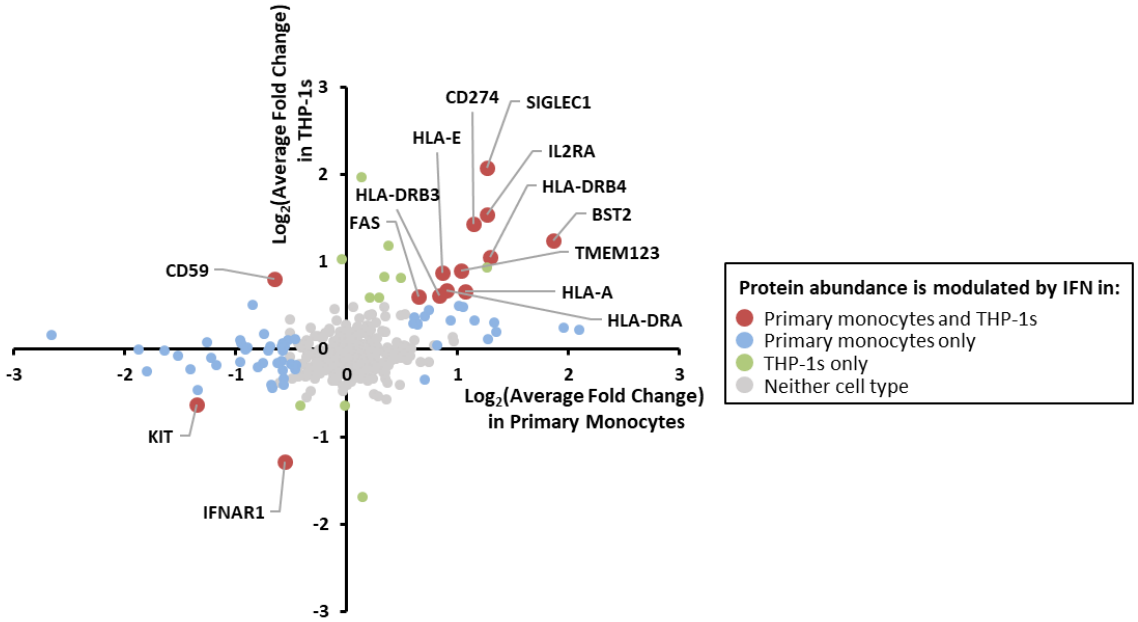


Figure 3.9 IFN-induced changes in primary monocytes and cultured THP-1s

Proteins were defined as being modulated by IFN in the primary monocytes if they met the 'sensitive criteria' (FC > 1 SD from the mean, and FC > 1 in all donors for upregulation), and were regulated in THP-1s if they had a FC > 1.5.

Pan Monocytes

In addition to the CD14+ monocytes from five donors, pan-monocyte populations were also examined from two more donors. As with the THP-1s, this sample was not fractionated so does not provide a complete dataset, but is useful for comparison and validation of the larger dataset. Whilst the MACS Monocytes Isolation Kit II extracted predominantly classical monocytes, pan-monocyte populations were enriched using the MACS Pan Monocyte Isolation Kit, which isolated classical (CD14++, CD16-), intermediate (CD14++, CD16+) and non-classical (CD14+, CD16++) monocyte populations. The same 'sensitive' filtering used previously was applied to the pan-monocytes to identify IFN modulated proteins (>1SD away from the mean and upregulated in both donors). As there were only two donors, no p-values were calculated and therefore 'stringent' filtering was not applicable. Of the proteins quantified in both CD14+ and pan- monocyte populations, 80 % of the proteins consistently upregulated in the CD14+ monocyte samples were also upregulated in pan monocytes (Figure 3.10A),

providing confidence in the data. Many of the proteins most dramatically downregulated by IFN in the CD14+ monocytes were also downregulated in the pan monocyte population, though often to a lesser extent (Figure 3.10B).

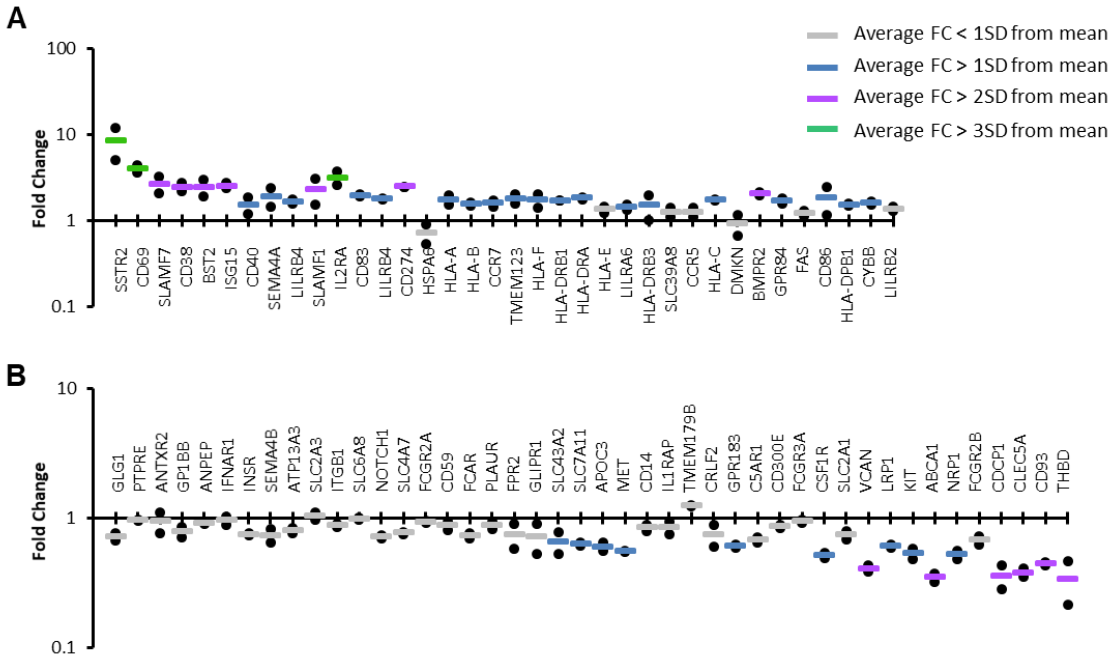


Figure 3.10 IFN induced changes at the surface of primary pan-monocytes

(A) IFN induced changes in pan-monocyte cells of proteins which were determined to be upregulated in the CD14+ monocytes (as in Figure 3.8B).

(B) IFN induced changes in pan-monocyte cells of proteins which were determined to be downregulated in CD14+ monocytes (as in Figure 3.8C).

Comparison to Transcriptomics

The Interferome provides a database of the effects of IFN on various cell types, at the transcriptomic level (Rusinova *et al*, 2013). Searching for previous datasets in the Interferome which investigated the effects of IFNα on monocytes identified one transcriptomic study (ID 306) (Smiljanovic *et al*, 2012). This study used microarrays to examine the effects of IFNα2a, IFNγ and TNFα treatment for 1.5 h on primary monocytes from healthy donors, or patients with RA or SLE. Smiljanovic *et al.* generated profiles of cytokine signatures so these could be compared to disease signatures. As this is a

transcriptomic study at a much earlier time point, and does not have the same enrichment for proteins expressed at the cell surface, some differences would be expected. However, this data was exported from the Interferome for a limited comparison.

In the Interferome data, multiple FCs were given for a single gene (from different probes on the microarray). An average was taken in these cases, to generate a single FC for each gene, and these were compared to the proteomic data generated here according to the gene name. Proteomic data was determined as being up or down regulated based on the 'sensitive' criteria described previously. A gene was considered to be modulated by IFN in the transcriptomic data if it was > 2 fold up or down regulated; this is the default cut off for IFN modulation by the Interferome. Note also, when searching the Interferome for a given gene, it only returns data where the difference in expression is significant at the FC searched for ($p < 0.05$). Therefore, 'no change' here means a FC < 2, but the p-value would still be significant; proteins, or some probes for given proteins, which truly showed 'no change' may therefore not show up in this comparison. For a complete comparison the data would need to be reanalysed more carefully, but here an initial overview is presented.

There were 250 proteins quantified in both datasets, the majority of which were not modulated by IFN in either study. Of these, 12 were commonly upregulated, including ISG15, SIGLEC1 and CD274. Two proteins were commonly downregulated, C5a anaphylatoxin chemotactic receptor 1 (C5AR1) and sodium/myo-inositol cotransporter (SLC5A3). Additionally, there were four proteins which were downregulated in this study and upregulated in the transcriptomic data. There were also a subset of proteins which were modulated by IFN in one study, but not the other. Several HLA-DR proteins were upregulated in the proteomic data but not the transcriptomic data. IFNα-induced upregulation of MHC class II has been previously observed in monocytes, though to a lesser extent than with IFNγ (Keskinen *et al*, 1997). In some cases, the discrepancy may be due to the thresholds applied. For example, BST2 is a known IFN-stimulated protein, and in this comparison does not appear as IFN-stimulated in the transcriptomic data, as it falls slightly below the two-fold cut off employed. Additionally, some differences may be expected due to the comparison of transcript and protein levels, as well as different donors, and differences in the IFN concentration and time point.

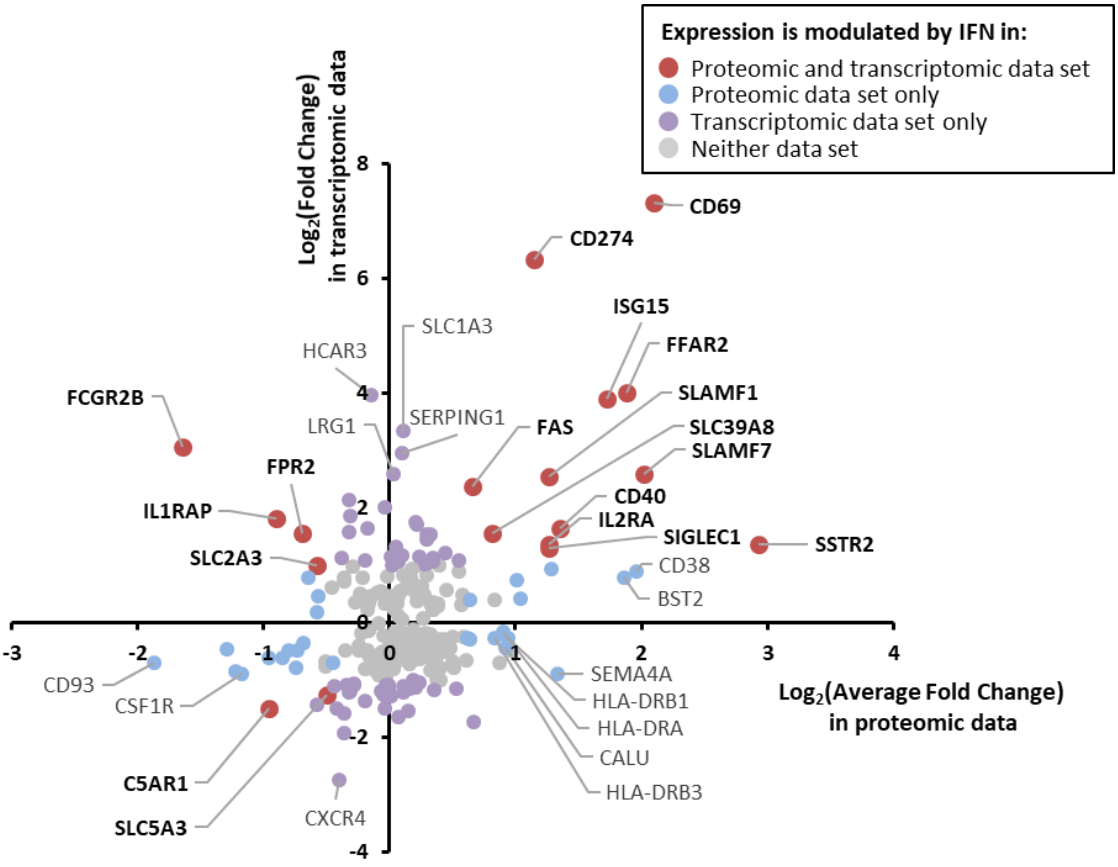


Figure 3.11 Comparison of proteins modulated by IFN in transcriptomic and proteomic data

Proteins from the proteomic data presented here were defined as up or downregulated according the sensitive criteria previously defined. Transcriptomic data was from (Smiljanovic *et al*, 2012) and genes were considered to be up or downregulated if they had an average FC>2 (though genes with FC < 2 would still have p<0.05 to be included in dataset).

3.6.2 CD4+ T Cells

Applying the same 'sensitive' filtering used for the monocytes, a much smaller subset of proteins were consistently up or down regulated in CD4+ T cells (Figure 3.12, Table 3.3). Again, this included positive controls such as BST2 and HLAs. Applying a BH-corrected paired two-tailed t-test did not identify any of the changes in T cells as being statistically significant, despite being consistent amongst the five donors. This may be due to over stringency of the multiple testing correction in the context of this proteomic data (Pascovici *et al*, 2016). It may also be partially due to heterogeneity in the population of

T cells. Additionally, although consistent, the FCs were on a small scale of magnitude, and the majority of proteins do not show consistent changes in abundance upon IFN stimulation. Despite this, identification of positive controls provides some confidence in the data.

A small subset of proteins were downregulated by IFN, including the IL-6 receptor (IL6R). This has previously been reported to be downregulated by IFNα in myeloma cells, where it prevents IL6 induced growth (Schwabe *et al*, 1994). Of the proteins identified as being upregulated by IFN, the IL-1 receptor antagonist (IL1RN) is also known to be regulated by IFN and is involved in modulating the inflammatory response (Arend *et al*, 1998; Liu *et al*, 1998), and C-type natriuretic peptide (CNP) is a known ISG involved in inhibiting HIV particle assembly (Wilson *et al*, 2012).

Some of the upregulated proteins have less well characterised relationships with IFN. Potassium voltage-gated channel subfamily A member 3 (KCNA3) is a heavily glycosylated 64 kDa protein with 6 transmembrane domains, which forms a homotetrameric complex. It is the major potassium channel responsible for regulation of the resting membrane potential and calcium signalling in T cells, with roles in T cell activation and proliferation, as well as IL-2 production (Spencer *et al*, 1997; Zhu *et al*, 2012; Shah *et al*, 2003). KCNA3 has previously been shown to be upregulated by JAK2, providing further evidence for its regulation by IFN (Hosseinzadeh *et al*, 2015). Probable G-protein coupled receptor 171 (GPR171) was also IFN-stimulated in this data. GPR171 is part of a family of proteins activated by extracellular nucleotides; It is suggested to be a negative regulator of myeloid differentiation (Rossi *et al*, 2013), and was identified as the receptor for BigLEN, a neuropeptide involved in food intake and metabolism (Gomes *et al*, 2013). GPR171 is also overexpressed in lung carcinoma tissues and thought to be involved in metastasis and migration (Dho *et al*, 2016).

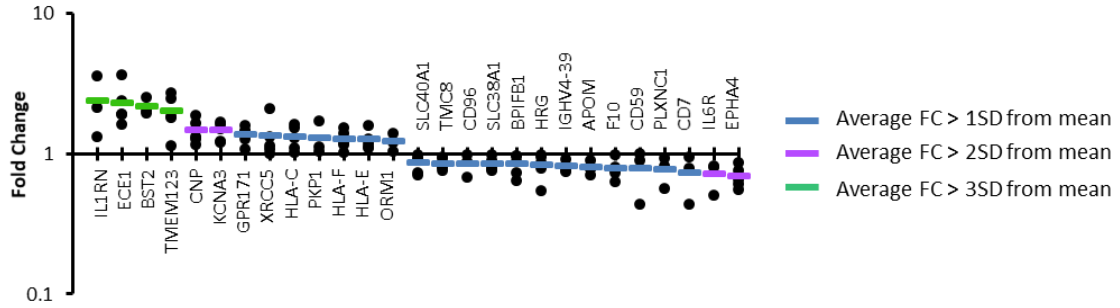


Figure 3.12 IFN induced changes at the surface of primary CD4+ T cells

FC of proteins up or downregulated in primary T cells. All proteins displayed had a FC > 1 SD from the mean, and were up or downregulated in all donors where quantified. A BH-corrected paired, two-tailed t-test did not identify any significantly modulated proteins.

Comparison to Transcriptomics

A previous investigation by Schlaak *et al.* (2002) used an array of 150 genes to investigate transcript level changes in expression of ISGs in various cells lines (fibrosarcoma HT1080, hepatoma HepG2 and HuH7, and T cell lymphoma Kit255), as well as primary T cells and DCs from several donors, upon stimulation with IFNα2a. There were 45 genes stimulated in at least one of the T cell donors in the Schlaak *et al* data. Of these, 16 were annotated PM proteins, and 4 of those were quantified in the CD4+ T cell data presented here (BST2, ISG15, collagen alpha-2 (COL1A2), and complement C1r subcomponent (C1R)). BST2 was upregulated in all donors in both datasets, and ISG15 was upregulated in all five donors in the Schlaak data, and 3/5 in this dataset. COL1A2 was upregulated in 1/5 donors in the Schlaak data and 2/5 here, whilst C1R was upregulated in 2/5 donors in the Schlaak data and none here. There are therefore several proteins showing concordant data between the proteomic and transcriptomic studies, with donor-dependent IFN stimulation. However, there is not enough overlapping data for a large comparison.

3.6.3 Comparison of Monocytes and T cells

The profile of IFN responses between monocytes and T cells appears quite different. There were 284 proteins quantified in both experiments, the majority of which were not IFN-stimulated in either cell type. Applying the sensitive filtering criteria (FC > 1SD from the mean, and FC > 1 in all donors), 11 proteins were IFN-stimulated exclusively in monocytes, 4 only in T cells, and 5 were commonly IFN-stimulated. These included the known HIV restriction factor BST2, HLA-C, E and F, and TMEM123 (Figure 3.13, Table 3.4). TMEM123 was also IFN-stimulated in samples from both pan-monocyte donors, and in the THP-1s. This is discussed further in section 3.8.

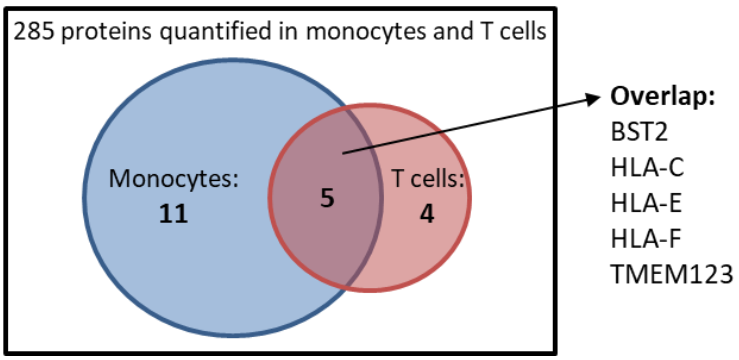


Figure 3.13 Overlap between IFN-stimulated proteins in monocytes and T cells

The proteins upregulated by IFN in primary monocytes and T cell was compared amongst the 285 proteins which were quantified in both datasets. Proteins were determined to be upregulated by IFN based on the sensitive criteria described previously and in Figure 3.8B.

Table 3.4 Proteins commonly modulated by IFN in monocytes and T cells

The FC in abundance of commonly modulated proteins upon IFN stimulation (IFN-stimulated / unstimulated) is given for each donor. Note, BST2 was only quantified in three donors in the T cell sample. D1 is donor 1, '-' indicates the protein was not quantified.

	CD14+ Monocytes FC upon IFN stimulation						CD4+ T cells FC upon IFN stimulation					
	D1	D2	D3	D4	D5	Average	D6	D7	D8	D9	D10	Average
BST2	4.6	5.7	2.0	3.3	2.6	3.6	1.9	2.0	2.5	-	-	2.2
HLA-C	1.5	1.9	1.7	1.8	1.5	1.7	1.0	1.5	1.1	1.3	1.6	1.3
HLA-E	1.6	2.4	1.6	1.9	1.5	1.8	1.1	1.3	1.2	1.1	1.6	1.3
HLA-F	1.8	2.3	2.0	2.3	1.7	2.0	1.0	1.4	1.2	1.2	1.5	1.3
TMEM123	2.3	2.8	1.7	2.1	1.4	2.1	1.1	1.8	2.7	1.9	2.5	2.0

Many more proteins changed in abundance in the monocytes compared to the T cells. This may be in part due to having quantified a greater number of proteins in the monocytes than in the T cells, however, 16 % of the proteins quantified in the monocytes were modulated by IFN, compared to 6 % of those in the T cells. Therefore, it may also be a biologically relevant phenomenon. A previous publication considered TNFα pre-treatment, followed by IFNβ stimulation of monocytes and T cells at the transcriptomic levels using microarrays. Henig *et al.* also observed more proteins to be involved in the IFN response in monocytes than T cells (Henig *et al*, 2013). Interestingly, they observed similar numbers of up and downregulated proteins in monocytes, whilst more were upregulated than down in T cells. Here, similar numbers of proteins change in both directions in both cell types. As this was a transcriptomic study, with a different subtype of IFN (and pre-treatment with TNFα), not looking specifically at the cell surface, completely concordant results would not necessarily be expected. Encouragingly however, they also identified CD38 as a protein which was IFN-stimulated in monocytes but not T cells, and validated this by flow cytometry.

Another study by Aso *et al*. combined transcriptomic data from multiple sources, examining stimulation with type I IFNs (α and β) on various HIV target cell types (monocytes, monocyte-derived macrophages and T cells) (Aso *et al*, 2019). This included several RNA sequencing and microarray studies, including the previously discussed research by Smiljanovic *et al*. They observed more proteins changing in abundance upon IFN stimulation of monocytes compared to T cells. Examining the ISGs that were common to both cell types, they also saw a greater magnitude of FCs in the monocytes. Aso *et al*. identified 1495 proteins which they defined as ISGs in at least one of the cell types, including 124 ISGs in T cells, 567 in monocytes and 1336 in monocyte-derived macrophages. Overlap between the proteomics data here and the meta-analysis by Aso *et al*. showed some commonly identified ISGs, and many commonly identified 'non-ISGs' (Figure 3.14). Encouragingly, Aso *et al*. identified TMEM123 as an IFN-stimulated protein in monocytes and monocyte-derived macrophages, however they did not find it to be IFN-stimulated in T cells.

One reason for the limited overlap may be the way the data was combined from multiple studies for each cell type by Aso *et al*., incorporating primary and cultured cell lines, and various type I IFNs. Somatostatin receptor type 2 (SSTR2) for example is designated as a non-ISG for monocytes, despite being 3.4 fold IFN-stimulated in primary monocytes by IFNα (data in this meta-analysis from the Smiljanovic study), as it was not IFN-stimulated in other studies incorporated in this analysis; however, these other studies used IFNβ or cultured monocyte cell lines. IL1RN was designated a non-ISG in T cells for similar reasons. Likewise, ISG15 was identified as an ISG in T cells by Aso *et al.*, but has not be counted as one in T cells in this study, despite having an average FC of 3.5, due to not being IFN-stimulated in two of the donors. A more thorough future comparison could investigate the individual contributing datasets from the Aso *et al.* research, and less stringent criteria for IFN stimulation in this proteomics data. Nonetheless, the global conclusions regarding more IFN responsive genes in monocytes compared to T cells, and changes of greater magnitude, are interesting, and in keeping with the observations here.

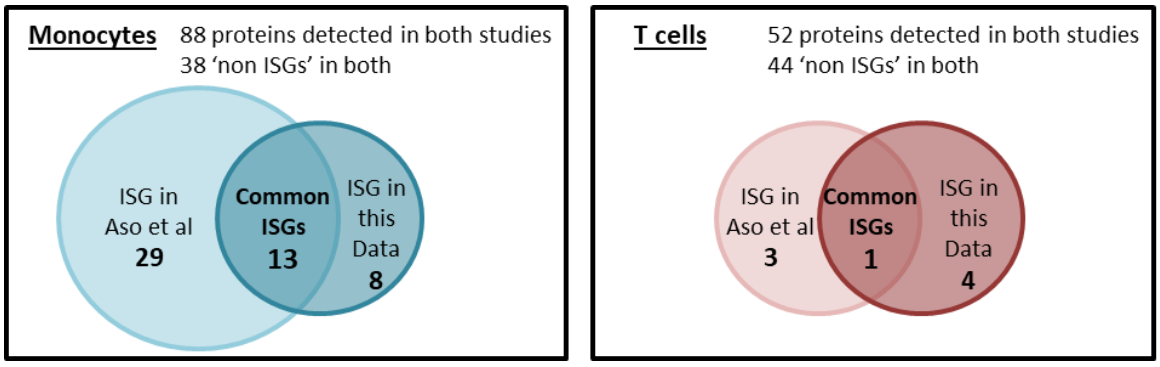


Figure 3.14 Comparison of proteomics data to transcriptomic meta-analysis

The proteomic data presented here for monocytes and T cells was compared to data by Aso *et al*. (2019), which combined multiple RNA sequencing and microarray based studies to look at IFN stimulation of monocytes and T cells (both primary and cultured cells) with type I IFNs. Proteins were considered as ISGs in this study if they met the sensitive criteria.

3.6.4 Variation between Donors

Importantly, there did not appear to be a systematically greater effect of IFN in some donors than other, the pattern of IFN-induced FCs appeared random (Figure 3.15). However, unlike the relatively invariant cell surface proteome, the IFN induced changes appeared to be more variable between donors, with a particularly low correlation coefficient in T cells (Figure 3.16). The particularly low correlation in the T cells may also be due, in part, to the majority of proteins not showing substantial changes in abundance upon IFN stimulation. Therefore, many of the proteins have small fluctuations around zero. Additionally, it could be influenced by different IFN responses of various T cell subsets.

However, a high level of variability in the IFN response between donors has been reported previously in primary leukocytes (Schlaak *et al*, 2002). In this instance, the cDNA array examined included 150 genes, composed of known ISGs and genes of interest. Of the 150 genes, 45 were induced >2 fold in at least one T cell donor, but only 10 were induced in all 5 donors. DCs were examined from three donors, and were also variable, though slightly less so than the T cells. Replicates of the cell line data however, were very reproducible, suggesting this is an effect of the donor variability, rather than the technique. It is also well established that the side effects and efficacy of IFN-therapy for viral hepatitis and cancer treatment are highly variable, suggesting differences between individuals (Cornberg *et al*, 2010; Ji *et al*, 2003; Lens & Dawes, 2002). Schlaak *et al*. suggest that variation between donors may be due to differential activation of STATs and IFN signalling pathways, or single nucleotide polymorphisms in the promotors or coding regions of IRFs, JAK or STAT proteins. This suggests further examination of proteins upregulated in just a subset of donors may still be valuable. Whilst just 13 proteins meet the sensitive criteria for consistent upregulation in T cells in this proteomic dataset, 112 proteins were upregulated more than 1.5 fold in at least one donor.

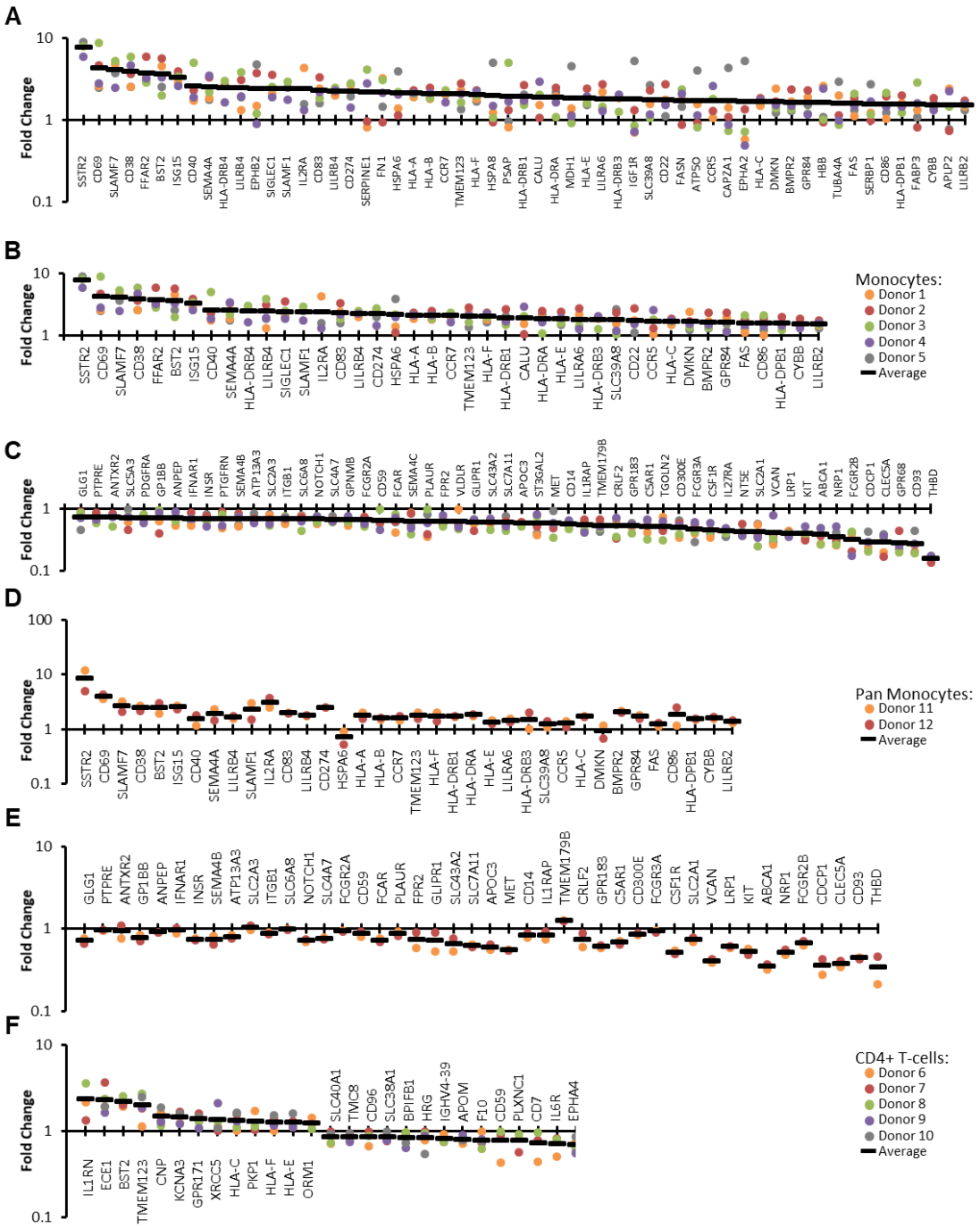


Figure 3.15 IFN-induced changes in individual donors of modulated proteins

(A) Proteins upregulated by IFN in monocytes, as in Figure 3.8A.

(B) Proteins consistently upregulated by IFN in monocytes, as in Figure 3.8B.

(C) Proteins consistently downregulated by IFN in monocytes, as in Figure 3.8C.

(D) Pan-monocyte data on proteins upregulated in monocytes, as in Figure 3.10A.

(E) Pan-monocyte data on proteins downregulated in monocytes, as Figure 3.10B.

(F) Proteins up or downregulated in primary CD4+ T cells, as in Figure 3.12.

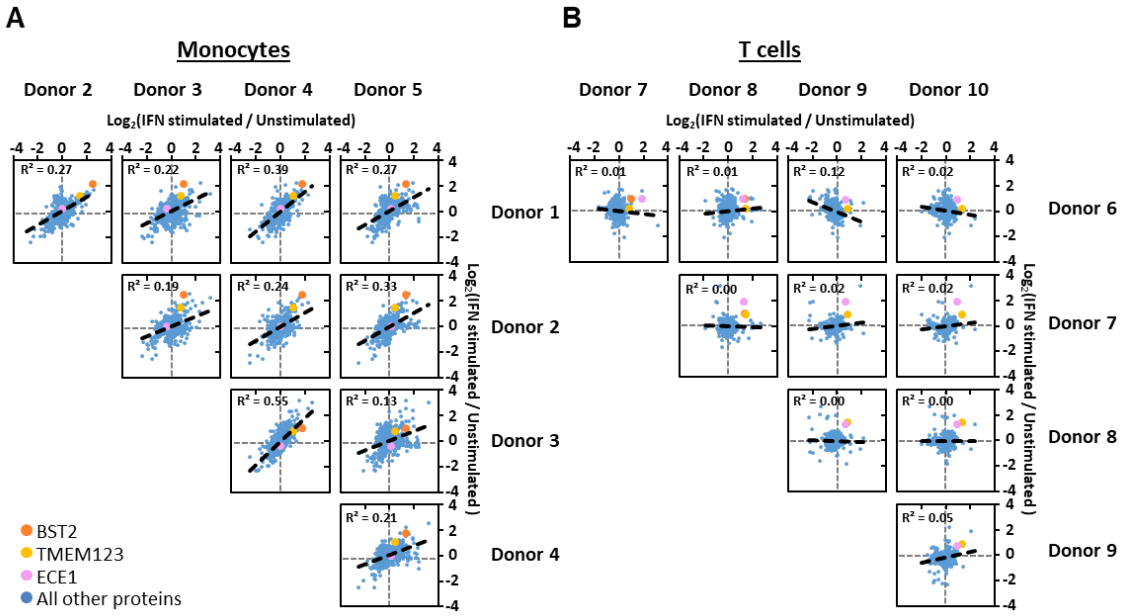


Figure 3.16 Variation in IFN induced changes between donors

Comparison of the IFN induced changes of each protein between donors, applying the same filtering as described for Figure 3.8B, for monocytes and T cells.

(A) Pairwise comparison of IFN induced proteomic changes between monocyte donors.

(B) Pairwise comparison of IFN induced proteomic changes between T cell donors.

3.7 Validation

A subset of the proteins upregulated in primary monocytes, for which cell surface flow cytometry antibodies were readily available, were selected for validation (Figure 3.17A). The IFN stimulation of the well characterised ISG and HIV restriction factor BST2 was demonstrated by both the proteomic data and by flow cytometry (Figure 3.17B). Other proteins were also validated by cell surface flow cytometry, including CD38 which is involved in cell adhesion, migration and calcium signalling (Frasca *et al*, 2006), and SIGLEC1 which provides negative feedback, controlling production of IFN (Zheng *et al*, 2015) (Figure 3.17C).

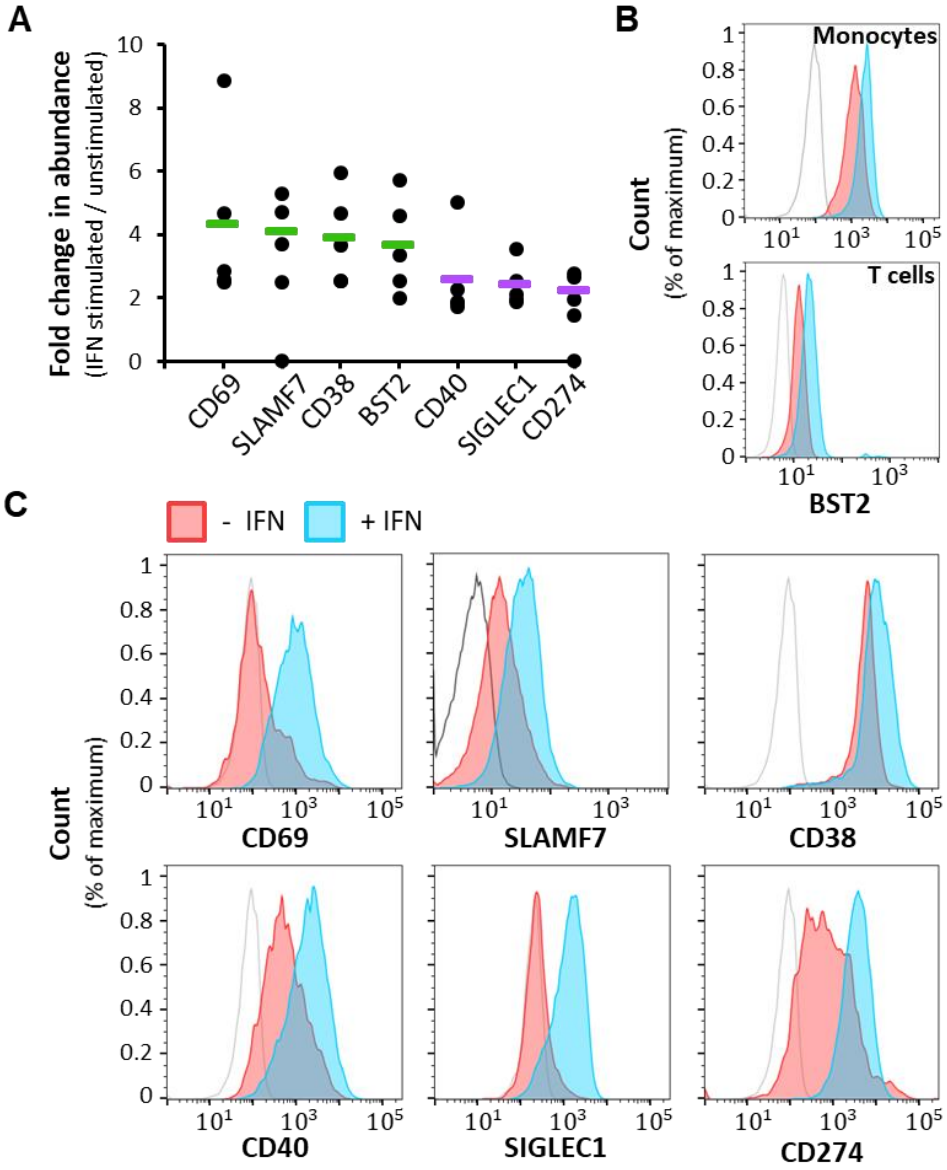


Figure 3.17 Validation of proteomic data on IFN stimulation of primary leukocytes

(A) Proteomic data on a subset of proteins that were IFN-stimulated in monocytes.

(B) Flow cytometry analysis of BST2 expression at the surface of primary monocytes (top) and CD4+ T cells (bottom), indicating IFN stimulation.

(C) Validation by flow cytometry of a selection of proteins stimulated by IFN from the proteomic data (as in A). CD69, CD38, CD40, SIGLEC1 and CD274 were examined in primary monocytes from one donor, whilst SLAMF7 was from a second donor. Red and blue lines show unstimulated and IFN-stimulated cells respectively. For SLAMF7 the grey line represents a control stained with the secondary antibody only. For all other samples, the antibody was conjugated to the fluorophore and the grey line represents the signal from an unstained sample.

3.8 TMEM123

3.8.1 IFN Stimulation of TMEM123

Aside from HLAs and the known HIV restriction factor BST2, TMEM123 was the only protein upregulated in both monocytes and T cells, in all donors where quantified (Figure 3.13, Figure 3.18A). It could therefore provide a pan-leukocyte marker for IFN stimulation, and may have as yet unrecognised roles in the immune response. IFN stimulation was validated by RT-qPCR in THP-1 cells (Figure 3.18B).

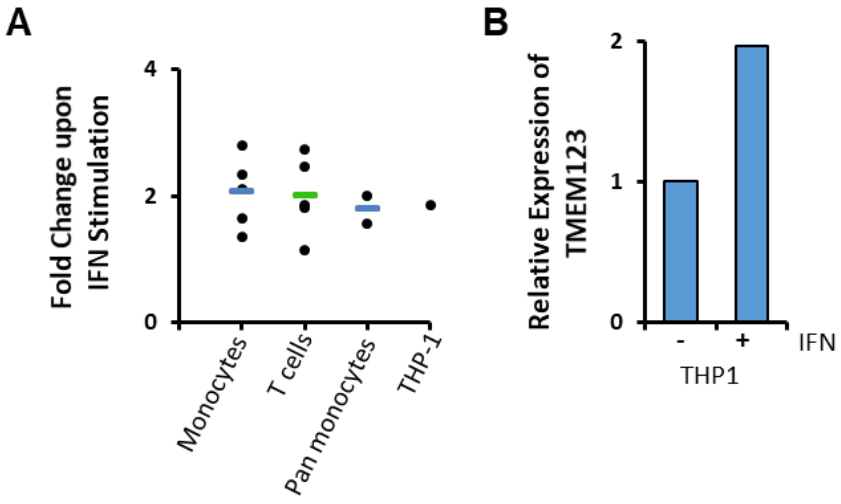


Figure 3.18 Validation of IFN stimulation of TMEM123

(A) FC in abundance of TMEM123 upon IFN stimulation in each cell type examined by proteomics.

(B) RT-qPCR validating IFN stimulation of TMEM123 in THP-1s.

3.8.2 TMEM123

Comparison of the IFN-stimulated proteins identified in monocytes and T cells revealed that the only protein upregulated in both cell types, aside from HLAs and the known HIV restriction factor BST2, was TMEM123. This protein may therefore be of interest as a pan-leukocyte marker of IFN stimulation. TMEM123 was originally identified by immunising mice with apoptotic jurkat cells and examining the antibodies raised. It is a type I transmembrane protein of 189 amino acids, with a molecular mass of 19 kDa,

though the observed mass varies substantially due to having many glycosylation sites. TMEM123 is also known as Porimin, as the antibody was found to cause pore formation at the PM in jurkats, inducing oncotic cell death. This particular form of cell death involves cell swelling and aggregation, formation of pores and membrane blebbing, but without DNA fragmentation or apoptotic bodies. It is unclear what biological stimulus could lead to porimin induced oncotic cell death. Cos7 and 293T cells transfected with porimin lost the ability to adhere to culture dishes, suggesting it may additionally be involved in cell adhesion, but these 293T cells did not undergo oncotic cell death following application of the anti-TMEM123 antibody (Zhang *et al*, 1998; Ma *et al*, 2001).

3.8.3 Testing of TMEM123 Antibodies

Since its identification in 1998, and cloning in 2001, TMEM123 has featured in very few publications. It was originally characterised using an in-house antibody, and we were not able to gain access to this (Zhang *et al*, 1998; Ma *et al*, 2001). Four different antibodies were tested (Table 3.5). The R&D antibody was conjugated to AF647, so was only tested for flow cytometry, whilst the others were tested for flow cytometry and immunoblot. Six different TMEM123 shRNA knockdown cell lines were produced in THP-1 cells, and validated by qPCR, and two of these were selected for testing of the antibodies. These had a 97 % and 83 % knockdown (Figure 3.19).

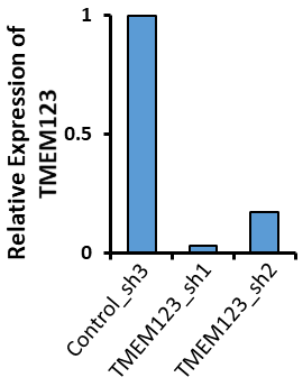


Figure 3.19 shRNA knockdowns of TMEM123

RT-qPCR validation of TMEM123 knockdown in two different shRNA THP-1 cell lines compared to a control shRNA. Knockdown was 97 % in TMEM123_sh1 and 83 % in TMEM123_sh2.

Table 3.5 Table of antibodies trialled for detection of TMEM123

Four antibodies were trialled for detection of TMEM123 by cell surface or intracellular flow cytometry and immunoblot. Applications tested by supplier: flow cytometry (FCM), immunohistochemistry (IHC), western blot (WB), immunofluorescence (IF). Residues were given in terms of: the 208 amino acid long isoform 1 (iso1), isoform 2 which is missing residues 34-52 (iso2), or an epitope within the stated region of mouse TMEM123 for the Santa Cruz antibody (mouse).

Supplier	Product Code	Species	Polyclonal or Monoclonal?	Residues Raised Against	Applications Tested by Supplier
R&D	FAB3010R	Mouse	Monoclonal	27-149 (iso2)	FCM
Santa Cruz	SC-377295	Mouse	Monoclonal	81-113 (mouse)	WB, IF
Abcam	ab81423	Rabbit	Polyclonal	145-194 (iso1)	WB
Novus	NB100-56371	Rabbit	Polyclonal	173-188 (iso2)	WB, IHC

Cell Surface Flow Cytometry

As the proteomic screen considered PM proteins, cell surface flow cytometry would be the ideal technique for validation and further experiments. Unfortunately, in all cases, the signal in the knockdown was the same as that in the control (Figure 3.20). Residues 27-166 of TMEM123 are extracellular, equivalent to residues 27-147 of isoform 2, a splice variant where residues 34-52 are missing. Therefore, the Novus antibody would not be expected to detect this extracellular region and work for cell surface flow cytometry. Likewise many of the potential epitopes for the polyclonal Abcam antibody may not be extracellular.

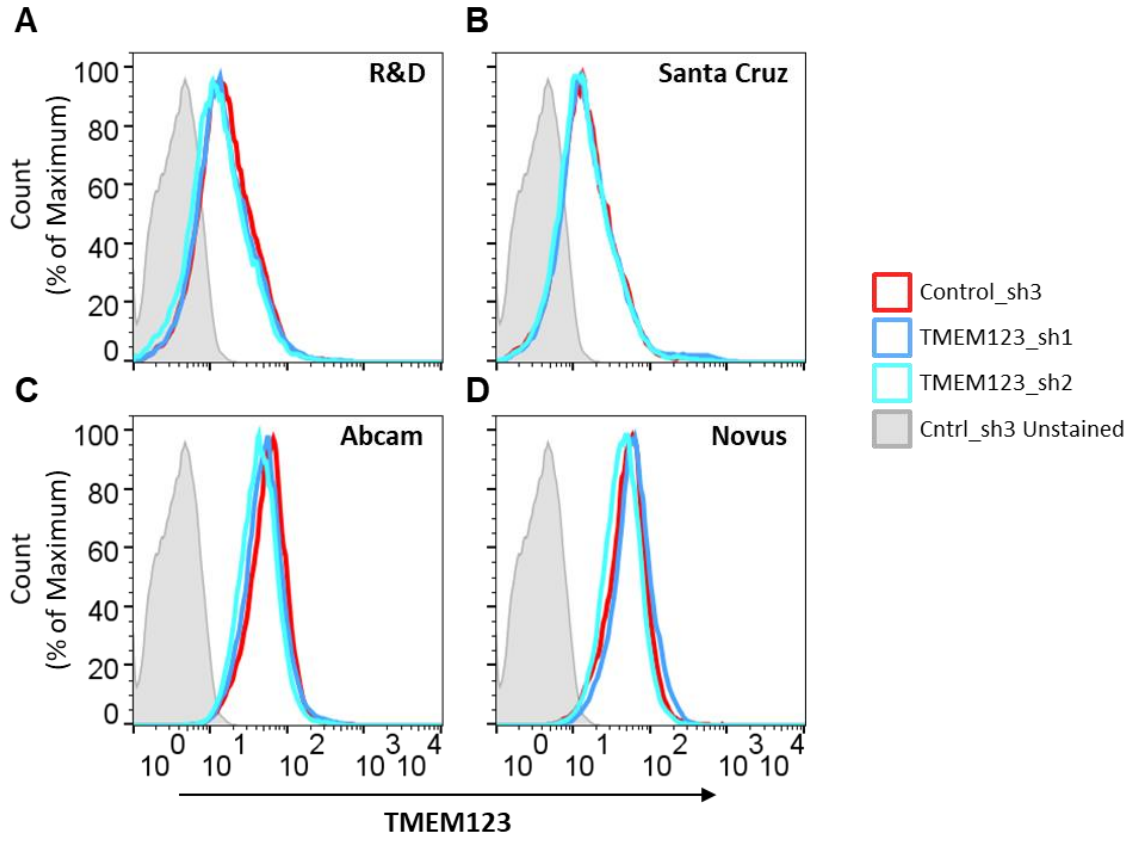


Figure 3.20 Testing TMEM123 antibodies on shRNA knockdowns THP-1 cells by cell surface flow cytometry

Four different primary antibodies were incubated with THP-1 cells expressing either a control shRNA or one targeting TMEM123. For (B) - (D), unconjugated antibodies were used. Although control samples were included where they were only stained with a secondary antibody, this was on IFN-stimulated cells, so the data is not shown, but good separation was observed.

(A) R&D antibody for TMEM123. This antibody was conjugated to AF647.

(B) Santa Cruz antibody for TMEM123. The sample was additionally stained with an anti-mouse secondary conjugated to AF647.

(C) Abcam antibody for TMEM123. The sample was additionally stained with an anti-rabbit secondary conjugated to AF647.

(D) Novus antibody for TMEM123. The sample was additionally stained with an anti-rabbit secondary conjugated to AF647.

Intracellular Flow Cytometry

The antibodies were additionally all tested for intracellular flow cytometry. The same signal was again observed for the shRNA knockdowns and the control THP-1s, suggesting the antibodies do not detect TMEM123 by flow cytometry.

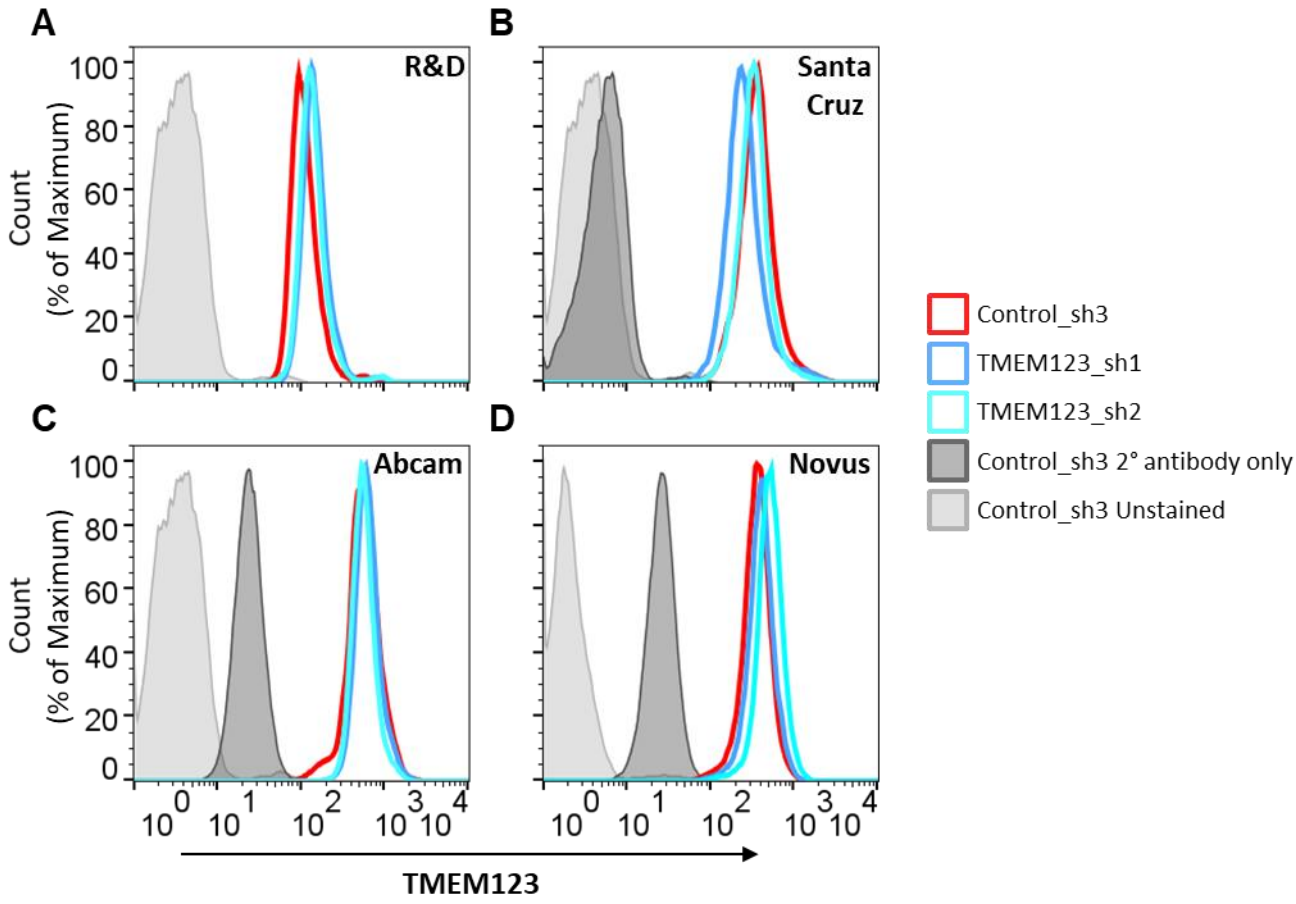


Figure 3.21 Testing TMEM123 antibodies on shRNA knockdown THP-1 cells by intracellular flow cytometry

THP-1 cells expressing either a control shRNA or one targeting TMEM123 were fixed and permeabilised, before being stained with one of four different antibodies for TMEM123.

(A) R&D antibody for TMEM123. This antibody was conjugated to AF647.

(B) Santa Cruz antibody for TMEM123. The sample was additionally stained with an anti-mouse secondary conjugated to AF647.

(C) Abcam antibody for TMEM123. The sample was additionally stained with an anti-rabbit secondary conjugated to AF647.

(D) Novus antibody for TMEM123. The sample was additionally stained with an anti-rabbit secondary conjugated to AF647.

Immunoblot

The Novus, Santa Cruz and Abcam antibodies were tested for detection of TMEM123 by immunoblot. This was initially performed on lysates made using RIPA or 2 % SDS, with 60 μg of protein being separated by gel electrophoresis, and transferred using a semi-dry system (Figure 3.22A). There were no bands detected using the Santa Cruz antibody, and multiple bands were observed using the Novus and Abcam antibodies.

On subsequent gels, 20 μg of protein was loaded in order to reduce noise. Additionally, the standard protocol involved denaturation of the protein by boiling at 95 °C for 5 mins. In the case of membrane proteins this can lead to aggregation, so heating at 60 °C for 20 mins or 37 °C for 30 mins was also tested. This was initially attempted using the semi-dry transfer system again, however the Novus antibody performed similarly, and Abcam and Santa Cruz staining did not show any bands. As no bands were seen with the Santa Cruz antibody on either test, it was not used again.

The same experiment was then performed using a wet transfer system (Figure 3.22B). In this case, the Novus antibody again produced many bands, with none at the expected size described by the supplier (Figure 3.22C), or correlating to the expected control and knockdown expression levels. The Abcam antibody provided some more promising results using RIPA lysates, with bands that seem to correlate to the control and knockdown. However, these bands suggest a much larger protein than expected, and the results were not always recapitulated when repeating using other lysates.

TMEM123 has a predicted molecular mass of ~19 kDa. Both the Novus and Santa Cruz example immunoblots (from the websites) suggest bands are detected at a substantially higher molecular weight, whilst the Abcam example probes a placenta lysate and detects a protein ~25 kDa. When TMEM123 was originally cloned, immunoblot analysis of transfected 293, Cos7 and HeLa cells found the molecular mass ranged from 55 kDa to 80 kDa, and an endogenous molecular mass of 110 kDa was confirmed in Jurkat cells. It was suggested that the differences were likely due to glycosylation (Ma *et al*, 2001). It is therefore difficult to determine whether the band observed using the Abcam antibody, or any of the ones seen using the Novus antibody, are at the correct molecular mass to indicate detection of TMEM123 in THP-1s.

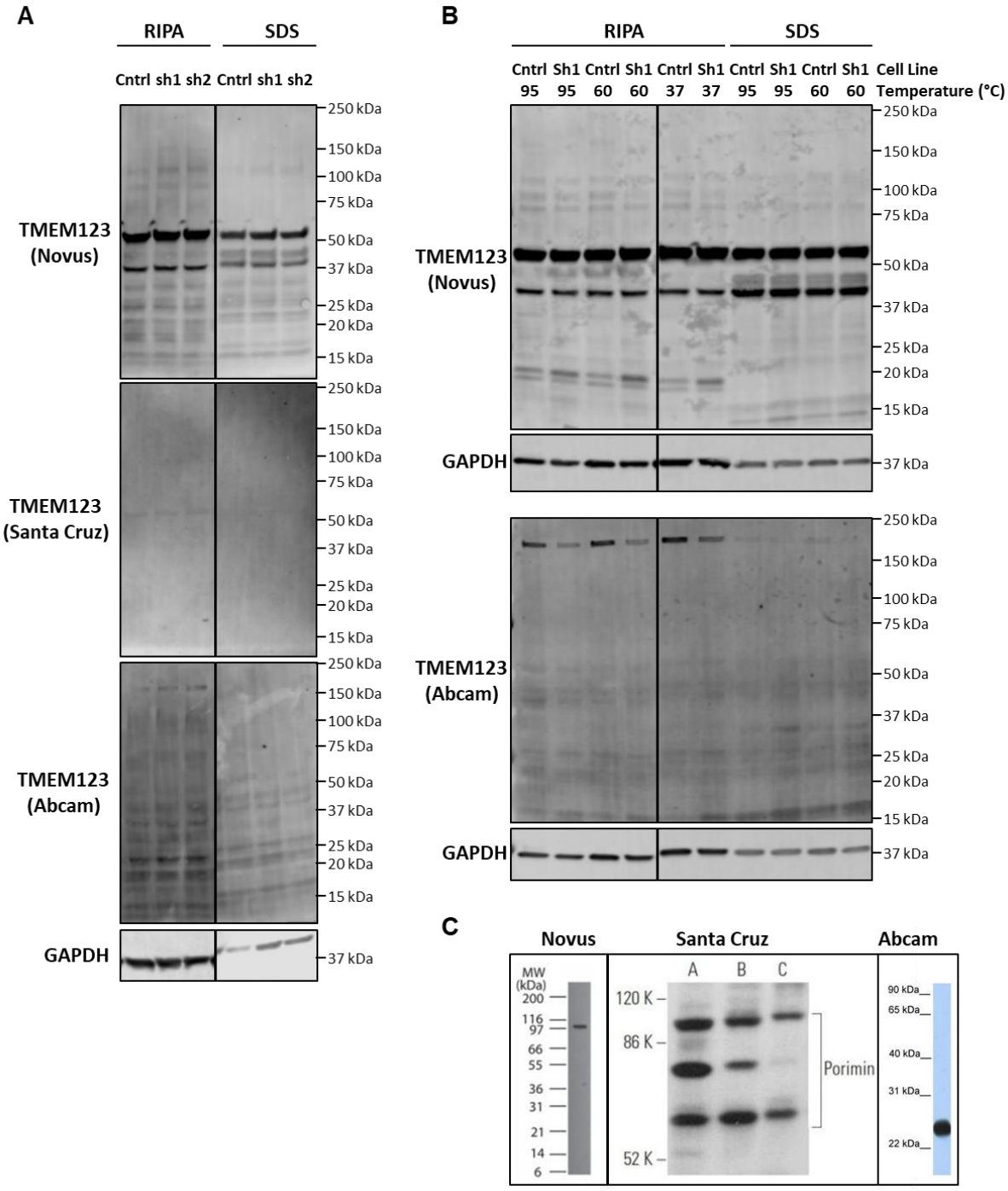


Figure 3.22 Testing TMEM123 Antibodies on shRNA knockdown THP-1 cells by Immunoblot

(legend on next page)

Figure 3.22 Testing TMEM123 Antibodies on shRNA knockdown THP-1 cells by Immunoblot

(A) An shRNA control cell line (control_sh3) and TMEM123 knockdown cell lines (TMEM123_sh1 and TMEM123_sh2), were lysed using either RIPA or 2 % SDS, and then boiled at 95 °C for 5 mins. 60 µg of protein from each lysate was separated by gel electrophoresis before transfer using a semi-dry transfer system. One membrane was stained with Novus anti-TMEM123 (rabbit) and Santa Cruz anti-TMEM123 (mouse), whilst the other was stained with Abcam anti-TMEM123 (rabbit) and anti-GAPDH (mouse). The dark line between the RIPA and SDS lysates indicates where superfluous lanes in the gel have been digitally removed.

(B) Cells expressing a control shRNA (sh3) or TMEM123_sh1 were lysed using either RIPA or 2 % SDS, and then incubated at 95 °C for 5 mins, 60 °C for 20 mins or 37 °C for 30 mins (as indicated). 20 µg of proteins from each lysate was separated by gel electrophoresis, and proteins were transferred to a membrane using a wet transfer system, prior to staining with the Novus or Abcam anti-TMEM123 antibodies and anti-GAPDH. The dark line between sh5 at 60 °C and control at 95 °C indicates where superfluous lanes in the gel have been digitally removed.

(C) Example immunoblots given on the website for the tested antibodies. The Novus antibody probes HL60 cell lysates, the Santa Cruz one is trialled on HeLa cells (A), RAW 264.7 cells (B) and BYDP (C) whole cell lysates, and the Abcam antibody is depicted in an immunoblot of placenta lysate.

3.9 Discussion

This study represents the first systematic analysis of the effects of type I IFN at the cell surface of primary leukocytes. Here, 607 and 485 cell surface proteins were quantified in primary monocytes and T cells respectively, providing a more comprehensive analysis of the cell surface proteome and the effects of IFN than was previously conducted. The use of TMT technology enabled multiplexing of the samples in order to investigate five donors in parallel, allowing identification of the most consistent effects and the investigation of donor-to-donor variation.

3.9.1 Cell Surface Proteome

A previous investigation of the T cell surface proteome identified 173 proteins using MS (Graessel *et al*, 2015). A thorough literature search by the authors suggested that 24 of

these proteins had not previously been associated with T cells. Nine of these were quantified in the PM data presented here: apolipoprotein C-III (APOC3), Sodium/potassium-transporting ATPase subunit alpha-1 (ATP1A1), transmembrane protein C16ORF54, ECE1, ecotropic viral integration site 2A protein homolog (EVI2A), cell surface A33 antigen (GPA33), neuroplastin (NPTN), ADP/ATP translocase 2 (SLC25A5), Sodium bicarbonate cotransporter 3 (SLC4A7). An additional four were found in the data presented here, if it was not filtered for PM annotated proteins: RING finger protein 149 (RNF149), suprabasin (SBSN), phosphate carrier protein mitochondrial (SLC25A3), transmembrane protein 2 (TMEM2). This corroborates some of the findings in this data, as well as that of Graessel *et al.*, and suggests that this proteomic data may be expanded further by not limiting the investigation to GO annotated PM proteins.

Analysis of the dataset presented here revealed that CD44, SPN and PTPRC were some of the most abundant proteins at the surface of the monocytes and T cells. CD44 is a highly glycosylated transmembrane protein, abundant at the surface of multiple leukocytes, and is involved in adhesion (Senbanjo & Chellaiah, 2017). SPN (also known as CD43) is also involved in adhesion, as well as immune cell activation (Rosenstein *et al*, 1999; Kyoizumi *et al*, 2004), and receptor-type tyrosine-protein phosphatase C (PTPRC, CD45) is another highly abundant proteins, involved in T cell activation (Stanford *et al*, 2012). Other proteins were more differentially expressed. For example, previous examination of the differential distribution of the GLUT1 and GLUT3 transporters was suggested to facilitate cellular function; GLUT3 has a higher affinity for glucose and its increased expression in monocytes may enhance its ability to combat infection in tissues with an impaired blood flow (Fu *et al*, 2004).

3.9.2 The Effects of IFN

Investigation of the IFN induced changes identified 41 consistently IFN-stimulated proteins in monocytes, and just 13 in T cells. More proteins were modulated by IFN and showed larger FCs in monocytes compared to T cells, consistent with previous transcriptomic analyses (Aso *et al*, 2019; Henig *et al*, 2013). These included multiple positive controls such as HLAs and BST2. Characterisation of the IFN response has

important implications in multiple fields. CD274 is also known as Programmed Death Ligand–1 (PD-L1), and was found to be upregulated in primary CD14+ and pan-monocytes as well as in THP-1s in this study. CD274 has previously been found to be stimulated by IFNα in a variety of murine immune cells and human DCs. Whilst many ISGs stimulated the immune response, CD274 is immunosuppressive, and this may have important implications for IFN based cancer therapies, with enhanced expression potentially contributing to tumour immune evasion (Bazhin *et al*, 2018).

Variability between Donors

TMEM123, BST2 and HLAs were some of the most consistently IFN-stimulated protein, both between cell types, but also between donors. Whilst the cell surface proteomes were remarkably invariant, much greater variability was observed in the IFN responses, particularly in T cells. One reason for the poor correlation may be that the majority of proteins were not IFN responsive. Many FCs were therefore just small shifts that centre around zero. Also, the CD4+ T cell population examined here is likely to be highly heterogeneous, and different between donors, possibly contributing to the different effects of IFN across donors. The presence of CD44, CD62L (L-selectin, SELL) and CCR7 suggest the majority were naïve or memory T cells. Additionally, activated T cell markers include CD25 (IL2RA) and CD69, neither of which were quantified in the CD4+ T cell data (Sondel *et al*, 2003; Tan & Surh, 2006). Although this does not definitively determine that CD25 and CD69 were not expressed, it gives a suggestion of the T cell populations present, and some confidence that the cells were not activated by the extraction and stimulation process. There may naturally also be more variability between IFN responses in T cell donors, a phenomenon which has been reported previously (Schlaak *et al*, 2002). The changes in monocytes showed a larger, more consistent IFN response between donors, though there was still a lot of variability. This highlights the benefit of multiplexing samples for the examination of multiple donors. An interesting follow up experiment would have been to examine the IFN responses of the same individuals on multiple occasions to see if it is consistent between replicates from the same donor.

TMEM123

Other than HLAs and BST2, TMEM123 was the only protein upregulated in both primary monocytes and T cells by IFNα2a. This may therefore provide a pan-leukocyte marker for IFN stimulation. The protein was only quantified by a single peptide in T cells, and three peptides in monocytes. It is possible that the peptides were misidentified and the observation requires validation. Encouragingly however, the peptides sequences only aligned to the two isoforms of TMEM123, and were not redundant to any other proteins. Little is currently known about TMEM123. Also known as Porimin, it was originally identified due to the ability of anti-porimin to cause oncotic cell death in jurkats (Zhang *et al*, 1998; Ma *et al*, 2001). Further characterisation of the protein would be interesting, though was unfortunately limited in this investigation by the lack of available reagents. Testing the antibodies was additionally complicated by not knowing the expected size of the protein in different cell lines. It would therefore be beneficial to retest them on a cell line overexpressing a tagged version of the protein, so the expected size could be confirmed. This process was started, however validation of the TMEM123 overexpression by RT-qPCR failed and due to time constraints this has not yet been investigated further.

Further Avenues for Analysis of the Data

Further data from the screen may be gained by extending the analysis beyond proteins with PM related GO annotations. Comparison to previous studies showed some proteins previously identified as PM were filtered out here (Graessel *et al*, 2015). Only requiring a 'membrane' annotation would include a much broader range of proteins. Many of these may not be specifically cell surface proteins, and may have been detected in the proteomics data as they were abundant intracellular contaminants, however incorporating them would reduce the chance of missing interesting proteins due to poor annotation. Additionally, only proteins meeting the described filtering criteria were analysed; reducing the stringency to incorporate more donor-to-donor variation may therefore provide further information.

Studying IFN-stimulated proteins alongside viral infection data may also aid in identification of ARFs. BST2 was upregulated in this data on monocytes and T cells. It

is an extensively studied HIV restriction factor, which prevents HIV exit from the cell by tethering it to the cell surface, and is antagonised by the HIV protein Vpu (Neil *et al*, 2008; Hammonds *et al*, 2010; Mitchell *et al*, 2009). It was one of the only proteins upregulated in all the cell types examined in this study. ECE1 was one of the most highly upregulated proteins quantified in all five T cell donors. This theme is discussed further in chapter 5.

This investigation provides a valuable characterisation of the effects of IFN in primary leukocytes across multiple donors, complementing and extending previous transcriptomic studies. It captures a variety of effects of cell type and donors, crucial for a complete understanding of the IFN response, valuable in development of IFN related therapies and for identification of ARFs.

CHAPTER 4: VACCINIA VIRUS SCREEN

Quantitative Temporal Proteomic Analysis of Vaccinia Virus Infection

The work in this chapter is adapted from the publication 'Quantitative Temporal Proteomic Analysis of Vaccinia Virus Infection Reveals Regulation of Histone Deacetylases by an Interferon Antagonist' published in Cell Reports, May 2019 (Soday et al, 2019). Therefore many of the figures here resemble those in the manuscript. This publication was produced in collaboration with the laboratory of Professor Geoffrey L. Smith, Department of Pathology (University of Cambridge). I am joint first author on this paper, and was responsible for processing the samples for proteomic analysis, as well as the analysis of all proteomic data and production of related figures. Dr Jonas Albarnaz (Smith Lab) performed the VACV infections and harvesting for proteomic experiments, as well as immunoblots for validation. Dr Yongxu Lu (Smith Lab) conducted experiments relating to HDAC5. Experiments performed by others are indicated throughout the text and figure legends in this chapter, as well as in the relevant methods.

4.1 Introduction and Aims

4.1.1 Introduction

VACV and other members of the *Poxviridae* family have a multitude of mechanisms for manipulating host immunity, dedicating a considerable proportion of their coding capacity to this (Burshtyn, 2013). Multiple VACV proteins are involved in inhibiting the IFN response, to which VACV is particular susceptible due to replicating in the cytoplasm. Furthermore, several proteins are involved in antagonising ISGs, and others may act as ligands for NK cells; for example the VACV protein N1 may limit NK cell activity, and A40 resembles a C-type lectin and some NK cell receptors, and contributes to virulence (Jacobs *et al*, 2008; Wilcock *et al*, 1999; Tscharke *et al*, 2002).

One of the criteria defined for restriction factors is antagonism by a viral factor, and this may include downregulation of the protein. Identifying proteins which are modulated by

viral infection may highlight those that are important in viral infection and immune evasion, as was previously demonstrated for HCMV (Weekes *et al*, 2014; Nightingale *et al*, 2018). In VACV infection, IFITM3 is downregulated in order to evade restriction, though a mechanism for the modulation of IFITM3 by VACV, or how its restricts infection, has not been determined (Li *et al*, 2018). A comprehensive analysis of the VACV proteome and which host proteins are modulated during infection could identify additional ways in which the virus evades the immune response. To this end, a proteomic time-course was conducted, investigating the changes in protein abundance throughout infection.

The time-course data can also be used to examine the kinetics of expression of viral proteins, which can assist in predicting their function, and defining host-virus interactions. As VACV is a model for host-virus interactions, this could highlight potential therapeutic targets for antiviral treatments, as well as improving understanding of VACV biology; this is essential for its development as a vaccine, both in terms of safety and optimizing immunogenicity.

The Western Reserve (WR) strain of VACV was used for infection of HFFF-TERTs in this investigation. This is a commonly used lab strain which, unlike some of the more attenuated strains of the virus, can replicate efficiently in mammalian cells. It is also considered to be one of the most virulent strains in animals, and can be used for *in vivo* experiments in mice. This virulence has been correlated with increased modulation of the immune response compared to less virulent strains, so it may be particularly valuable for identifying facets of the immune response (de Freitas *et al*, 2018). The HFFF-TERT cell line was chosen in order to be comparable with proteomic investigations of other virus (Weekes *et al*, 2014; Soh *et al*, 2019). As many restriction factors are able to target multiple viruses (Table 1) comparison of this data can be used to identify proteins which may be important in infection.

4.1.2 Summary

In this chapter, a quantitative, multiplexed, proteomic screen over the whole time-course of VACV infection is described. Around 9000 host proteins and ~80% of VACV proteins

were quantified, providing a comprehensive and dynamic view of proteomic changes during infection. VACV downregulated 265 proteins more than two fold, and additional investigation using the proteasome inhibitor MG132 revealed that 69 % of these were specifically targeted for proteasomal degradation. Many of these were cell surface proteins, and enrichment analysis revealed targeting of certain classes, in particular collagens, cadherins and innate immune proteins. The viral proteins were also analysed, and four temporal classes identified, which correlate well with previous transcriptomic studies and known functions. The data was used to predict an interaction between the early viral protein C6 and the host protein HDAC5. C6 was found to target HDAC5 for degradation, and HDAC5 was demonstrated to restrict viral replication.

4.1.3 Aims

1) Perform an unbiased, systematic and quantitative analysis of proteomic changes in the host and virus across a single replication cycle of VACV infection.
2) Determine which proteins may be of particular importance during infection by looking at enriched groups of proteins, overlap with other viruses, and examination of which proteins are being specifically targeted for proteasomal degradation.
3) Classify viral proteins into temporal classes according to their kinetics.
4) Examine how the data can be used to predict host-virus interactions.

4.2 Quantitative Temporal Analysis of VACV Infection

4.2.1 Optimisation of the Infection Time Course

In an initial experiment, Dr Jonas Albarnaz (Smith Lab, Department of Pathology, University of Cambridge) infected HFFF-TERTs with the WR strain of VACV at a high MOI (5.0), and harvested samples for whole cell protein analysis. Mock samples were harvested at 0, 4 and 8 hpi, and VACV infected samples at 2, 4, 6, 8 and 12 hpi. Mock and infected samples that were additionally treated with the DNA replication inhibitor AraC were harvested at 4 h. From a single shot analysis, 1607 proteins were quantified including 48 viral proteins.

The experiment technically appeared to work well; hierarchical clustering showed that mock and infected samples clustered together, and proteins that were modulated by infection largely changed in a linear fashion. However, the fold downregulation of host proteins observed was modest and viral proteins expressed late during infection still appeared to be increasing in abundance by the end of the time-course. It was therefore decided to improve the set up by considering up to 18 h of infection.

Additionally, AraC was expected to limit expression of viral proteins dependent on viral DNA replication. Comparison of the 4 hpi time point with and without AraC did not show much change, as these viral proteins were expressed later in infection and were therefore not expressed highly at 4 hpi. The AraC treated samples were therefore included at a later time point in subsequent experiments.

4.2.2 Experimental Outline

Based on the initial experiment, the outline of the time-course was optimised, extending it to 18 hpi and investigating the effects of AraC at 6 hpi (Figure 4.1). At this stage we were also fortunate to have access to an 11th TMT, allowing the 10-plex experiment to be expanded. Mock infected samples were harvested at 3 time points (0, 6 and 18 h), and 6 time points of VACV infection were now included, with samples harvested at 2, 4, 6, 8, 12 and 18 hpi. Mock and infected samples additionally treated with AraC were harvested

at 6 h. Infection was again of HFFF-TERTs at a high MOI (5.0), and flow cytometry confirmed that > 95 % of cells were infected. The time-course was conducted three times, with infections and harvesting performed by Dr Jonas Albarnaz. The 11 samples for each time-course were given to me for further processing. They were digested into peptides, labelled with TMT and subjected to MS3 analysis.

The expression profiles of 8,991 human proteins and 172/216 viral proteins were quantified, with 7,316 human and 160 viral proteins quantified in all three replicates of the time-course (Table 4.1). The FC of each protein compared to the average of the three mock samples was examined by hierarchical clustering, and as expected, mock, early and late viral infection time points all clustered separately (Figure 4.2). Changes were largely linear with those of the greatest magnitude occurring late during infection.

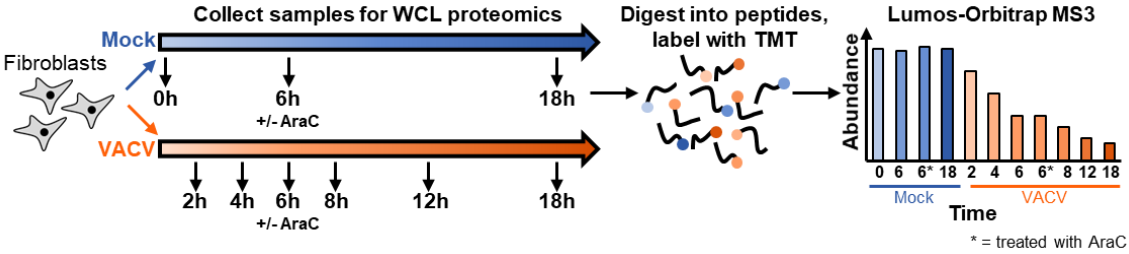


Figure 4.1 Schematic of workflow for proteomic analysis of VACV infection

Cells were either mock treated or infected with VACV at an MOI of 5. One mock and one infected sample at 6 h were treated with AraC, a viral DNA replication inhibitor. The 11 samples were labelled with TMT, pooled and analysed by MS3 analysis.

Table 4.1 Table of peptides and proteins quantified in each proteomic experiment on VACV discussed in this chapter

The time-course was repeated in triplicate harvesting samples at various time points throughout infection. The MG / C6 data includes samples of WT infection with and without MG132, and also investigated the effects of deleting the viral protein C6. For each experiment, the number of proteins quantified is given.

	Time Course 1	Time Course 2	Time Course 3	Any Time Course	All Time Courses	MG / C6 data
Total peptides	120345	124779	112310			123980
Human proteins	8229	8310	8076	8991	7316	8263
Viral proteins	166	169	169	172	160	173

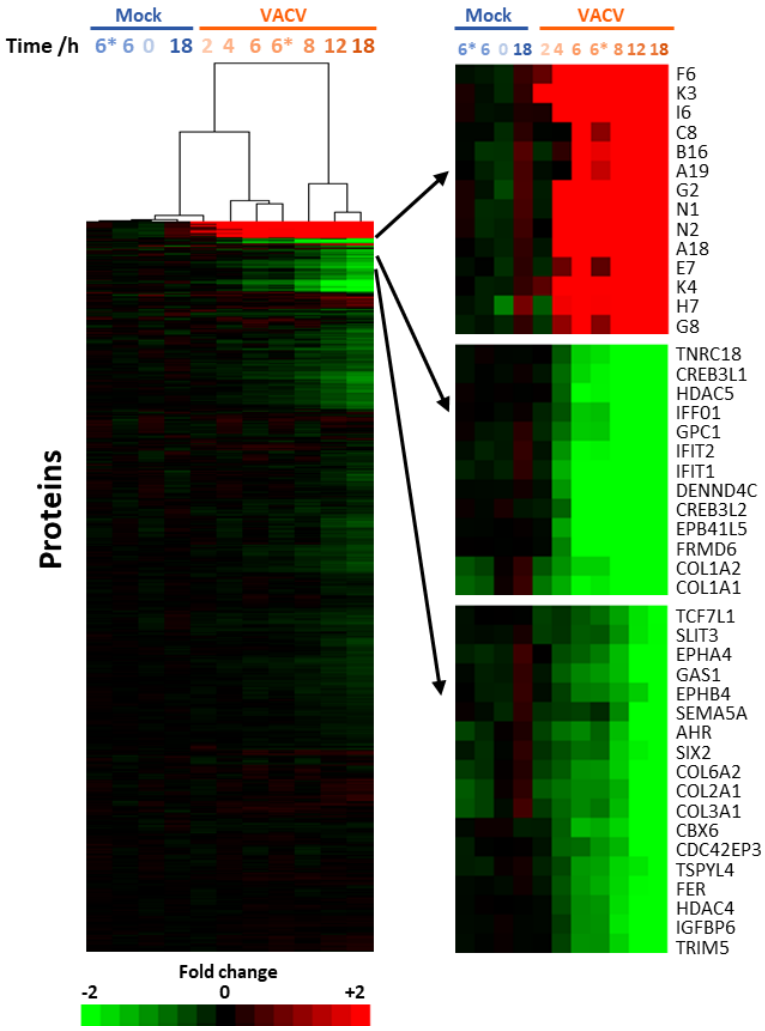


Figure 4.2 Hierarchical clustering of VACV infection time-course

Hierarchical clustering of proteins quantified in all three biological replicates, with fold change compared to the average of the three mock samples. Some subclusters are highlighted on the right, showing particularly up- (in red, top panel of viral proteins) or down- (in green, middle and bottom panels) regulated proteins.

4.3 Modulation of Host Proteins by VACV Infection

The three mock time points exhibited very similar profiles (Figure 4.3A). Therefore in all further analyses of the triplicate time-course data in this chapter, FC of proteins was compared to the 18 h mock sample. A mean FC for each time point was calculated by taking the FC at the given time point compared to the 18 h mock, and averaging this across all replicates where the protein was quantified. The various plots of expression profiles in this chapter, as in Figure 4.3B, display this mean FC.

In order to examine which host proteins may be important in infection, the proteins downregulated during infection were examined. For a comprehensive analysis of downregulated proteins, a 'sensitive' filter was applied, examining proteins that had a mean FC > 2 at any time point. A shorter, 'stringent' list identified proteins which were quantified in all three replicates, had a mean FC >2 and a BH-corrected p-value <0.05 (Table 4.2, Appendix Table 2, Appendix Table 3). In practice, the proteins highlighted by the sensitive list were used for most analyses in order to encompass all changes, and to avoid exclusion of proteins because they were not quantified in all three time-courses.

Applying the sensitive criteria, 265 human proteins were downregulated >2 fold, and 70 human proteins were upregulated; these included cellular tumour antigen p53 (Tp53), proto-oncogene FOS and early growth response protein 1 (EGR1), proteins known to be modulated by VACV (de Magalhaes *et al*, 2001; Silva *et al*, 2006; Wali & Strayer, 1999; Yoo *et al*, 2008), Figure 4.3B). The majority of changes could be seen by comparing the 18 h VACV infected sample to the 18 h mock, as the greatest magnitude of change was observed at this latest time point in infection (Figure 4.3C). Visual inspection of this 18 h data on a dot plot highlights the downregulation of multiple collagens and IFIT proteins. The complete dataset is included in the supplemental excel workbook, Table S1, of the published manuscript (Soday *et al*, 2019). The excel 'plotter' in this enables interactive generation of graphs displaying the relative abundance of all proteins quantified in the proteomic datasets presented in this chapter.

Table 4.2 Number of host proteins modulated by VACV infection

The number of proteins modulated by VACV infection according to 'sensitive' or 'stringent' criteria is given.

	Sensitive Criteria (Mean FC > 2 at any time point)	**Stringent Criteria** (Quantified in all 3 time courses; mean FC > 2 and p <0.05 at any time point)
Total Proteins	8991	7316
Downregulated	265	142
Upregulated	70	4

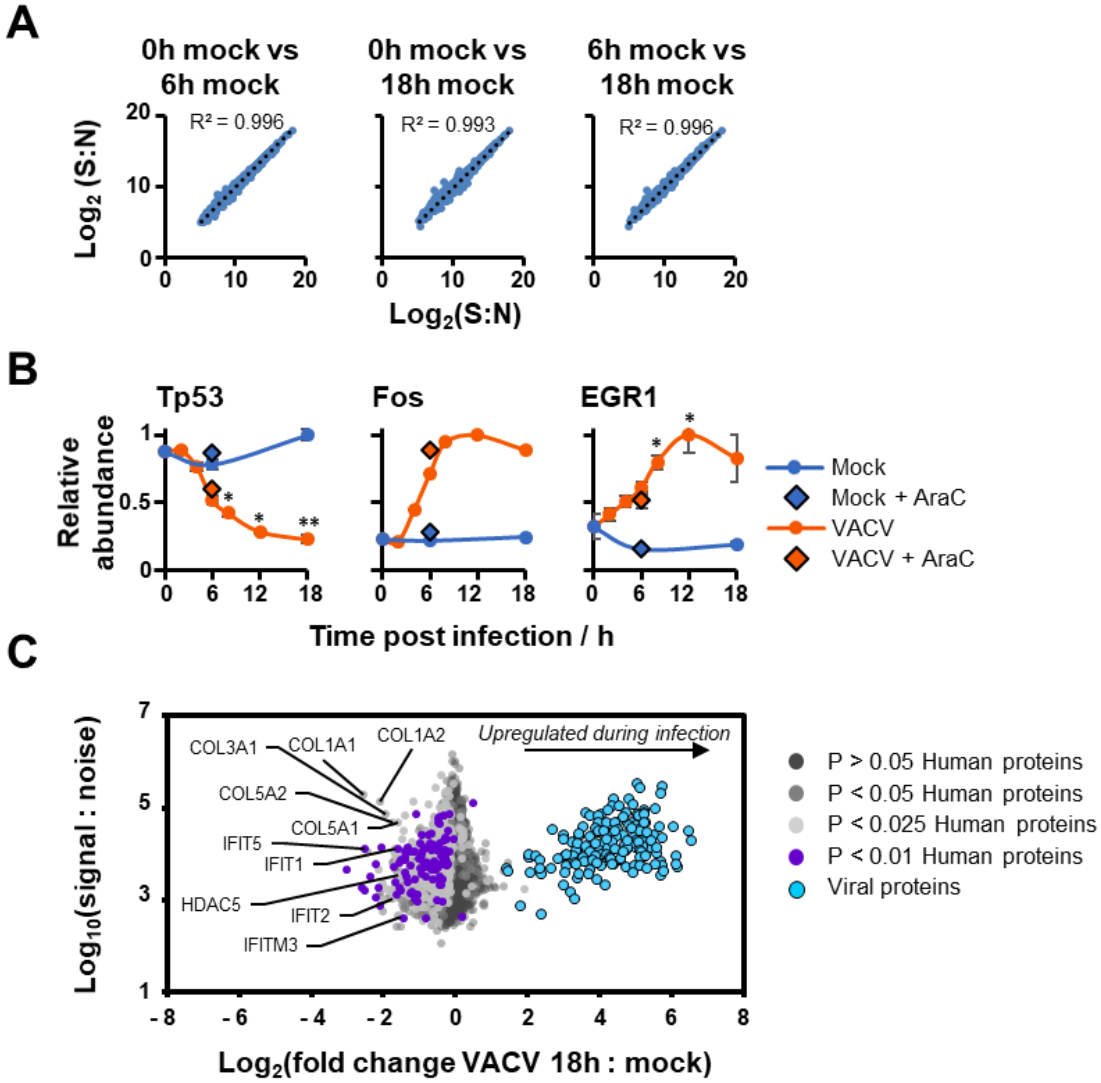


Figure 4.3 Quantitative temporal analysis of VACV infection

(A) Correspondence between mock samples. Each graph shows log2(normalised S:N) for proteins quantified in all three biological replicates.

(B) Expression profiles of proteins previously reported to be modulated by VACV. Data are represented as mean ± SEM; *p<0.05, **p<0.01. Error bars and statistics are not included on the plot for Fos as this protein was not quantified in all three replicates.

(C) Scatterplot of proteins quantified in all three biological replicates at 18 hpi compared to the 18 h mock; each dot on the plot denotes an individual protein. The FC (on the x-axis) was determined by averaging the FC in each replicate. The y-axis displays the normalised S:N, indicating the accuracy of quantitation. P-values were estimated using a BH-corrected two-tailed t-test.

4.4 Downregulation of Cell Surface Proteins and Antiviral Factors

Reasoning that proteins modulated by VACV may be of particular importance, several approaches were employed to perform a more systematic, unbiased exploration of proteins downregulated during infection.

4.4.1 Enrichment Analysis of Downregulated Proteins

DAVID enrichment analysis was used to identify enriched groups of proteins amongst those downregulated during infection according to the sensitive criteria (Huang *et al*, 2009a, 2009b). This examined protein annotations from various databases, including GO terms, Interpro domains and UniProt keywords, to find features enriched in the downregulated proteins compared to the total sample of proteins quantified. Interestingly, multiple clusters referred to downregulation of cell surface receptors and ligands, including terms such as 'extracellular', 'cell attachment site' and 'immunoglobulin' (Figure 4.4A). This suggests that the cell surface may be a key site for VACV modulation. Enrichment analysis also highlighted terms relating to collagens, innate immunity and cadherins amongst others. Some examples of the expression profiles of proteins identified in these clusters are shown (Figure 4.4B) and several of these were validated by immunoblot (Figure 4.6C, immunoblots performed by Dr Jonas Albarnaz). In total, 11 collagens and 3 related proteins were quantified in the collagen cluster, 13 proteins were in the innate immunity cluster and the 4 protocadherins shown were identified in the cadherin cluster.

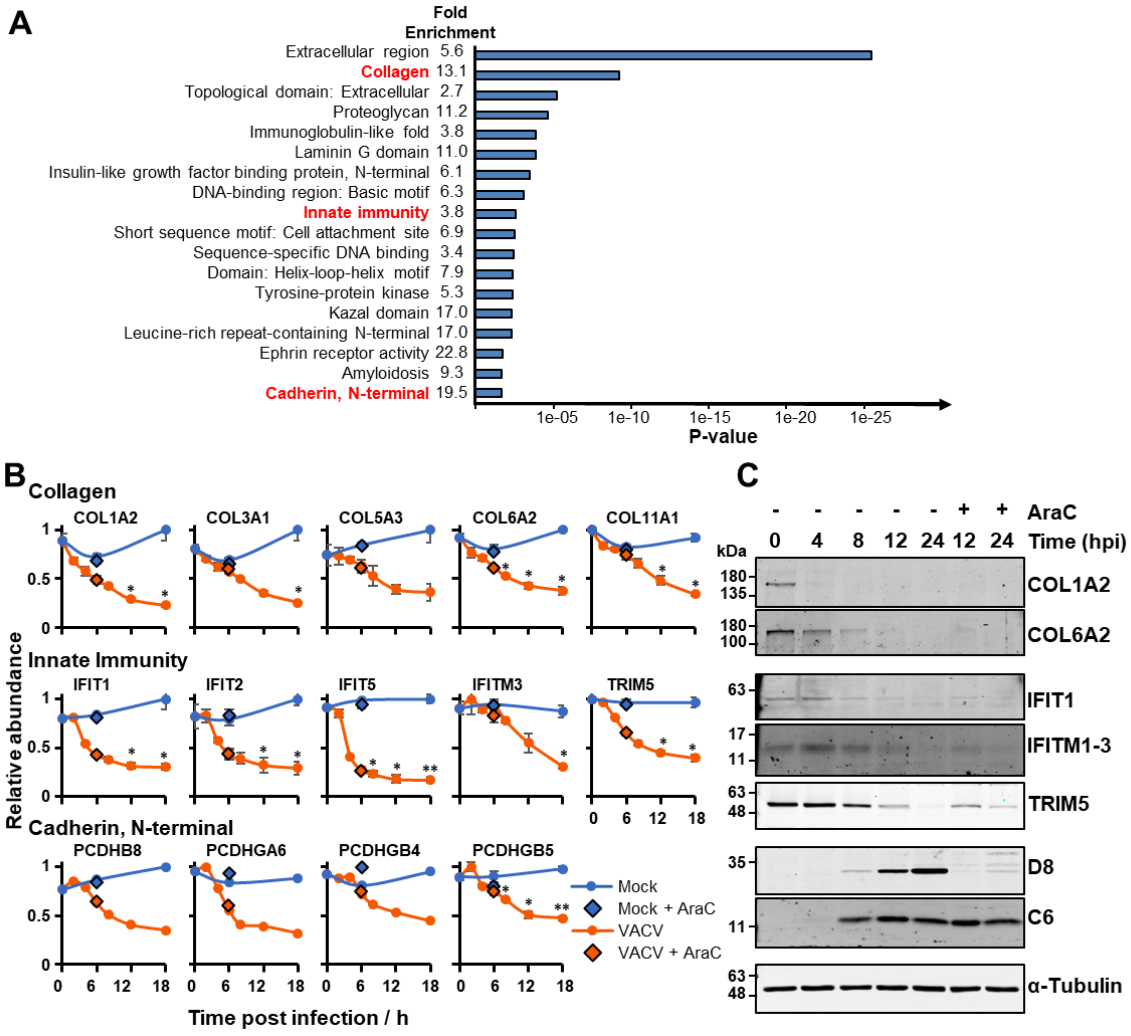


Figure 4.4 Downregulation of multiple collagens, protocadherins and innate immune mediators during VACV infection

(A) DAVID functional enrichment analysis of all proteins downregulated by VACV (using sensitive criteria). A background of all quantified human proteins was used. Displayed are representative terms from each cluster with a BH corrected $p < 0.05$. A similar analysis on upregulated proteins did not reveal any significantly enriched clusters.

(B) Examples of proteins from the collagen, innate immunity and cadherin clusters. Data are represented as mean ± SEM (n=3); p-values compare the indicated time point to the 18 h mock using a two-tailed t-test, *p<0.05, **p<0.01. Error bars and statistics are not included for protocadherins PCDHB8, PCDHGA6 and PCDHGB4 as these proteins were not quantified in all three replicates.

(C) Validation of expression profiles of some proteins shown in (B) by immunoblot. HFFF-TERTs were infected with VACV at an MOI of 5, and the viral proteins D8 and C6 represent late and early viral gene expression respectively (experiments performed by Dr Jonas Albarnaz).

4.4.2 Overlap with HCMV

Proteins of particular importance in defending the host against an infection may be targeted by multiple viruses (Schoggins *et al*, 2011; Schreiner & Wodrich, 2013). This dataset on VACV infection was overlapped with previous data on the unrelated dsDNA virus, HCMV (Figure 4.5A). The three VACV time-courses were included, with downregulated proteins defined as those fulfilling the sensitive criteria. Two HCMV time-courses were analysed alongside this, WCL1 and WCL2 from Weekes *et al.* (2014). The set-up of these experiments was as discussed in the section 1.4.2). For the HCMV data, the FC of proteins was calculated both compared to mock samples and also early time points of infection; this is because some IFN-induced proteins were initially upregulated, prior to being targeted by the virus. Proteins were considered to be downregulated by HCMV in this comparison if they had a FC > 2 in both time-courses, or were downregulated in one time-course and not quantified in the other.

The 7289 proteins that were quantified in at least one of the three VACV time-courses and one of the two HCMV time-courses were examined. Of these, 85 proteins were commonly downregulated (Figure 4.4A). DAVID enrichment analysis on these 85 proteins, against a background of the 7289 proteins quantified in both studies, identified similar classes of proteins to those seen previously, again showing enrichment for cell surface proteins, as well as collagens and innate immunity (Figure 4.5B).

The term proteoglycan highlighted the downregulation of multiple syndecans (SDC1, SDC2 and SDC4) by both viruses. Interestingly, syndecans have varied roles in viral infection. HSV-1 infection increases expression of SDC1 and 2, and knockdown of these reduces viral entry and plaque formation, suggesting they may be beneficial to the virus (Bacsa *et al*, 2011). SDC1 is a receptor for HCV attachment (Shi *et al*, 2013). VACV and HCMV appear to negatively regulate these proteins, so they may have alternative roles in infection with these viruses.

The previous proteomic study of HCMV infection identified protocadherins as being downregulated by the virus and provided initial evidence that some members of this family are activating NK cell ligands. Additionally, all of the 11 collagens downregulated by VACV were also downregulated by HCMV. Certain collagens may be ligands for the

inhibitory leukocyte-associated Ig-like receptor-1 (LAIR-1), with ligand binding leading to inhibition of immune cell activation (Lebbink *et al*, 2006); collagen downregulation may therefore represent a cellular response to infection.

Only 10 proteins were commonly upregulated by both viruses, and no terms were significantly enriched using DAVID analysis. Interestingly however, one of these proteins was APOE which is known to have roles in HSV, HIV and HCV infection, potentially assisting cell entry (Tudorache *et al*, 2017).

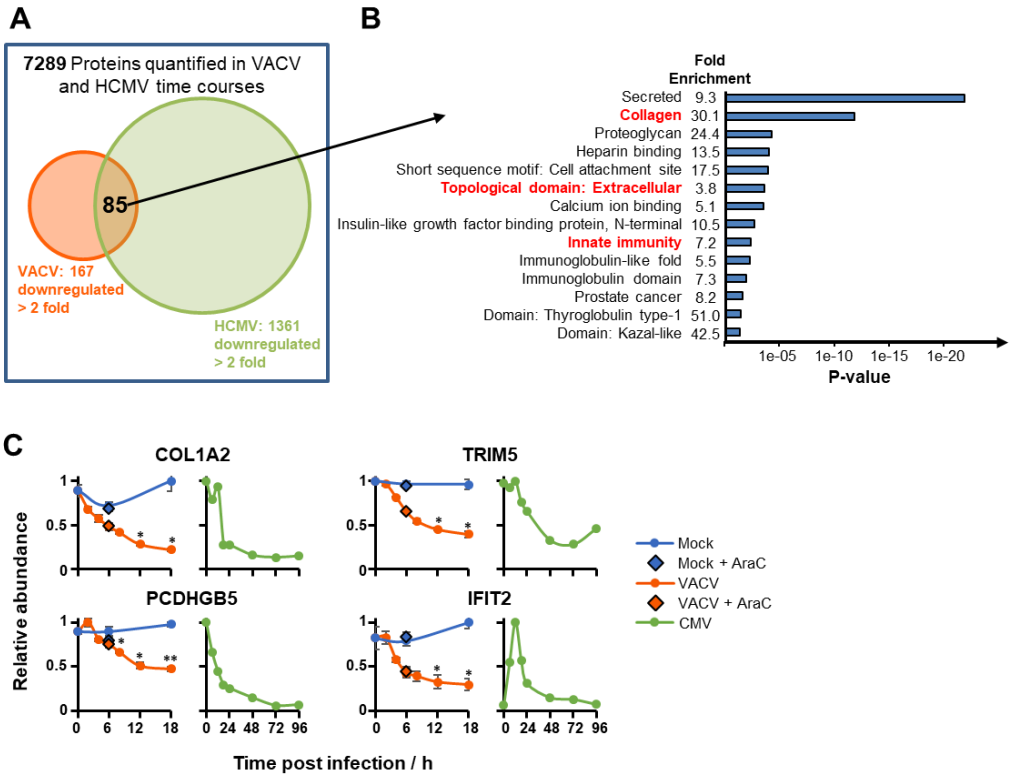


Figure 4.5 Co-regulation of proteins by VACV and HCMV

(A) Overlap between VACV and HCMV WCL proteomic data (Weekes *et al*, 2014). Downregulation of proteins in the VACV dataset used the sensitive criteria, and downregulation of proteins in the HCMV time-course required the protein to have a FC > 2 compared to mock or an early infection time point in both time-courses, or in one with the protein not quantified in the other.

(B) DAVID enrichment analysis of the 85 commonly downregulated proteins against a background of the 7,289 proteins quantified in both studies. Displayed are representative terms from each cluster and the associated BH-corrected p-values.

(C) Example profiles of proteins from enriched clusters. For the VACV data, plots are represented as mean ± SEM (n=3); *p<0.05, **p<0.01. For HCMV, plots are from the WCL2 experiment.

4.4.3 Downregulation of Innate Immune Proteins

DAVID enrichment analysis highlighted that both VACV and HCMV target multiple molecules in the category 'innate immunity' (Figure 4.4, Figure 4.5), in particular IFITs, IFITMs and TRIM proteins. Examination of these may identify additional mechanisms of immune evasion and assist identification of DNA virus restriction factors.

Strikingly, VACV rapidly downregulated all four canonical IFITs 1, 2, 3 and 5 (Figure 4.4B). The IFITs are induced by IFN, and have homologues in multiple vertebrate species; they have previously been implicated in restricting diverse RNA viruses including Rift Valley fever virus, vesicular stomatitis virus, and influenza A (Vladimer *et al*, 2014). IFITs restrict viral infection by inhibiting translation via binding of eukaryotic initiation factor 3 (eIF3), by recognition of RNA lacking 2' O methylation of the 5' guanosine cap, and of uncapped RNA with a 5' triphosphate. Direct interaction between IFITs and viral proteins may also be a mechanism of restriction; IFIT1 has been shown to directly inhibit DNA replication of HPV by binding and sequestering the viral protein E1 (Daffis *et al*, 2010; Pichlmair *et al*, 2011; Terenzi *et al*, 2008; Diamond & Farzan, 2013). Targeting of IFITs by both VACV and HCMV suggests that these proteins have additional mechanisms for restricting DNA viruses. Subsequent to this research, the VACV protein C9 was demonstrated to be responsible for targeting multiple IFITs (Liu *et al*, 2019).

Additionally, IFITM3 is downregulated by VACV in this study. It has also recently been reported to be downregulated by VACV to avoid antiviral restriction (Li *et al*, 2018).

Of the 29 TRIM proteins quantified, TRIM 5, 13, 25, 26 and 56 were downregulated during VACV infection. TRIM5 was also downregulated by HCMV (Figure 4.5C), and is known to restrict retroviruses. TRIM56 inhibits multiple RNA viruses (Liu *et al*, 2014; Rahm & Telenti, 2012). It also induces mono-ubiquitination of the cytosolic DNA sensor cGAS, and mice deficient in TRIM56 exhibited impaired production of type I IFNs and increased susceptibility to lethal infection by HSV-1 (Seo *et al*, 2018). Downregulation of TRIM56 by VACV may therefore represent a mechanism for immune evasion.

4.4.4 Modulation of Candidate Immunoreceptors

Enrichment analysis highlighted the downregulation of multiple collagens and protocadherins by VACV, with these molecules potentially having roles as NK cell ligands (Figure 4.4, Figure 4.5). Poxviruses have many mechanisms for manipulating host immunity, and regulates multiple ligands involved in modulating NK or T cell recognition (Burshtyn, 2013). Analysis of this data revealed downregulation of multiple known NK and T cell ligands, such as HLA-A, B and C and TNF receptor superfamily member 10B (TNFRSF10B). Nectin2, the ligand for activating NK receptor DNAX accessory molecule-1 (DNAM-1) was also downregulated, as were UL16 binding protein 2 (ULBP2) and MHC class I polypeptide-related sequence A (MICA), ligands for the NK receptor, NKG2D (Vivier *et al*, 2008; Smith *et al*, 2013b) (Figure 4.6A).

As many NK and T cell ligands belong to a small number of classes of proteins, candidate ligands were identified by considering proteins with associated InterPro functional domain annotations relating to immunoglobulins, C-type lectins, cadherins, TNF receptors, collagen, butyrophylin and MHC related molecules (Vivier *et al*, 2008; Hunter *et al*, 2012). This approach was previously used in investigations of HCMV infection (Weekes *et al*, 2014). Proteins identified which had these functional domains and were also modulated by VACV, are therefore potentially important. This identified 37 proteins which were both 2 fold downregulated and had a relevant InterPro annotation (Table 4.3). All the collagens and cadherins quantified are displayed in Figure 4.6B and C, a number of these were downregulated > 2 fold meeting the previously described sensitive criteria, and many were downregulated to some lesser extent. Other downregulated proteins included the tyrosine protein kinase receptor AXL, endosialin (CD248) and other molecules involved in adhesion, signalling and immunity (Figure 4.4, Figure 4.6).

Further analysis of these candidate immunoreceptors and their possible functions was beyond the scope of the current project, however the Smith Lab are investigating this area. An additional collaborative project with them is now focussing on changes at the PM of VACV infected cells. This should be particularly interesting given the enrichment of cell surface related terms amongst proteins modulated by VACV, and this may shed light on additional candidate immunoreceptors and viral targets.

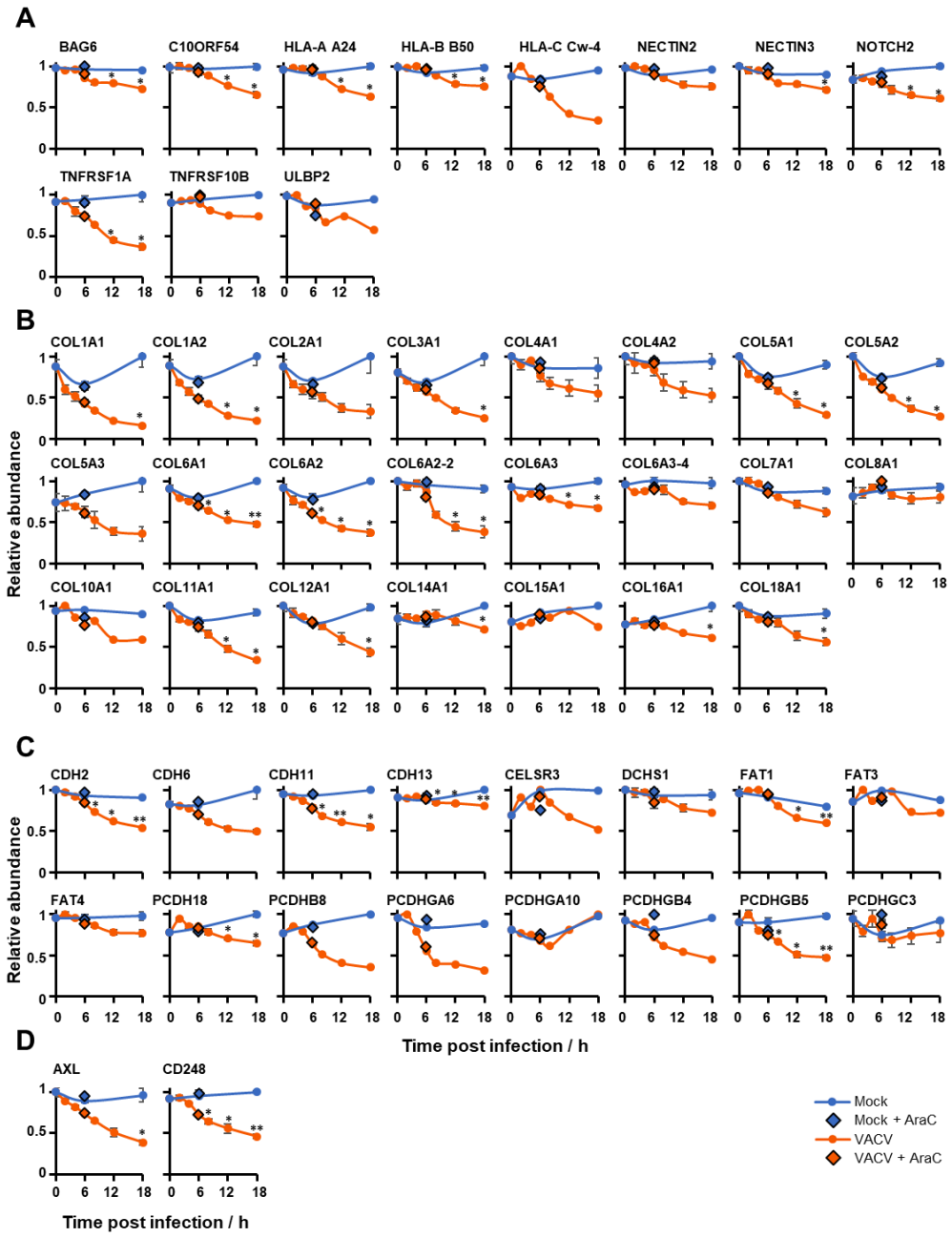


Figure 4.6 Modulation of known and putative immune ligands by VACV infection

Data are represented as mean ± SEM; *p<0.05, **p<0.01. Error bars and statistics are not included on the plots for HLA-C, TNFRSF10B, ULBP2, COL10A1, COL15A1, CELSR3, FAT3, PCDHB8, PCDHGA6, PCDHGA10, PCDHGB4 as these proteins were not quantified in all three replicates.

(A) Modulation of known NK or T cell ligands

(B) All collagens quantified

(C) All protocadherins quantified

(D) Downregulation of AXL and CD248

Table 4.3 Candidate immunoreceptors modulated by VACV infection

Candidate immunoreceptors identified according to Interpro domains (Ig denotes the term immunoglobulin), and viral downregulation ('Max FC' gives maximal downregulation).

InterPro Term	Gene name	Description	Max FC
Ig	ADAMTSL3	ADAMTS-like protein 3	3.09
MHC, Ig	ALPK2	Alpha-protein kinase 2	2.05
Ig	AXL	Tyrosine-protein kinase receptor UFO	2.48
Collagen	CCBE1	Collagen and calcium-binding EGF domain-containing protein 1	2.14
C-type lectin	CD248	Endosialin	2.18
Cadherin	CLSTN1	Calsyntenin-1	2.48
Collagen	COL11A1	Collagen alpha-1(XI) chain	2.70
Collagen	COL12A1	Collagen alpha-1(XII) chain	2.22
Collagen	COL1A1	Collagen alpha-1(I) chain	5.90
Collagen	COL1A2	Collagen alpha-2(I) chain	4.32
Collagen	COL2A1	Collagen alpha-1(II) chain	2.45
Collagen	COL3A1	Collagen alpha-1(III) chain	3.84
Collagen	COL5A1	Collagen alpha-1(V) chain	3.04
Collagen	COL5A2	Collagen alpha-2(V) chain	3.35
Collagen	COL5A3	Collagen alpha-3(V) chain	2.57
Collagen	COL6A1	Collagen alpha-1(VI) chain	2.11
Collagen	COL6A2	Collagen alpha-2(VI) chain	2.68
Ig	EBF1	Transcription factor COE1	2.34
Ig	F13A1	Coagulation factor XIII A chain	2.16
Ig	FGFR1	Isoform 21 of Fibroblast growth factor receptor 1	2.34
Ig	FLT1	Isoform 2 of Vascular endothelial growth factor receptor 1	2.22
C-type lectin, Ig	HAPLN3	Hyaluronan and proteoglycan link protein 3	2.35
MHC, Ig	HLA-C	HLA class I histocompatibility antigen, Cw-4 alpha chain	2.78
MHC, Ig	HLA-C	HLA class I histocompatibility antigen, Cw-12 alpha chain	2.22
Ig	IGFBP7	Insulin-like growth factor-binding protein 7	2.78
Ig	MXRA8	Matrix-remodeling-associated protein 8	3.29
Ig	NEO1	Neogenin	2.07
Cadherin	PCDHB8	Protocadherin beta-8	2.83
Cadherin	PCDHGA6	Protocadherin gamma-A6	2.78
Cadherin	PCDHGB4	Protocadherin gamma-B4	2.11
Cadherin	PCDHGB5	Protocadherin gamma-B5	2.06
Ig	PDGFRA	Platelet-derived growth factor receptor alpha	2.16
Ig	PDGFRL	Platelet-derived growth factor receptor-like protein	2.42
C-type lectin	SUSD5	Sushi domain-containing protein 5	2.37
TNF	TNFRSF1A	Tumor necrosis factor receptor superfamily member 1A	2.61
Butyrophylin	TRIM5	Tripartite motif-containing protein 5	2.39
Ig	TXNIP	Thioredoxin-interacting protein	2.42

4.5 Temporal Analysis of VACV Viral Protein Expression

As well as quantifying nearly 9000 human proteins, the majority of the VACV proteins were also quantified over the whole course of infection, enabling a detailed temporal investigation of viral protein expression. Previous classification of viral genes into temporal classes was predominantly based on RNA-sequencing (RNA-seq) analysis and very limited proteomic data. This investigation therefore complemented the previous studies, and enabled correlation between expression profiles of viral and cellular proteins, giving insights into host-virus interactions.

4.5.1 There were Four Temporal Classes of Viral Protein Expression

Around 80% of all predicted VACV proteins were quantified in this study, and the number of distinct classes determined using the k-means method. The approach was trialled using 1-15 classes. For each number of classes, all proteins were assigned to a class, and a 'cluster centroid' was calculated. The distances of every protein from its cluster centroid was summed, and this was plotted against the number of classes. This summed distance necessarily decreases as the number of classes increases, but this difference reduces with each added group. When there are too many classes, the summed distances become minimal and simply represents overfitting of the data. The point of inflection therefore indicates up to how many classes for which there is structure in the underlying data, estimating the true number of classes. Using this approach, the point of inflection in this dataset fell between four and six classes, suggesting there were at least four distinct temporal classes (Figure 4.7A). Plotting of the class centroid profiles displays the expression profile of the average protein in that class, from early Tp1 proteins (yellow) through to late Tp4 proteins (blue, Figure 4.7B).

In order to provide additional confidence to the classification, viral expression at 6 hpi was examined in the presence of AraC. Late viral genes were initially termed 'post replicative' (PR) in an early RNA-seq based classification system, due to the requirement of viral DNA replication for their expression (Yang *et al*, 2010). As AraC is an inhibitor of viral DNA replication, this would be expected to limit expression of the later classes, without impacting proteins expressed early during infection. The FC between protein

expression during VACV infection with or without AraC was calculated and examined in the context of each class (Figure 4.7C). As expected, none of the Tp1 proteins were inhibited by AraC. Conversely, more than 70 % of the latest class of proteins, Tp4, showed at least 1.5 fold downregulation with AraC. However, a subset of proteins in Tp4 were not inhibited by AraC. This may be due to the somewhat arbitrary 1.5 fold cut off employed, or by the time point chosen. Despite including this sample at a later time point than the 4 hpi used in the initial optimisation experiment, it may still be too early to see maximal effect; many of the late proteins were not yet highly expressed at 6 hpi, and a greater window for downregulation with AraC may exist at a later time point.

Viral protein signals were normalised to the maximum expression across all time points, so expression was on a scale of 0 to 1. Each class of proteins was then subjected to hierarchical clustering for visualisation on a heatmap. A clear pattern of expression can be seen from the early to late classes (Figure 4.7D). Examples of proteins from each class are shown underneath (Figure 4.7E). No p-values are given on these profiles for viral proteins, as any signal observed in the mock samples was only at the level of noise, so comparing to the 18 h mock (as with the human proteins) would not give any meaningful information. A list of all viral proteins detected and their assigned classes is given in Appendix Table 4.

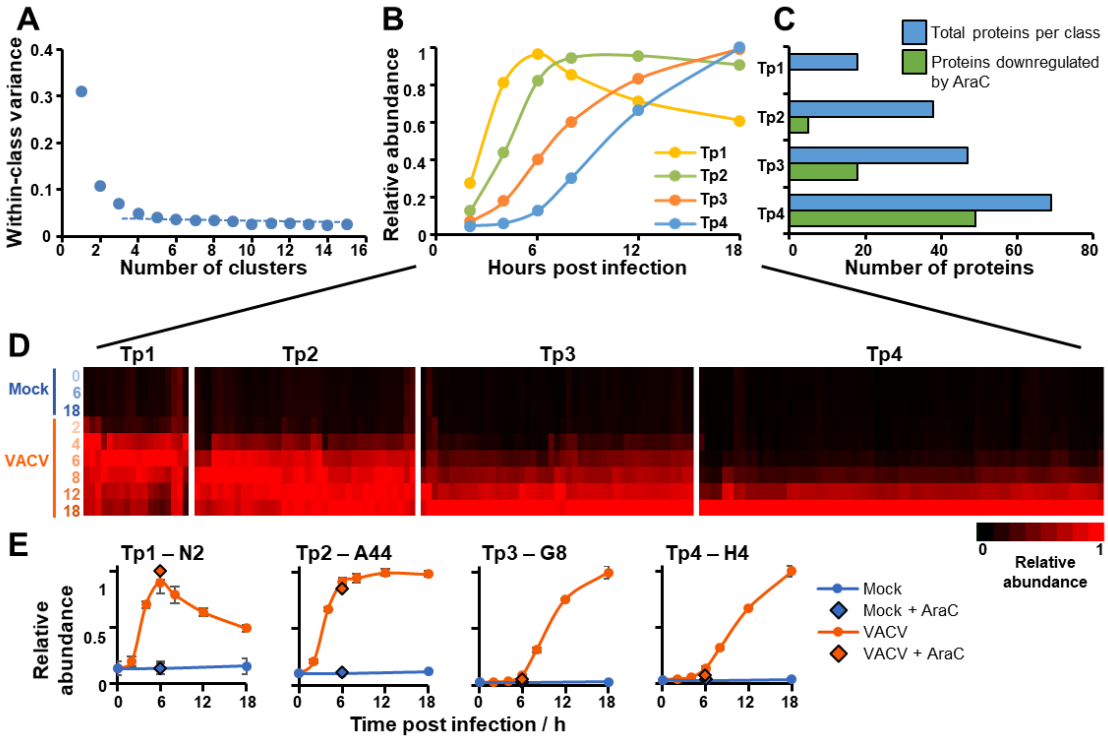


Figure 4.7 Definition of temporal classes of VACV gene expression

(A) The number of k-means classes is plotted against the summed distance of each protein from the class centroid, in order to determine the true number of classes.

(B) Expression profiles of the class centroids.

(C) Number of proteins per temporal class, and the number of proteins in each class whose expression was reduced > 1.5 fold at 6 hpi by the addition of AraC.

(D) Hierarchical clustering of all proteins in each temporal class, as assigned by k-means. Expression was normalised to the maximum across the measured time points.

(E) Expression profiles of example proteins, representative of each class. Data shows the mean ± SEM (n=3).

4.5.2 Comparison of viral protein classes to previous studies

Previous transcriptional studies have utilised microarrays (Assarsson *et al*, 2008), RNA-seq (Yang *et al*, 2010, 2011) and a combination of RNA-seq with ribosome profiling (Yang *et al*, 2015). There has additionally been a more limited proteomic investigation (Croft *et al*, 2015).

RNA Sequencing Based Classifications

RNA-seq studies were conducted on infected HeLa cells, harvested at 0 through to 4 hpi. This study defined three classes of viral proteins, 'early', separated into E1.1 and E1.2, and PR (Yang *et al*, 2010). An extension of this study further dissected the PR category by utilising a virus with inducible expression of the late transcription factor G8 (Yang *et al*, 2011). This yielded a classification systems of the four transcriptional classes: E1.1, E1.2, intermediate (I) and late (L). Further RNA-seq data was included in the most recent ribosome profiling investigation, with samples up to 8 hpi (Yang *et al*, 2015).

Comparison of protein level data to these RNA-seq based classifications showed a high level of similarity. The RNA-seq data from the ribosome profiling publication provided expression data at 2, 4 and 8 hpi. This was normalised to the maximum in the same way as the proteomics data so it could be included in the hierarchical clustering of each class. The heat map therefore shows the relative expression at each time point for a given protein in the RNA-seq data alongside its relative expression at each time point in the proteomic screen (Figure 4.8A). Visual comparison suggests good concordance between the proteomic data and the RNA-seq.

More extensive comparison identified that 18/18 Tp1 proteins were either E1.1 or E1.2 transcripts, Tp2 proteins were also predominantly early transcripts, and 60/69 Tp4 proteins were I or L transcripts ($p < 0.0001$, Fisher's exact test; Figure 4.8B, Appendix Table 4). This high level of correspondence between studies using different approaches suggests the temporal classes of VACV protein expression defined here were likely to be biologically relevant. Additionally, as most class assignments were the same from proteomic and transcriptomic data, this suggests that regulation of expression was likely to be chiefly exerted at the level of transcription.

However, interesting differences were also highlighted. Some early transcripts had their maximal expression at the protein level later during infection, and were defined here as Tp3 proteins (Figure 4.8A). These included E4 and G5.5, which are subunits of the DNA-dependent RNA polymerase, as well as G2, which has roles PR transcription elongation. There may therefore be additional mechanisms of temporal regulation of viral gene expression, in addition to that observed at the transcriptional level.

Proteomic Studies Based Classifications

A previous proteomic study examined infection of a murine bone marrow-derived dendritic-like cell line DC2.4 (Croft et al, 2015). Two time-courses were conducted, the first from 0.5 – 9.5 hpi at 3 h intervals, and the second from 0.5 to 8.5 hpi at 2 h intervals. In total, 101 viral proteins were quantified, and 4 temporal classes defined. There were 47 proteins which showed concordant classification between the 2 time-courses, with the others being mismatched or only quantified in 1. Comparison of the proteomic data presented here to these 47 proteins showed very similar classification, with 4/5 Tp1 proteins assigned to temporal class 1 of the Croft et al data, and 19/20 Tp4 proteins classified as temporal class 3 or 4 (Figure 4.8C, Appendix Table 4).

Viral Protein Class and Function

A final comparison was between the assigned classes of viral proteins, and their functions. The functions were predominantly as summarised by Yang *et al*. (2010), with some updated functions added from other sources (literature reviewed and functions added by Dr Jonas Albarnaz, described in Appendix Table 4). The majority of early expressed proteins were involved in 'host interaction', encompassing vaccinia growth factor and immune evasion proteins. 'DNA replication' related largely to Tp2 proteins, including enzymes involved in nucleotide precursor synthesis, DNA replication and DNA processing. The majority of Tp4 proteins were assigned to 'virion association', including virion structural proteins or those involved in virion morphogenesis (Figure 4.8D).

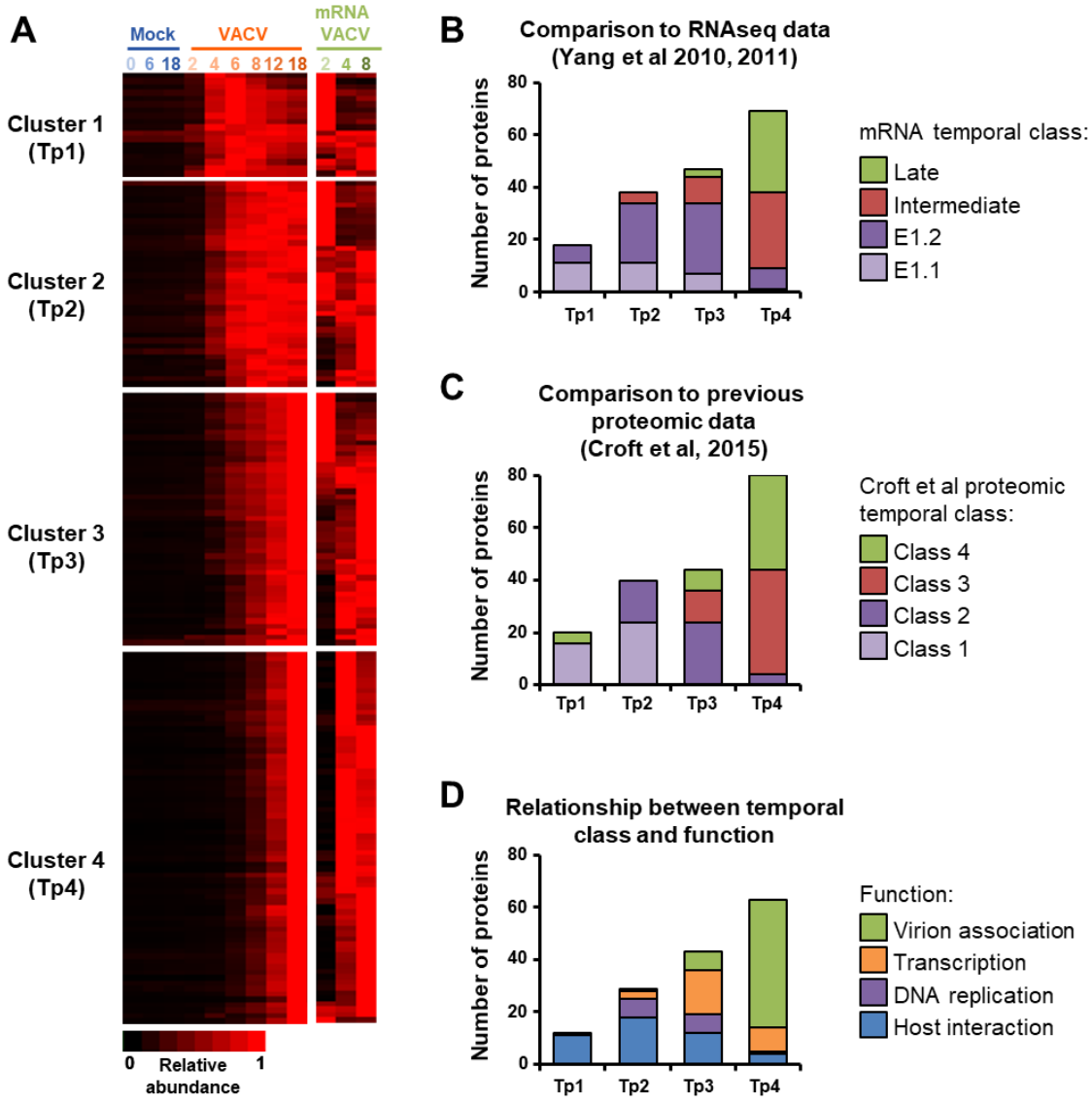


Figure 4.8 Comparison of viral protein temporal classes to previous studies

(A) Comparison between temporal expression profiles of each protein from this study, and the equivalent expression profile in a transcriptional analysis (Yang *et al*, 2015). Proteins were grouped according to assigned classes from this data, and expression was normalised to the maximum across the measured time points. Each class of proteins was subjected to hierarchical clustering. Full details of comparisons in this figure can be found in Appendix Table 4.

(B) Comparison of viral protein classes (this study) and transcriptional classes (Yang *et al*, 2010, 2011).

(C) Comparison of viral protein classes from this study, with those assigned in a previous proteomic study by (Croft *et al*, 2015). As there were two datasets from the previous investigation, only the 47 viral proteins with concordant assignments in both experiments were included in the comparison.

(D) Functional analysis of viral proteins in each class, based on information from (Yang *et al*, 2010) and other sources, as detailed in Appendix Table 4.

4.6 Systematic Analysis of Protein Degradation During VACV Infection

4.6.1 Experimental Outline

In order to examine the mechanism by which VACV led to downregulation of multiple classes of proteins, a multiplexed proteomic approach used the proteasome inhibitor, MG132, to investigate which proteins were subject to active degradation. Cells were infected at an MOI of 5 in biological triplicate, for 12 h. Infection was allowed to proceed for 2 h, as required to enable virus uncoating in the host cells, prior to the addition of MG132, or mock treatment with DMSO. Uninfected samples were also either mock treated or treated with MG132; in this case just a single sample of each was harvested (Figure 4.9). Infections and harvesting were performed by Dr Jonas Albarnaz.

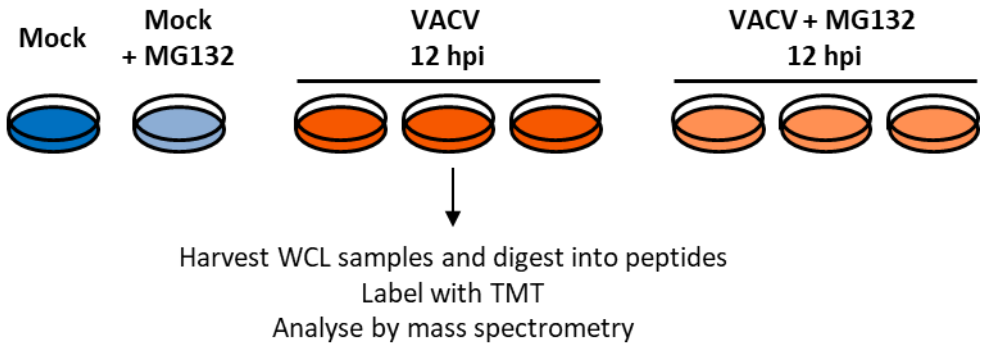


Figure 4.9 Schematic of MG132 experiment examining VACV induced protein degradation

Cells were infected in the presence or absence of MG132 (10 μM added 2 h after infection), and harvested at 12 hpi for WCL proteomic analysis (in triplicate). Mock treated cells were prepared similarly (single replicate).

4.6.2 Proteasomal Degradation of Human Proteins

In this experiment, 8263 human proteins were quantified. In each case, a BH-corrected p-value was estimated using a two-tailed t-test comparing the triplicate infected samples with or without MG132. However, it was important to separate proteins which were proteasomally degraded from those that have a high baseline turnover and showed altered abundance during MG132 treatment of uninfected cells. An MG132 'rescue ratio' was therefore calculated by comparing protein abundance during VACV infection ± MG132 (average of the three replicates) with the protein abundance during mock infection ± MG132 (Figure 4.10B). A minimum value of one was used for the change in abundance of mock ± MG132 ('a' in the rescue ratio), in order to avoid artificial inflation of the rescue ratio for proteins where any decline was observed in the abundance.

Of the proteins downregulated >2 fold in this experiment, 69 % had a rescue ratio > 1.5, $p < 0.05$ (Figure 4.10A). This suggests that proteasomal degradation was one of the predominant mechanisms employed by VACV to modulate expression of host proteins. All of the IFITs, TRIMs and Ephrin receptors which were downregulated > 2 fold were rescued by MG132 (Figure 4.10B). In contrast, some proteins were downregulated > 2 fold, but not rescued by MG132; whilst isoforms 1 and 2 of collagen alpha-2(VI) chain (COL6A2) were rescued, a further six downregulated collagens did not meet the criteria. This suggests alternative mechanisms of downregulation were also acting. The complete dataset of proteins quantified in this experiment, and their rescue ratios, is included in supplemental table S6 of the manuscript (Soday *et al*, 2019).

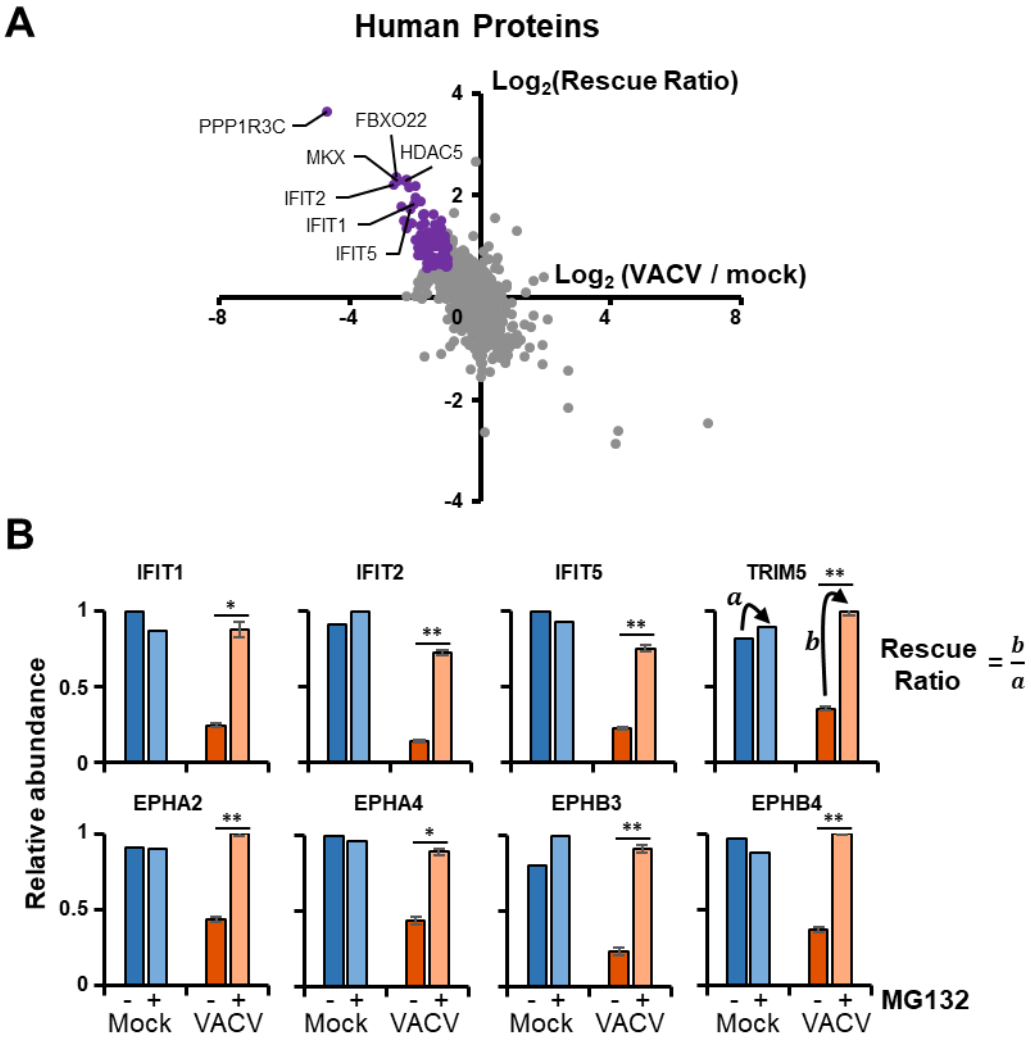


Figure 4.10 Systematic analysis of proteasomal degradation of host proteins

(A) Scatter Plot of all human proteins quantified in MG132 proteomic investigation. Rescue ratio was calculated as: (protein abundance during VACV infection with MG132 / abundance during infection without MG132) (b) / (protein abundance during mock infection with MG132 / abundance without MG132) (a). The denominator (a) was set a minimum of 1. VACV infected samples (± MG132) were examined in triplicate, whilst mock samples (± MG132) were in single replicates, in order to fit the constraints of the 11-plex experiment. P-values were estimated for each protein using a BH-corrected two-tailed t-test, comparing the triplicate infected samples with and without MG132. Proteins highlighted in purple were downregulated > 2 fold, have a rescue ratio > 1.5 and $p < 0.05$.

(B) Examples of proteins rescued by MG132. Infected samples are displayed as the mean ± SEM. P-values were calculated as described in (A); *$p < 0.01$, **$p < 0.0005$.

4.6.3 Proteasomal Degradation of Viral Proteins

MG132 inhibited expression of late but not early viral proteins, consistent with previous reports (Satheshkumar *et al*, 2009; Teale *et al*, 2009). In this case, the rescue ratio does not take into account changes in mock infection, as any quantification of viral proteins here is only at the level of noise. The viral rescue ratio was therefore calculated as mean expression (across the three replicates) during VACV infection in the presence or absence of MG132. Distinct patterns of MG132 response were observed between the different classes of viral proteins (Figure 4.11).

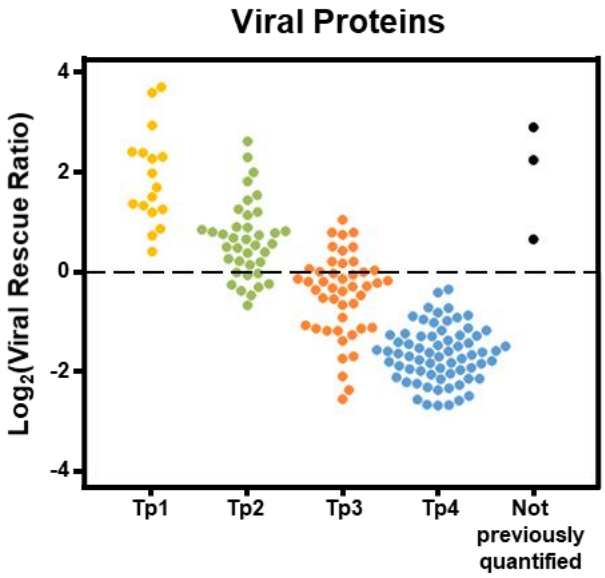


Figure 4.11 Effect of MG132 treatment on viral protein expression

Effects of proteasome inhibition by MG132 on expression of viral proteins is displayed, with proteins separated according to their assigned temporal classes. The viral rescue ratio was determined by the relative abundance during VACV infection with or without MG132, using the mean expression values from the three replicates. The three proteins indicated in black were not quantified in the proteomic time-courses and therefore did not have an assigned class.

Most Tp1 and some Tp2 proteins were upregulated by MG132, whilst Tp3 proteins were largely unaffected and Tp4 proteins were inhibited (Figure 4.11). The increase in Tp1 and Tp2 proteins may be a consequence of the inhibition of viral DNA replication and expression of PR genes, leading to accumulation of early viral mRNAs and their products. It may also be partially explained by enrichment for these classes of proteins in 'host

interaction' functions (Figure 4.8D); if the protein is usually co-degraded with a host factor, it will accumulate in the presence of MG132.

Three viral proteins were detected in the MG132 experiment that were not quantified in the time-courses, and therefore did not have a class assigned. They were A57, VACWR014 and VACWR012/VACWR207, and they had a viral rescue ratio of 6.9, 3.8 and 1.4 respectively. This suggests the first two were likely to be Tp1 or Tp2 early proteins, whilst the latter was most likely to be Tp2 or Tp3.

4.7 Identification of Candidate Virus-Host Interactions: HDAC5

HDAC5 was rapidly degraded during infection, with both VACV (Figure 4.12A) and also HCMV (Weekes *et al*, 2014) (Figure 4.12B). In addition, it was one of the most substantially rescued proteins by MG132 (Figure 4.12C). This suggests it may be important during infection with multiple viruses.

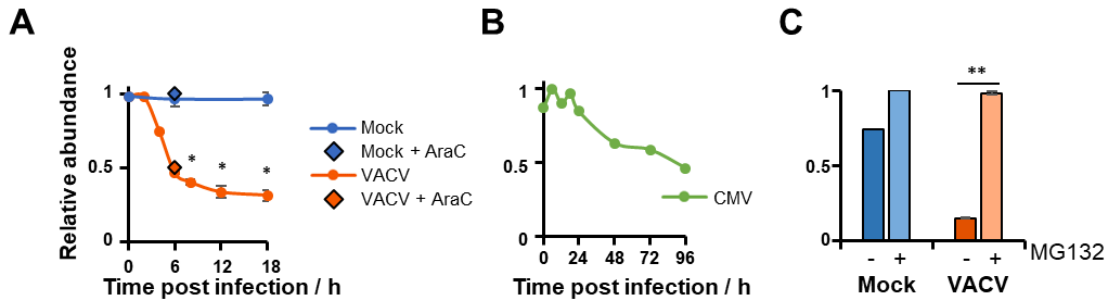


Figure 4.12 HDAC5 expression during VACV infection

(A) HDAC was downregulated during infection with VACV compared to mock. Data are represented as mean ± SEM (n=3); *p < 0.05.

(B) HDAC5 was downregulated over 96 h of infection with HCMV (Weekes *et al*, 2014).

(C) HDAC5 was rescued by MG132. Mock samples were single replicates, and infected samples were harvested in triplicate. Infected samples are displayed as the mean ± SEM. P-values were estimated using a BH-corrected two-tailed t-test; **p < 0.0005.

4.7.1 HDAC5 was Targeted by VACV C6 Protein

With more than 200 VACV proteins, identifying which viral factor is responsible for targeting a host protein can be challenging. HDAC5 degradation was observed in the presence of AraC, suggesting the protein was likely to be from an early class. An initial investigation was performed by considering the profile of HDAC5 degradation. Matching the inverse profile of HDAC5 to that of the central profiles of each temporal class indicated it was likely to be one of the 38 proteins from the Tp2 class (Figure 4.13A).

More comprehensive investigation was used to match the profile of HDAC5 to Tp2 viral proteins known to have roles in innate immune signalling. Along with additional data from the Weekes and Smith laboratories, investigating multiple mutant viruses (unpublished), C6 appeared to be a good candidate with matching kinetics (Figure 4.13B).

In order to examine this further, an unbiased proteomic comparison of infection with WT VACV compared to a deletion mutant lacking the *C6L* gene which encodes the viral protein C6 (Unterholzner *et al*, 2011), vΔC6, was performed. Cells were infected with WT and mutant virus in triplicate (by Dr Jonas Albarnaz), in order to assess which proteins were rescued by deletion of C6 and therefore were targeted by C6 for degradation. This identified HDAC5 as the major target (Figure 4.13C). Cerebellar degeneration-related protein 2 (CDR2) and the E3 ubiquitin ligase RING-finger protein 114 (RNF114) were also rescued by deletion of C6, though to a lesser extent. Downregulation of HDAC5 during infection, and rescue by deletion of C6, was readily validated by immunoblot; additionally, there was no effect of infection on the control protein HDAC1, confirming the specific regulation of HDAC5 by C6 (Figure 4.13D, immunoblot by Dr Yongxu Lu, Smith Lab, Department of Pathology, University of Cambridge). Deletion of C6 prevented the degradation of HDAC5, therefore C6 was necessary for its downregulation. Additional experiments by Dr Yongxu Lu were able to demonstrate sufficiency of C6 for downregulation of HDAC5, using inducible expression of C6 and immunoblot analysis. He was also able to validate that degradation was proteasome dependent, with an immunoblot confirming rescue by of HDAC5 by the addition of MG132 (data shown in Soday *et al*. 2019).

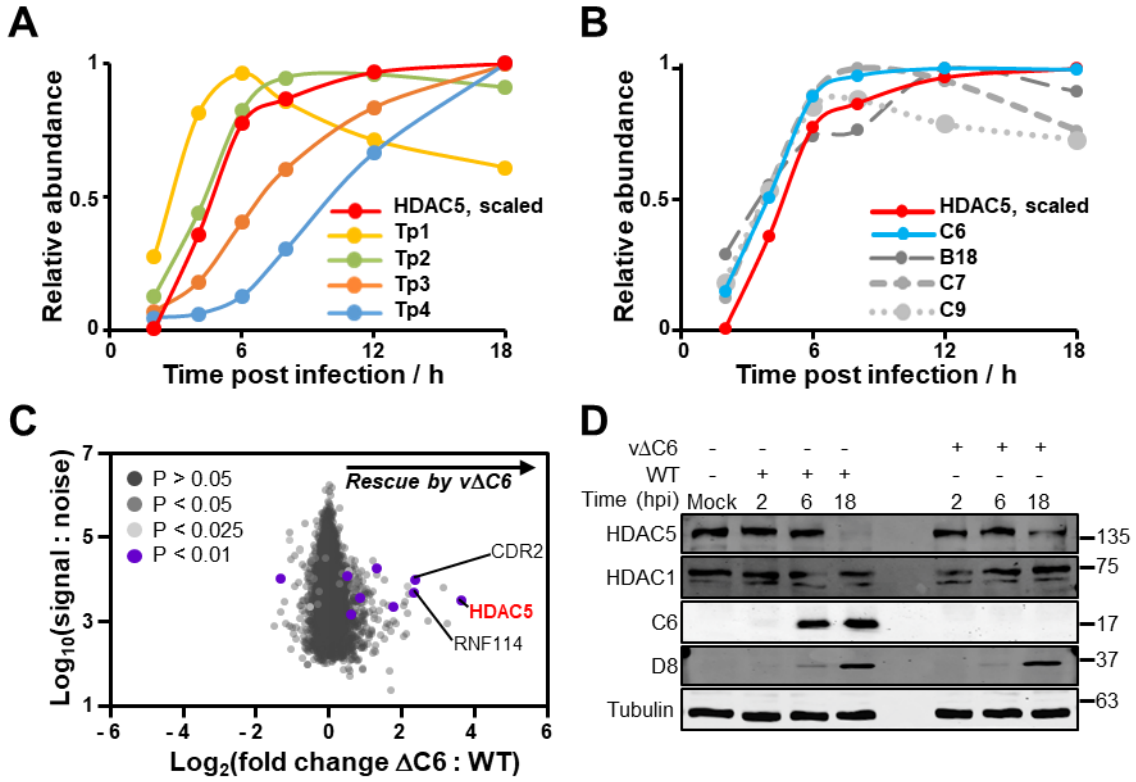


Figure 4.13 HDAC5 was targeted by the VACV protein C6

(A) Expression profiles of viral protein class centroids, alongside an inverted profile of HDAC5 scaled from 0 to 1.

(B) Inverted and scaled profile of HDAC5 (as in (A)), alongside a selection of Tp2 proteins with known roles in the regulation of IFN or ISGs.

(C) HFFF-TERTs were infected in triplicate with WT VACV or a deletion mutant lacking the gene *C6L* (vΔC6), at an MOI of 5 for 12 h (Dr Jonas Albarnaz). The scatter plot shows all quantified proteins, with BH corrected p-values estimated using a two-tailed t-test.

(D) Representative immunoblot (n=3) demonstrating that C6 specifically targeted HDAC5, and had no effect on HDAC1. Infection was at an MOI of 5. Immunoblot by Dr Yongxu Lu.

4.7.2 HDAC5 Restricts VACV and HSV-1 Infection

As HDAC5 was downregulated rapidly and early during infection by multiple large DNA viruses, it may have a role as an ARF. In order to test this, Dr Yongxu Lu (Smith lab) generated various cell lines either overexpressing HDAC5 or with a HDAC5 knockout, and examined the effects on replication of VACV and HSV-1.

He generated U2OS cell lines with inducible overexpression of HDAC5 (Figure 4.14B). The overexpression and control cells were infected with A5-GFP VACV at low MOI, and the supernatant harvested at two days post infection and titrated. Overexpression of HDAC5 led to reduced viral titre. The same experiment was performed using VP26-GFP HSV-1, in this case harvesting supernatants for titration at three days post infection, demonstrating HDAC5 restriction of HSV-1 replication (Figure 4.14A). A similar experiment was performed on HDAC5 knockout populations in HeLa cells. Knockouts were produced using the CRISPR/Cas9 system and single cell cloned to generate monoclonal populations. CRISPR guides targeted exon 4 (HDA5 knockout (H5KO)-1) and exon 3 (H5KO2) of HDAC5 (Figure 4.14D). Knockout of HDAC5 led to increased viral titre for both VACV and HSV-1, suggesting it was restricting viral replication (Figure 4.14C). Reintroduction of HDAC5 into the CRISPR knockout cells restored this restriction, whereas transduction with an empty vector (EV) did not (Figure 4.14E). Two further monoclonal population of HDAC5 knockouts (H5KO3, H5KO4, both targeting exon 4) were generated in HEK-293Ts (Figure 4.14H), and these also enhanced replication of both viruses (Figure 4.14G).

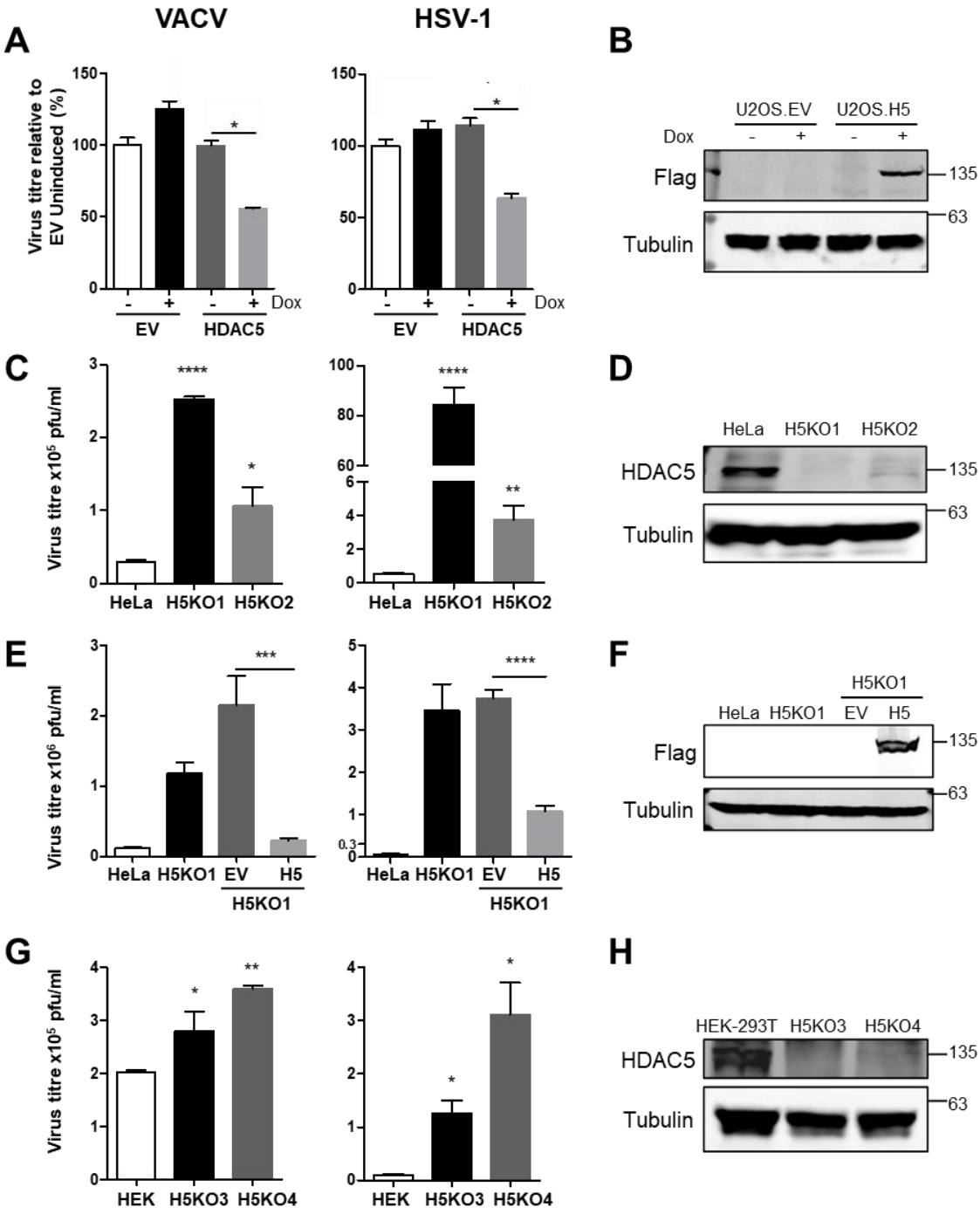


Figure 4.14 HDAC5 restricted infection with VACV and HSV-1

All infections were performed at an MOI of 0.001, for two days with VACV, and three days with HSV-1. Data are represented as mean ± SEM. P-values were estimated using a two-tailed t-test (n=3) where *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001. Experiments were performed and figures produced by Dr Yongxu Lu.

(legend on next page)

Figure 4.14 HDAC5 restricted infection with VACV and HSV-1

(A) Replication of VACV and HSV-1 was reduced in U2OS cells inducibly expressing HDAC5-FLAG. Induction was with 100 ng/ml doxycycline (Dox) overnight prior to infection.

(B) Immunoblot of HDAC5-FLAG expression in U2OS cells used in (A)

(C) Replication of VACV and HSV-1 was enhanced in HDAC5 CRISPR/Cas9 knockout HeLa cells, compared to parental cell lines.

(D) Immunoblot confirming knockout of HDAC5 in two HeLa monoclonal CRISPR populations, H5KO1 and H5KO2, as used in (C).

(E) Reintroduction of HDAC5 in the CRISPR knockout HeLa cells restored restriction, whilst transduction with an empty vector (EV) did not.

(F) Immunoblot confirming expression of HDAC5-Flag on reintroduction into H5KO1 knockout cell line, as used in (E).

(G) Replication of VACV and HSV-1 was enhanced in HDAC5 CRISPR/Cas9 knockout HEK-293T cells, compared to parental cell lines.

(H) Immunoblot confirming knockout of HDAC5 in two HEK-293T monoclonal CRISPR populations, H5KO3 and H5KO4, as used in (G).

4.8 Discussion

This study quantified ~9000 human proteins and ~80% of all VACV proteins over a single replication cycle. This has greatly extended the findings from previous research, which were limited by the technology available (Bartel *et al*, 2011; Chou *et al*, 2012). Here, the use of TMT-based MS3 analysis has provided a comprehensive and precise quantitative analysis of proteomic changes throughout infection, and therefore a valuable resource for researchers in the poxvirus and immunity fields.

4.8.1 Downregulation of Host Proteins

Global Host Shutoff

Identifying host proteins which were downregulated by viral infection was a key part of this investigation; VACV has numerous strategies for modifying the immune response in order to enable its cytoplasmic replication, and therefore targeting of a protein by VACV may suggest important roles in immunity. Global shutoff of host protein synthesis is a well documented phenomenon in VACV infection (Moss, 1968; Rice & Roberts, 1983). It is thought to reduce competition for the cellular machinery required for protein synthesis by host mRNAs and therefore direct resources to viral protein synthesis, as well as limiting expression of antiviral proteins. Shutoff is achieved by viral proteins D9 and D10 removing the 5' cap from host mRNAs leading to their degradation by 5'-3' exoribonuclease 1 (XRN1) (Parrish *et al*, 2007; Parrish & Moss, 2007), and through VACWR169 inhibiting translation initiation (Strnadova *et al*, 2015). Global shutoff was not readily observed in this proteomic data. This may be due to the experimental set up used here, which aimed to have an equal total amount of protein in each channel. This was achieved by using a BCA to measure the amount of protein from each sample, and then electronic normalisation during data processing. This normalisation may mask an overall reduction in the amount of host protein in a particular sample.

Specific Downregulation of Host Proteins

In this study, just 265 proteins were downregulated; this was less than 3 % of the human proteins quantified. Global shutoff is unlikely to account for changes only seen in such a small proportion of the proteins quantified, suggesting specific regulation of these proteins. Further evidence of this comes from the rescue of ~69 % of proteins downregulated more than 2 fold by the addition of MG132.

As MG132 is predominantly a proteasome inhibitor, rescue of many of these proteins will mean they were proteasomally degraded. However, MG132 also has less potent effects as an inhibitor of calpains and lysosomal cathepsins, so these pathways may also be involved in the protein degradation (Kisselev *et al*, 2012). A more specific proteasome inhibitor, such as bortezomib, could be used to confirm proteasomal degradation. Additionally, MG132 treatment has previously been shown to prevent formation of cytoplasmic viral factories, and expression of late viral genes (Teale *et al*, 2009); this is in concordance with the data presented here (Figure 4.11). Therefore, proteins rescued by MG132 may also represent targets of late viral proteins. It would be interesting to compare proteins which were rescued by MG132 with those rescued by AraC, an inhibitor of late viral gene expression, to delineate which host proteins may be targets of late viral proteins. Although this investigation did include AraC treated samples, these were harvested at 6 hpi, whilst the MG132 treated samples were harvested at 12 hpi. Not all the late viral proteins were highly expressed at 6 hpi, and most host proteins had not yet reached their maximal downregulation, therefore a later time point may be more suitable for this comparison.

Overlap of Proteins Downregulated by VACV and HCMV

It is interesting that so few proteins were specifically downregulated during VACV infection compared to during infection with HCMV. Both are large dsDNA viruses with a similar number of canonical genes. Many of the HCMV proteins are well characterised in targeting multiple host proteins. Expression of HCMV US2 in THP-1 cells identified a multitude of proteins being downregulated, such as MHC molecules and integrins. Additionally, infection with a US2 deletion mutant rescued 21 proteins compared to WT HCMV infection, including multiple integrins and immunoglobulin superfamily

members; this suggests it has many targets (Hsu *et al*, 2015). It is possible that each VACV protein has a more specific role and fewer targets, however examination of C6 here shows it targets at least three proteins. Another possible explanation is the comparatively complex viral lifecycle of herpesviruses. As they exhibit latent infection as well as lytic, and the life-cycle involves entry and exit of the nucleus, increased modulation of the infected cell may be required to achieve this.

HDAC5

Despite the vast difference in the number of modulated proteins, there were 85 proteins which were commonly downregulated by both VACV and HCMV. With different viruses possessing varied mechanisms for regulating host immunity, convergence of these on downregulation of the same protein provides additional confidence that they may be of interest. HDAC5, a class IIa histone deacetylase, was one such protein. This class of proteins are transcriptional repressors, however one of their best documented roles is binding to and repressing the myocyte enhancer factor-2 (MEF2) transcription factors. MEF2 has varied roles in development, differentiation and cell survival (Verdin *et al*, 2003; Lemercier *et al*, 2000). Although HDAC5 has not previously been reported to restrict viral replication, it has been observed to interact with viral proteins. The HSV-1 immediate protein ICP0 is able to interact with HDAC5, preventing its repression of MEF2. The function of this isn't defined, but it may have roles in allowing MEF2 to aid neuronal cell survival during viral reactivation (Lomonte *et al*, 2004). Additionally, HDACs may bind to the EBV protein EBNA2, preventing its transcriptional activation of the EBV oncoprotein latent membrane protein-1 (LMP1) (Portal *et al*, 2006). HDACs have also been investigated in the context of EBV infection as MEF2 has binding sites in the promoter for its immediate early gene trans-activator *BZLF1*, expression of which leads to reactivation of the virus. MEF2 binding is thought to recruit class IIa HDACs to the promoter, and maintain EBV in its latent state (Gruffat *et al*, 2002).

In this investigation, HDAC5 was found to restrict both VACV and HSV-1 replication. These viruses replicate in distinct cellular compartments; VACV replicates in viral factories in the cytoplasm, whereas HSV-1 replicates in the nucleus. The ability of HDAC5 to shuttle between these compartments may therefore be important for its function here. The Smith Lab is currently investigating potential roles of HDAC5 in

VACV infection, it may involve inhibition of viral promoters, or modulating innate immune signalling pathways. It would be interesting to determine whether this antiviral restriction is dependent on MEF2. Histone deacetylase 4 (HDAC4) is also targeted by VACV protein C6, and limits replication; this is thought to be due to roles in promoting type I IFN signalling, through interaction with STAT2 (Lu *et al*, 2019).

IFITs

As well as HDAC5, multiple immune mediators were found to be degraded by VACV infection. These included all canonical IFIT proteins, and some of the TRIM proteins. Following the publication of this data, the downregulation of IFIT proteins was attributed to the VACV protein C9 (Liu *et al*, 2019). The C9 protein contains ankyrin repeats, usually involved in protein-protein interactions, and an F-box which forms part of the SCF complex, resulting in ubiquitination and proteasomal degradation of associated proteins. Proteomic studies were used to investigate the effects of C9 expression in IFN-stimulated cells. The F-box was deleted in order to determine which proteins were associated with C9, by allowing them to accumulate rather than getting targeted for degradation. This identified IFITs (1, 2, 3 and 5) as the primary targets for interaction with and degradation by C9. This agrees well with the data presented here showing downregulation of the IFITs, and rescue with proteasome inhibition by MG132.

Liu *et al.* investigated the role of IFITs in VACV infection, and confirmed that IFN sensitivity of a C9 deletion mutant was dependent on expression of IFIT1 and 3. Furthermore they found that IFN treatment during infection with a C9 deletion mutant inhibited synthesis of intermediate and late proteins. An alternative VACV mutant with an inactivated ribose methyl-transferase is unable to convert the IFIT-sensitive 5’ mRNA cap to a form which is resistant to IFITs. This mutant is unable to synthesise early as well as intermediate and late proteins, due to recognition of mRNAs by IFITs. Expression of C9 and therefore degradation of IFITs rescued this mutant. However, expression of additional methyl transferase was not able to rescue viral protein synthesis in the C9 mutant, suggesting the IFITs have additional mechanisms, later in viral replication, beyond those involving the 5’ cap. They suggested that as the expression of intermediate and late genes is dependent on viral DNA replication the IFITs may be inhibiting VACV at this stage. This may be in a similar manner to the role of IFITs in restricting HPV

infection by binding the E1 protein, but no mechanism was identified (Liu *et al*, 2019). It could be interesting to overlap the list of downregulated proteins from this study with the proteomic investigations of C9 to see if any additional targets are identified.

Collagens

Although more than two thirds of the downregulated proteins were rescued by proteasome inhibition with MG132, such as HDAC5 and the IFITs, some proteins were not rescued. A major group of these was the collagens. In the MG132 experiment, there were 8 collagens downregulated more than 2 fold, and 11 more than 1.5 fold. However, just the two quantified isoforms of COL6A2 were rescued by MG132. In the recent analysis by the Weekes lab of proteins degradation in HCMV infection, 10 of the 11 collagens downregulated by HCMV were transcriptionally regulated, rather than proteasomally (Nightingale *et al*, 2018). Transcriptional regulation during VACV infection is also a possibility, and RNA samples harvested alongside the proteomics ones could be used to confirm this. If this is the mechanism, it remains unclear whether the two viruses regulate the collagens in a similar manner, or whether this is a cellular response to infection.

4.8.2 Temporal Regulation of Viral Proteins

VACV orchestrates a careful cascade of temporally controlled viral protein expression. The DNA-dependent RNA polymerase and transcription factors for early genes are packaged within the virion. Stage specific transcription factors then recognise promoters of different temporal classes of viral proteins, to regulate their expression. Additionally, expression of intermediate and late genes relies on DNA replication (Baldick Jr. *et al*, 1993; Broyles & Fesler, 1990; Yang *et al*, 2013; Moss & Salzman, 1968; Keck *et al*, 1990). In this study, quantification of ~80 % of the viral proteins allowed k-means clustering into temporal classes, and identified 4 distinct profiles of expression, Tp1-Tp4. These correlated well with RNA-seq based assignments into the E1.1, E1.2, I and L classes (Yang *et al*, 2010, 2011, 2015), and with a previous, more limited, proteomic study (Croft *et al*, 2015). Additionally, it fitted well with the expected cascade of functions at various points during infection, with the earliest class of proteins functioning

largely in host interaction, moving through to viral replication and transcription, and finally virion associated proteins.

The rapid changes in the viral proteins present may be assisted by turnover of the mRNAs, and proteasomal degradation. Inhibition of the proteasome led to accumulation of early viral proteins, and limited expression of late ones. The proteasome is known to be required for VACV DNA replication, and therefore for the expression of intermediate and late viral proteins which is dependent on this (Satheshkumar *et al*, 2009; Teale *et al*, 2009; Mercer *et al*, 2012). Mercer *et al*. identified that the proteasome is required for viral uncoating (hence the addition of MG132 at 2 hpi in this investigation). However they also suggested that a second mechanism is involved in inhibiting viral DNA replication. They found that Cullin 3 (Cul3) and RING-box protein 1 (Rbx1), two components of an E3 ubiquitin ligase complex, were required for late viral gene expression and for normal levels of viral replication. A mechanism for the involvement of the proteasome in viral DNA replication has not yet been elucidated. The data presented here was consistent with these findings, showing early viral proteins accumulate with MG132 treatment, whilst the majority of intermediate proteins and all late viral protein expression was reduced with MG132. The reasons for dependence on the proteasome is unclear, it may be necessary for degradation of host antiviral factors such as the HDACs and IFITs discussed previously, in order to allow viral replication.

CHAPTER 5: CANDIDATE ANTIVIRAL RESTRICTION FACTORS

This chapter presents follow up work exploring results from the proteomic screens. Some of the cell lines produced, and restriction assays utilised, where developed by other members of the Weekes lab (in particular Dr Katie Nightingale, Dr Luis Nobre and Alice Fletcher-Etherington,, University of Cambridge), as indicated in the text, figure legends and methods.

5.1 Introduction and Aims

5.1.1 Identifying Candidate Antiviral Restriction Factors

Various characteristics of restriction factors were defined by Duggal *et al* (2012) (see Introduction 1.2.1), and two key features are used here to identify candidate ARFs: (1) stimulation by IFN (2) downregulation by viral factor. (Duggal & Emerman, 2012) From the work so far presented in this thesis, proteomic data is available on IFN stimulation at the surface of primary leukocytes, and on VACV infection in HFFF-TERTs. As well as this, proteomic studies previously published by the Weekes lab and others provide data on additional viral infections, in particular WCL and PM data on HCMV infection (Weekes *et al*, 2014) and HIV infection (Matheson *et al*, 2015; Greenwood *et al*, 2016). Overlap of the IFN screen with these viral datasets was performed to identify candidate antiviral restriction factors, and Endothelin Converting Enzyme 1 (ECE1) was identified as the major target. ECE1 is introduced further in section 5.3.

5.1.2 Aims

1) Identify candidate ARFs through the overlap of IFN and viral infection datasets.
2) Once a candidate restriction factor has been identified examine:
    a. How the virus downregulates the protein.
    b. Whether the protein acts to restrict viral infection, and if so, how.

5.2 Identifying Candidate Antiviral Restriction Factors

5.2.1 Overlap of IFN Simulation and Viral Infection Data

Overlap of the IFN stimulation with the VACV infection proteomic screens presented here, along with proteomic data on other viral infections, provided a basis for identifying candidate restriction factors (Table 5.1, Table 5.2). IFN stimulated proteins in both monocytes and T cells were examined ('sensitive' criteria, requiring the protein to have an IFN-induced FC > 1SD from the mean and a FC > 1 in all donors, as described in Chapter 3). Alongside this, infection with HCMV, VACV and HIV was analysed. For each virus, the maximal downregulation of each protein is given, and this was determined by considering FC at each time point examined by proteomics, and taking the minimum:

1. HCMV WCL and PM (Weekes *et al*, 2014): For WCL1 and PM1, the average expression at each time point was compared to the average mock and to 24 hpi. For WCL2 and PM2, the expression at each time point was compared to the average mock, 12 hpi and 18 hpi. Comparison to early infection time points was in order to account for proteins that were initially induced by IFN and then downregulated by the virus.
2. VACV WCL (Soday *et al*, 2019): Expression at 18 hpi was compared to the 18 h mock.
3. HIV WCL (Greenwood *et al*, 2016), HIV PM (Matheson *et al*, 2015): Expression at each time point was compared to an uninfected 0 h sample.

It is important to note that the data from the IFN and viral proteomic screens were matched on the basis of the gene symbol, and therefore may represent data on different isoforms of the protein. Blank spaces in the table indicate that the protein was not quantified in the given screen. Additionally, whilst the IFN data is in primary monocytes and CD4+ T cells, the HCMV and VACV proteomic time courses were performed in HFFF-TERTs, and the HIV one in CEM-T4s, a cultured T cell line.

Table 5.1 Effect of viral infection on proteins which were IFN stimulated in monocytes

IFN and viral induced FCs were calculated as explained in the text above. Blank spaces indicate that the protein was not quantified in the given screen. Viral data from (Weekes *et al*, 2014; Soday *et al*, 2019; Matheson *et al*, 2015; Greenwood *et al*, 2016).

Gene Symbol	FC on IFN Stimulation	HCMV				VACV			HIV	
		WCL1	WCL2	PM1	PM2	WCL1	WCL2	WCL3	WCL	PM
SSTR2	7.62									
CD69	4.29									0.93
SLAMF7	4.06									
CD38	3.88			0.89					0.87	0.64
FFAR2	3.69									
BST2	3.64			0.52	1.02				0.48	0.11
ISG15	3.30	1.03	0.32	0.93	1.20	0.99	0.73	0.85	0.98	
CD40	2.55		1.29	1.26	2.54	0.78	0.92	0.84		
SEMA4A	2.53									0.19
HLA-DRB4	2.45									
LILRB4	2.43								0.74	
SIGLEC1	2.41									
SLAMF1	2.41									
IL2RA	2.40									
CD83	2.29									0.20
LILRB4	2.23								0.74	
CD274	2.21		0.77	0.92	1.64		1.26	0.97		
HSPA6	2.14		1.44	0.70	3.43	1.03	1.26	1.29		
HLA-A	2.10	0.43	0.22	0.45	0.53	0.94	0.71	0.73	0.63	0.13
HLA-B	2.07					0.81	0.72	0.78		
CCR7	2.07									0.15
TMEM123	2.05			0.48	0.50					
HLA-F	2.02		0.36		0.77	0.57	0.80	0.65		
HLA-DRB1	1.92									
CALU	1.90	0.79	1.03	0.71	2.95	1.01	1.03	1.08	0.95	0.95
HLA-DRA	1.87									
HLA-E	1.82	0.49	0.68	1.12	2.42	1.27	0.74	0.48	1.16	
LILRA6	1.81									
HLA-DRB3	1.79									
SLC39A8	1.76			0.96	12.6				1.13	0.45
CD22	1.76									
CCR5	1.70									0.39
HLA-C	1.68			0.56	0.36	0.66	0.47	0.40	0.99	0.64
DMKN	1.65									0.67
BMPR2	1.64		0.89	1.25	1.38	0.87	1.19	0.70		0.85
GPR84	1.63									0.40
FAS	1.58	0.25	0.29	0.67	0.49	0.83	0.70	0.81	1.08	0.60
CD86	1.56									
HLA-DPB1	1.56									
CYBB	1.53									
LILRB2	1.52									

Table 5.2 Effect of viral infection on proteins which were IFN stimulated in T cells

IFN and viral induced FCs were calculated as explained in the text above. Blank spaces indicate that the protein was not quantified in the given screen. Viral data from (Weekes *et al*, 2014; Soday *et al*, 2019; Matheson *et al*, 2015; Greenwood *et al*, 2016).

Gene Symbol	Fold Change on IFN Stimulation	CMV				VACV			HIV	
		WCL1	WCL2	PM1	PM2	WCL1	WCL2	WCL3	WCL	PM
IL1RN	2.35		0.22							
ECE1	2.30	0.42	0.24	0.51	0.31	0.66	0.80	0.69	0.78	0.60
BST2	2.16			0.52	1.02				0.48	0.11
TMEM123	2.01			0.48	0.50					
CNP	1.46	0.99	0.88	0.80	0.98	0.79	0.80	0.83	0.64	0.74
KCNA3	1.45									
GPR171	1.37									
XRCC5	1.33	0.92	0.81	0.64	0.39	1.01	1.02	1.11	0.98	0.61
HLA C	1.32			0.56	0.36	0.66	0.47	0.40	0.99	0.64
PKP1	1.29		0.22							0.78
HLA F	1.26		0.36		0.77	0.57	0.80	0.65		
HLA E	1.26	0.49	0.68	1.12	2.42	1.27	0.74	0.48	1.16	
ORM1	1.22									

5.2.2 Candidate Antiviral Restriction Factors

Monocytes

BST2 (tetherin) is a well characterised, IFN-stimulated restriction factor for HIV (Perez-Caballero *et al*, 2009; Hammonds *et al*, 2010); the overlap shows it was both upregulated by IFN and also downregulated during HIV infection as has been previously observed (Neil *et al*, 2008; Van Damme *et al*, 2008). It is interesting that there were several proteins in the monocyte dataset which were quantified exclusively in the PM HIV data, and not the other viral time-courses. This is likely due to the enrichment for PM proteins, and because the HIV infection was of CEM-T4 cells rather than fibroblasts, and some of the proteins quantified are expressed predominantly in leukocytes. CCR7 was stimulated by IFN in primary monocytes and downregulated at the cell surface in HIV infection; it is thought to be downregulated by the HIV protein Vpu in order to limit migration to lymph nodes (Ramirez *et al*, 2014). However, not all of the proteins that fall into this overlap are known to have antiviral activity; they may be important in infection or immunity for

other reasons. C-C chemokine receptor type 5 (CCR5), for example, is known to act as the HIV coreceptor. Rather than being antiviral, upregulation by IFN may in fact result in an expanded cell tropism for HIV (Stoddart *et al*, 2010). Downregulation of CCR5 is by the HIV protein Nef, and may be in order to prevent superinfection (Michel *et al*, 2005). FAS was upregulated by IFN in primary monocytes, and downregulated by multiple viruses. It was previously observed to be downregulated at the cell surface during both adenovirus and HCMV infection, in order to evade Fas induced apoptosis (Shisler *et al*, 1997; Seirafian *et al*, 2014).

T cells

Fewer proteins were consistently upregulated by IFN in primary CD4+ T cells. IL1RN competes with IL-1 for binding to its receptor, preventing its pro-inflammatory functions (Arend *et al*, 1998). IL1RN was quantified in three of the five donors, and upregulated by IFN in all of these. It also appears to be downregulated by HCMV infection, although it should be noted that it was only quantified here with a single peptide, and has a complex and variable effect during infection with multiple viruses (Kline *et al*, 1994; Hill-Batorski *et al*, 2015; Yoon *et al*, 1999). X-ray repair cross-complementing protein 5 (XRCC5), also known as Ku80, was upregulated by IFN and downregulated to some extent at the cell surface by infection with HCMV or HIV; examination of the individual datasets showed these effects were modest. Ku80 forms part of the DNA-PK complex, which has been associated with DNA sensing and innate immune functions (Ferguson *et al*, 2012; Morchikh *et al*, 2017). However, it also has roles at the cell surface in cell-cell interactions and adhesion (Monferran *et al*, 2004). CNP was also downregulated to some extent by HIV, and is a known ISG which restricts HIV infection (Wilson *et al*, 2012).

Candidate ARFs targeted by HCMV presented the most tractable targets, due to the establishment of HCMV restriction assays within the Weekes lab. Although TMEM123 may be an interesting candidate for follow up studies, the lack of prior literature and reagents limited this (see chapter 3.8). ECE1 was upregulated in all five donors in primary T cells, and downregulated during HCMV infection, and to some extent by the other viruses; it was therefore chosen for further investigation.

5.3 Endothelin Converting Enzyme 1 (ECE1)

ECE1 is a zinc metallopeptidase and type II integral membrane protein. It was first identified as a 758 amino acid polypeptide, expressed abundantly in epithelial cells and responsible for the production of the vasoactive molecule Endothelin (ET), by cleavage of an inactive intermediate. It shows sequence similarity to the peptidase Neprilysin and the human Kell blood group protein (Xu *et al*, 1994). There are four isoforms, with some localised to the cell surface where they are present as homodimers, and some intracellular. ECE1 has also been shown to be shed from endothelial cells (Kuruppu *et al*, 2007).

5.3.1 Isoforms

The four isoforms of ECE1 are all generated from different promoters in a single gene. They share a common, large C terminal domain containing the HEXXH zinc binding motif and have differing N terminal cytoplasmic tails (Figure 5.1). They are all thought to cleave ET with equal efficiency, but they vary in their subcellular localisation, and expression across different tissues (Valdenaire *et al*, 1999; Schweizer *et al*, 1997).

The first three isoforms of ECE1 discovered were ECE-1a, 1b and 1c. ECE-1a is the most highly expressed at the cell surface, followed by ECE-1c. ECE-1b is found at intracellular compartments (Schweizer *et al*, 1997; Barnes *et al*, 1998). ECE-1d was identified later, and found to have an intermediate distribution, present both at the PM and intracellularly in endosomal structures (Valdenaire *et al*, 1999). Finer definition of their intracellular localisations showed that ECE-1b concentrates in late endosomes/multivesicular bodies, whilst ECE-1d is in recycling endosomes. However, the isoforms may heterodimerize, in which case the localisation of ECE-1b acts as the dominant regulatory signal, resulting in intracellular localisation of other isoforms (Muller *et al*, 2003). ECE-1a may also be targeted to the nucleus (Jafri & Ergul, 2003). The tissue distribution of each isoform also differs; ECE-1a for example has a high level of expression in endothelial cells, but not smooth muscle cells (Valdenaire *et al*, 1999).

Figure of ECE1 isoforms and gene organisation removed for copyright reasons. Copyright holder is John Wiley and Sons.

Figure 5.1 Isoforms of ECE1 and gene organisation

The organisation of the 5' end of the ECE1 gene is shown, and how the four exons contribute to each isoform (based on Valdenaire *et al*, 1999). Exons 4-19, which are common to all isoforms and constitute the extracellular domain, have been omitted. Transcription is driven from the promoters (P) upstream of the four exons, and these transcripts are spliced to generate the isoforms. The length of each isoform in amino acids is given, along with the predominant localisation (sources of this information given in text).

5.3.2 Function

Endothelin

The main function of ECE1 is to cleave the inactive precursor big-ET into its mature vasoactive form. There are three isoforms of ET, with ET-1 being cleaved at Trp21-Val22 (Létourneau *et al*, 2000). ET-1 is usually produced by endothelial cells, but can also be produced by other cells including macrophages (Ehrenreich *et al*, 1990, 1993). Endothelin acts via the ETA and ETB receptors, with the exact function depending on the localisation and type of the receptor, but the overall effect is of ET-1 as a potent vasoconstrictor and pro-inflammatory peptide (Houde *et al*, 2016). It is also known to induce expression of monocyte chemotactic protein-1 (MCP1) and IL-8 (Helset *et al*, 1994), and the endothelin system involving ECE1, ETB and ET-1 is involved in monocyte diapedesis of the brain endothelial cell barrier (Reijerkerk *et al*, 2012).

Cleavage of other peptides

Analysis of a purified, recombinant soluble form of ECE1 showed it is also able to cleave other biologically active peptides, including Bradykinin (Pro7-Phe8), angiotensin I (predominantly pro7-phe8 but also some other residues), neurotensin (pro10-tyr11, leu2-tyr3) and substance P (SP, gln6-phe7, gly9-leu10), as well as having various cleavage sites in other peptides. The study demonstrated that ECE1 cleavage sites are usually found on the amino side of hydrophobic residues. Substrates smaller than hexapeptides were not efficiently hydrolysed (Johnson *et al*, 1999). The cleavage of substrates by ECE1 is pH dependent, with hydrolysis of big ET-1 being most efficient at a neutral pH, but smaller substrates being cleaved at an acidic pH (Fahnoe *et al*, 2000).

Receptor Recycling and Resensitisation

ECE1 regulates receptor recycling and resensitisation through metabolism of peptide ligands in endosomes (Figure 5.2). SP binds the neurokinin 1 receptor (Nk_1R) at the PM, leading to activation of NF-κB and proinflammatory cytokines. Translocation of β–arrestin to the membrane triggers receptor desensitisation and internalisation, and then recruitment of proto-oncogene tyrosine-protein kinase Src, dual specificity MAPK-kinase (MEK) and extracellular signal-regulated protein kinase 1/2 (ERK1/2), leading to MAPK signalling. ECE1 is able to degrade SP in the early endosomes, terminating this signal and allowing recycling of Nk_1R (Roosterman *et al*, 2007). Inhibiting ECE1 therefore leads to increased signalling through ERK1/2, which triggers phosphorylation of the nuclear death receptor, nuclear receptor subfamily 4 group A member 1 (Nur77) and cell death (Cottrell *et al*, 2009). It also attenuates inflammation by preventing the recycling and resensitisation of the receptor (Cattaruzza *et al*, 2013). ECE1 is also involved in recycling of other receptors, degrading calcitonin gene related peptide (CGRP) to promote calcitonin receptor-like receptor / receptor activity-modifying protein 1 (CLR/RAMP) recycling (Padilla *et al*, 2007), and degrading glucagon-like peptide 1 (GLP-1) to recycle and resensitise the GLP-1 receptor (Lu & Willars, 2019). ECE1 has also been implicated in the cleavage and recycling of TLR9 in endosomes, regulating TLR9 mediated cytokine release (Julian *et al*, 2013).

Other functions and homologues

Finally, ECE1 may have other functions. It has been implicated in many diseases, and homologues may also suggest additional roles. The homologue M13 is found in invertebrates. This suggests that the enzyme is evolutionarily conserved, and as invertebrates do not have a closed circulatory system emphasises roles outside of blood pressure regulation (Macours *et al*, 2004).

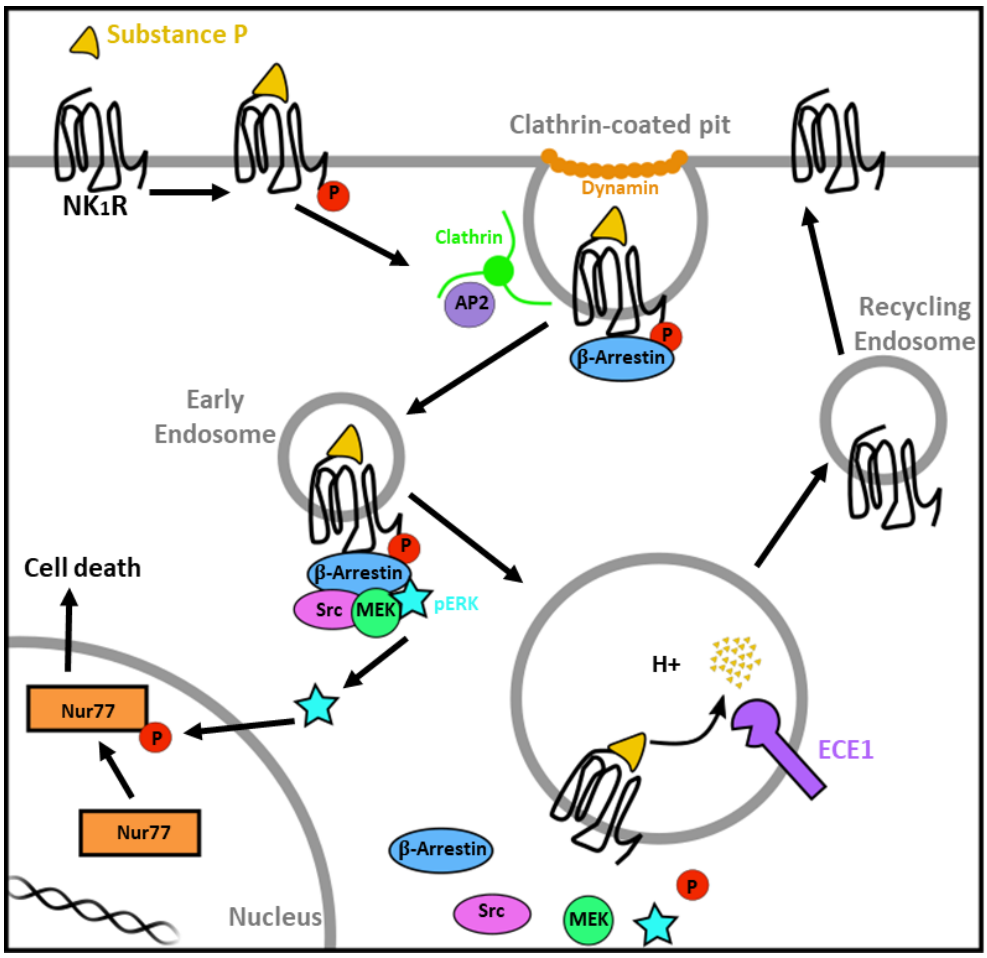


Figure 5.2 ECE1 involvement in substance P degradation and receptor recycling

Substance P activates NK_1R, inducing phosphorylation of the receptor and translocation of β-arrestin to the membrane. This triggers receptor desensitisation and endocytosis, as well as recruitment of the components of a MAPK signalling module (Src, MEK, ERK1/2). In the acidified endosome, ECE1 degrades substance P, causing the complex to disassemble, and dephosphorylation of NK_1R. The receptor is recycled to the PM, mediating resensitisation. In the absence of ECE1, sustained signalling through ERK leads to phosphorylation and activation of Nur77, leading to cell death. Mechanism of ECE1 regulation of SP-stimulated ERK activation and cell death proposed by (Cottrell *et al*, 2009), and figure reproduced based on this research originally published in the Journal of Biological Chemistry © American Society for Biochemistry and Molecular Biology.

5.3.3 Regulation

Expression

Previous research has implicated a variety of factors in regulating ECE1 expression, which may vary dependent on context and cell type. It is upregulated in endothelial cells by the cell surface TNF-like weak inducer of apoptosis (TWEAK), through the modulation of AP-1 and NF-κb transcription factors (Martínez-Miguel *et al*, 2017). In addition, expression of ECE1 in human umbilical vein endothelial cells could be induced by some chemokines; MCP1 was the most potent stimulator, but macrophage inflammatory protein 1α and 1β, IL-8, and RANTES (regulated on activation, normal T cell expressed and secreted) all stimulated ECE1, as did a mixture of cytokines (IL-1β, TNF-α, IFN-γ) (Molet *et al*, 2000). A study in mice found that poly(I:C) transfection led to increased *ECE1* mRNA in skin, and that this effect could be blocked by knockout of TLR3, but not the IFN receptor (Farina *et al*, 2011). This suggests that ECE1 is upregulated by the TLR3 pathway and this is independent of IFN. ECE1 is also upregulated by shear stress, hyperochesterolemia, during statin therapy and due to high glucose levels. It is downregulated by the influx of calcium triggered by ETB receptor activation, acting as a negative feedback loop when ET-1 concentrations are high (Kuruppu & Smith, 2012). There are therefore many interacting factors and a great deal of complexity in the regulation of ECE1 expression.

Localisation

In CHO cells, PKC mediated phosphorylation triggers ECE-1a and ECE-1b internalisation. PKA mediated phosphorylation only led to ECE-1b internalisation as ECE-1a does not have this phosphorylation site. Conversely, PMA mediated phosphorylation or high glucose concentrations, lead to trafficking of ECE-1c to the cell surface (Kuruppu & Smith, 2012; Kuruppu *et al*, 2012). In the AtT-20 neuroendocrine cells, ECE-1a and 1c are at the PM, and 1b and 1d in the endosomal system. Formation of heterodimers with ECE-1b was able to regulate localisation of other isoforms (Muller *et al*, 2003).

5.3.4 ECE1 in Disease and Infection

Levels of ECE1 or of ET have been linked to viral myocarditis in a murine model, with increasing levels of ECE1 in murine heart tissue following infection with encephalomyocarditis virus (Ono *et al*, 1999). Increased levels of ECE1 and ET in the placenta are also observed in infection associated pre-term birth (Olgun *et al*, 2010; Wang *et al*, 2010), and in chronic rhinitis (Furukawa *et al*, 1996). The HIV proteins gp120 and trans-activator of transcription (TAT) have been reported to upregulate ET-1, and this may be related to HIV-associated pulmonary hypertension, though the role of ECE1 here has not been investigated (Ehrenreich *et al*, 1993; Chauhan *et al*, 2007). ECE1 has also been linked to cell invasion in multiple cancers, including prostate cancer (Lambert *et al*, 2008; Hong *et al*, 2011), breast cancer (Smollich *et al*, 2007) and colorectal cancer (Pérez-Moreno *et al*, 2019) amongst others, with the different isoforms having different effects (Tapia & Niechi, 2019). ECE1 also has a role in Alzheimers (Eckman *et al*, 2003), and cardiovascular disease (Martínez-Miguel *et al*, 2009), as well as an inactivating mutation of ECE1 being associated with Hirschsprung disease (Hallett *et al*, 1992).

5.3.5 Potential Roles as a Restriction Factor

Based on its known functions, there are several ways that ECE1 might be predicted to restrict viral infection. As ECE1 is able to cleave peptides, a potential mechanism could involve hydrolysis of peptides necessary for viral replication, or potentially even of the virus itself at the cell surface or in endosomes. Alternatively, ECE1 may be involved in the recycling and resensitisation of immune receptors. A major role of ECE1 is in cleavage of big-ET to its mature form, so another possibility is that it is the increased level of ET that is important. If a role in infection was established, there would therefore be several interesting avenues for exploration.

5.4 ECE1: A Candidate Antiviral Restriction Factor

5.4.1 Validation of IFN Stimulation

Upon IFN stimulation, ECE1 was upregulated more than 1.5 fold in all five donors (Figure 5.3A). The average FC with IFN stimulation was greater than that observed for BST2, a known restriction factor. This proteomic data was validated by RT-qPCR in primary CD4+ T cells (Figure 5.3B), and by immunoblot in two cultured T cell lines, Jurkats and SUPT1s (Figure 5.3C). Identification of a suitable antibody for this is discussed in section 5.5.

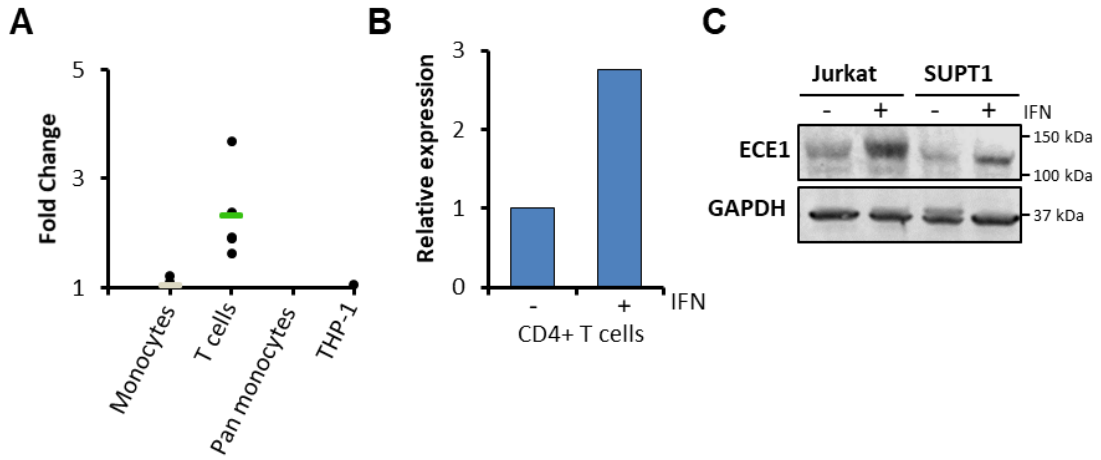


Figure 5.3 Validation of IFN stimulation of ECE1

(A) FC of ECE1 upon IFN stimulation in each cell type examined by proteomics.

(B) RT-qPCR validating IFN stimulation of ECE1 in primary CD4+ T cells.

(C) Immunoblot validating IFN stimulation of ECE1 in two cultured T cell lines, Jurkats and SUPT1s.

5.4.2 How Does HCMV Downregulate ECE1?

In addition to being stimulated by IFN, ECE1 was highlighted as a candidate restriction factor due to also being downregulated by viral infection. Over 96 h of infection of fibroblasts with HCMV, ECE1 was substantially downregulated at both the whole cell and PM level (Weekes *et al*, 2014) (Figure 5.4).

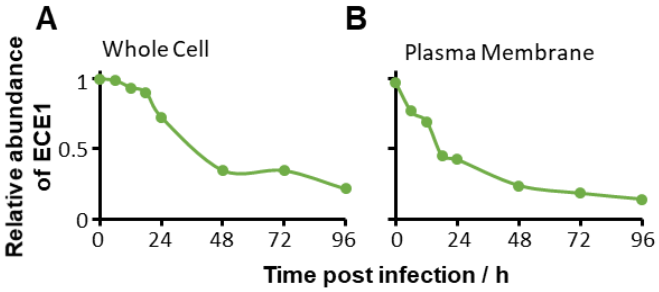


Figure 5.4 ECE1 is downregulated during infection with HCMV

Data from Weekes *et al*., 2014.

(A) Downregulation of ECE1 at the whole cell level during HCMV infection

(B) Downregulation of ECE1 at the cell surface during HCMV infection

Further data was examined to investigate how HCMV might downregulate ECE1. Dr Katie Nightingale (Weekes Lab, Department of Medicine, University of Cambridge) in collaboration with Dr Richard Stanton (Cardiff University) performed a proteomic screen whereby HFFF-TERTs were infected with HCMV mutants, each with a different block of genes deleted (Nightingale *et al*, 2018). Deletion of the US1-US11 block led to increased expression of ECE1 compared to WT infection or other mutants (Figure 5.5A). This suggests that a gene in this block is responsible for downregulation of ECE1 during infection. Furthermore, a proteomic examination of the HCMV protein US2 was previously published (Hsu *et al*, 2015). In this, fibroblasts were infected with WT HCMV or a US2 deletion mutant, and examined by MS. A host of proteins were rescued by deletion of US2, including multiple integrin family members, and ECE1 (Figure 5.5B). US2 is therefore necessary for ECE1 downregulation. An immunoblot of fibroblasts expressing US2 demonstrated that US2 was also sufficient for degradation of ECE1 (Figure 5.5). The US2 expressing cell line was kindly provided by Dr Luis Nobre (Weekes Lab, Department of Medicine, University of Cambridge).

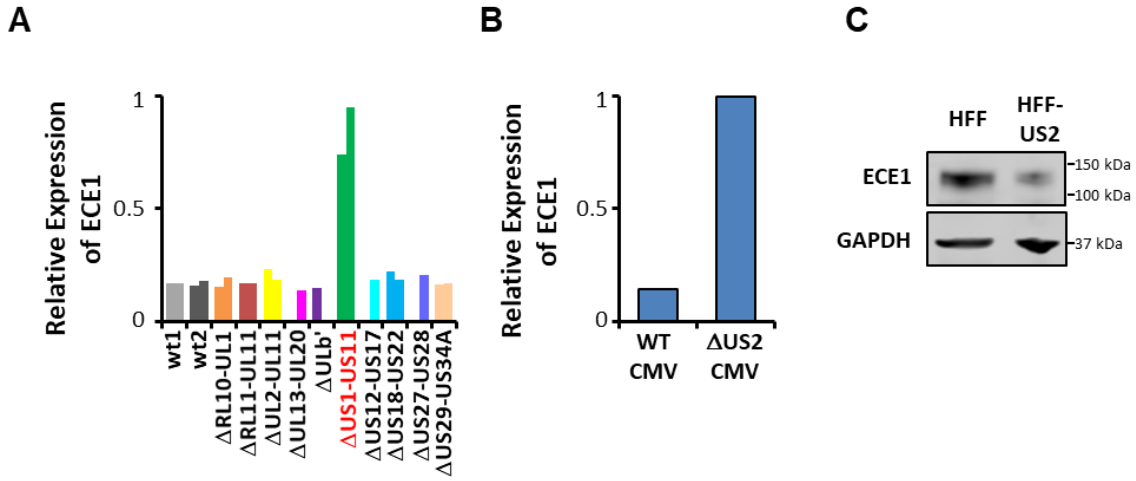


Figure 5.5 The HCMV proteins US2 is necessary and sufficient for degradation of ECE1

(A) HFFF-TERTs were infected with either WT HCMV or various mutant viruses with different blocks of genes deleted, before WCL analysis by TMT-based proteomics. Deletion of the US1-11 gene block resulted in increased expression of ECE1 compared to infection with the WT virus (data from Nightingale *et al*, 2018).

(B) ECE1 expression is increased during infection with a US2 deletion mutant compared to infection with WT HCMV, demonstrating that US2 is necessary for downregulation of ECE1 during infection (data from Hsu *et al*, 2015).

(C) Immunoblot of WT fibroblasts or cells expressing the viral protein US2, demonstrating that US2 alone was sufficient for downregulation of ECE1.

5.5 ECE1: Reagents

5.5.1 ECE1 Antibodies

Several different unconjugated antibodies were trialled for detection of ECE1 (Table 5.3). Ideally, the PM proteomic data would be validated by cell surface flow cytometry, however intracellular flow cytometry and immunoblot were also tested.

Table 5.3 Table of antibodies trialled for ECE1 detection

Four antibodies were trialled for detection of ECE1 by cell surface or intracellular flow cytometry and immunoblot. Applications tested by supplier: flow cytometry (FCM), immunohistochemistry (IHC), western blot (WB). The Proteintech antibody was still undergoing quality control checks, so some information was not available.

Supplier	Product Code	Species	Polyclonal or Monoclonal?	Residues Raised Against	Applications Tested by Supplier
Invitrogen	PA5-24563	Rabbit	Polyclonal	528-556	FCM, IHC, WB
Abcam	ab71829	Rabbit	Polyclonal	700-770	WB
R&D	MAB17841	Rat	Monoclonal	90-770	FCM
Proteintech	N/A	Rabbit	-	-	-

5.5.2 Testing ECE1 Antibodies by Flow Cytometry

Cell Surface Flow Cytometry

As the proteomic data on IFN stimulation of ECE1 was from PM samples, cell surface flow cytometry would be the ideal technique for validation and further investigation. The Invitrogen, Abcam and R&D antibodies were tested for this application. ECE1 is a 770 amino acid single pass type II membrane protein, with residues 90-770 being extracellular, and all antibodies were raised against extracellular epitopes (Table 5.3). Unfortunately, none of the antibodies showed substantial separation in signal compared to unstained samples or those stained with only a secondary antibody (Figure 5.6A). Furthermore, the signal from CRISPR/Cas9 ECE1 knockout cells was the same as that observed in ECE1 overexpressing and control cells (Figure 5.6B, cell lines discussed in section 5.6.2 and 5.6.4). Multiple concentrations of the antibodies were tested and similar results were observed. One caveat to this data is however, that the ECE1 overexpressing cell line was later determined to be isoform ECE-1b, which is predominantly intracellular. Additionally, it can heterodimerise with other isoforms and is the dominant protein in determining their location.

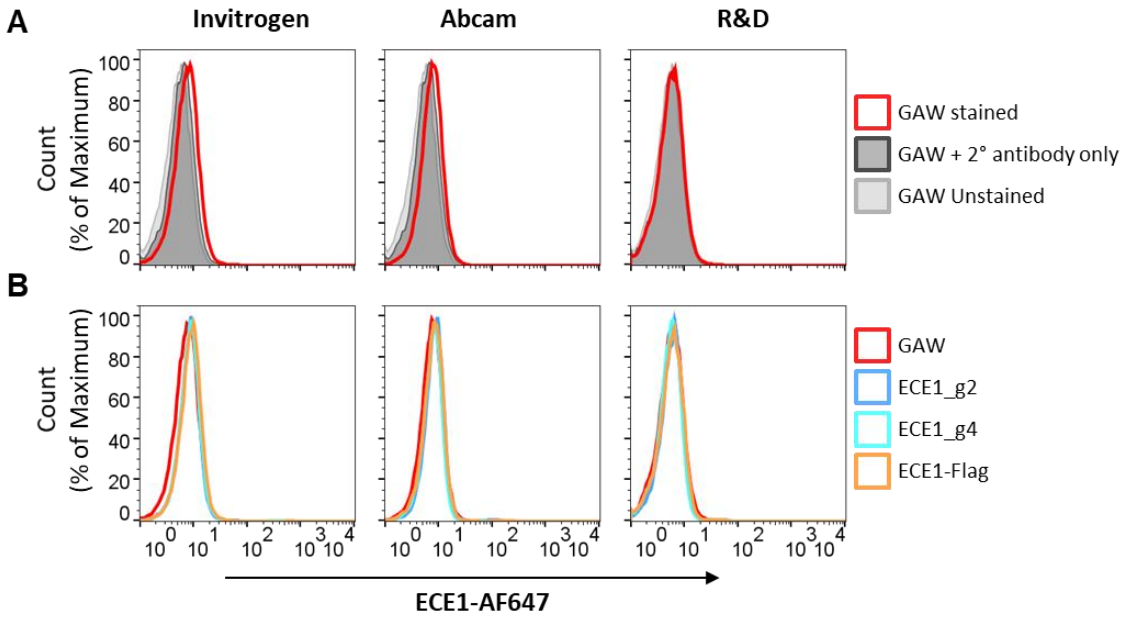


Figure 5.6 Testing ECE1 antibodies for cell surface flow cytometry on HFFF-TERTs

(A) Cells expressing the GAW control pHAGE-PSFFV plasmid were analysed by cell surface flow cytometry. They were either unstained (light grey), stained with only the relevant AF647-conjugated secondary antibody (rabbit for Invitrogen and Abcam, rat for R&D; dark grey), or stained with both primary and secondary antibodies (red).

(B) Overlay of signal from staining of GAW control cells (red) compared to CRISPR knockout polyclonal populations (blue, ECE1_g2, ECE1_g4) and cells overexpressing ECE1-Flag (orange), using the stated antibodies.

Intracellular Flow Cytometry

The same antibodies were additionally tested for intracellular flow cytometry. In this case, the HFFF-TERTs were fixed using 4% paraformaldehyde and then permeabilised using methanol, prior to antibody staining. Despite all three antibodies showing a more promising distinction between the stained samples and those incubated with just the secondary antibody (Figure 5.7A), none showed consistent correlation between the level of ECE1 expected in the CRISPR and overexpression cell lines, and the signal detected by flow cytometry (Figure 5.7B). As an antibody had been validated for immunoblot (see section 5.5.3), and no promising results were obtained for cell surface flow cytometry, no further examination of antibodies for detection of ECE1 by flow cytometry was conducted.

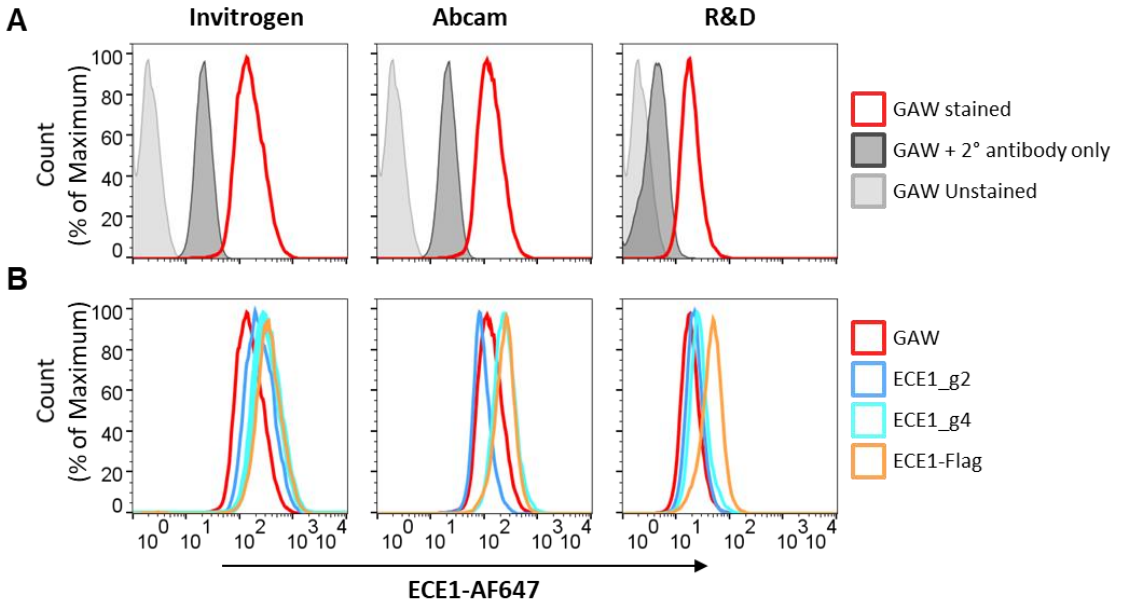


Figure 5.7 Testing ECE1 antibodies for intracellular flow cytometry on HFFF-TERTS

(A) Cells expressing the GAW control pHAGE-PSFFV plasmid were analysed by intracellular flow cytometry. They were either unstained (light grey), stained with only the relevant AF647-conjugated secondary antibody (rabbit for Invitrogen and Abcam, rat for R&D; dark grey), or stained with both primary and secondary antibodies (red).

(B) Overlay of signal from staining of GAW control cells (red) compared to CRISPR knockout polyclonal populations (blue, ECE1_g1, ECE1_g4) and cells overexpressing ECE1-Flag (orange), using the stated antibodies.

5.5.3 Testing ECE1 Antibodies by Immunoblot

The four antibodies described in Table 5.3 were trialled on three different cancer cell lines for detection of ECE1 by immunoblot. DLD1 and SW48 cells are both colorectal cancer cell lines (kindly gifted by Dr Matthew Hoare, Cancer Research UK, Cambridge) and in proteomic data available in the Weekes lab, DLD1 cells were shown to have higher expression of ECE1 than SW48 cells (Weekes and Gygi *et al*, unpublished). In a separate experiment, MDA-MB-231 cells showed high expression of ECE1 relative to other breast cancer cell lines (Weekes and Gygi *et al*, unpublished), and ECE1 should therefore be readily detectable in this cell line (kindly gifted by Proffessor Paul Lehner).

All four antibodies were trialled for detection of ECE1 by immunoblot (Figure 5.8A). Using the R&D antibody, no bands were observed. As it was not advertised by the

supplier as being suitable for readout by immunoblot, no further optimisation was performed with the R&D antibody. The Proteintech antibody resulted in many bands; as it had not yet passed quality control, further testing with this was also abandoned. The Invitrogen and Abcam antibodies both gave promising results, and as the Abcam antibody produced bands of the expected size this was used for all further experiments (Figure 5.8B). Immunoblot analysis of the subsequently produced overexpressing and knockout cell lines provided additional confidence on detection of ECE1 by the Abcam antibody (eg Figure 5.10, Figure 5.15).

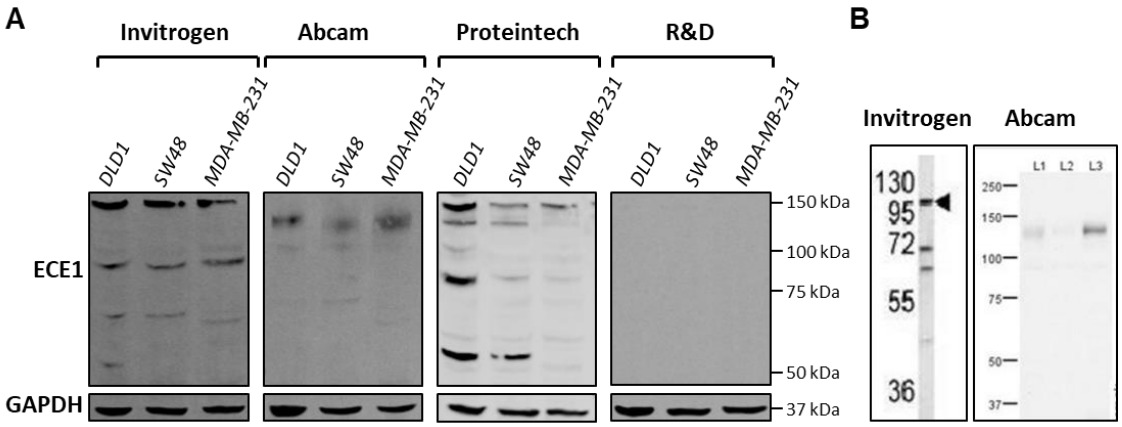


Figure 5.8 Immunoblot analysis of ECE1 antibodies

(A) Immunoblot of indicated anti-ECE1 antibodies on various cancer cell lines.

(B) Example immunoblots given on the manufacturer's website for ECE1 from the Invitrogen and Abcam anti-ECE1 antibodies. The Invitrogen antibody was tested on A2058 cell lysates, whilst the Abcam antibody probed HUVEC (L1), MEF1 (L2) and NIH-3T3 (L3) lysates.

5.6 ECE1: Does it Restrict HCMV Infection?

5.6.1 Restriction Assay Methodology

In order to examine the possibility of ECE1 being a novel ARF, restriction assays were performed to assess the effects of ECE1 expression on HCMV infection. All cell lines were produced in HFFF-TERTs as these are permissive to infection with the Merlin strain of HCMV, and previous data demonstrated that the virus targets ECE1 in this cell line. ECE1 overexpression cells were produced, as well as knockdowns using either shRNAs

or the CRISPR/cas9 system. For the restriction assay with overexpression cell lines, gateway (GAW) control cells and cells overexpressing ECE1 were seeded in triplicate for each MOI to be analysed. The cells were infected with HCMV in which the immediate-early protein UL36 was fused to GFP at the C terminus with a self-cleaving P2A peptide, so GFP would be released on translation. UL36 was chosen for these assays as it was one of the most abundantly expressed viral proteins early in infection, and the fusion of GFP did not disturb function (Nightingale *et al*, 2018). After 24 h the samples were harvested and fixed for analysis by flow cytometry (Figure 5.9). If ECE1 is a restriction factor, a reduced percentage of GFP positive cells would be expected to be observed in the overexpressing cells compared to controls. The opposite would be true of the knockdown cells, where an increased percent infection would be expected if the protein were a restriction factor.

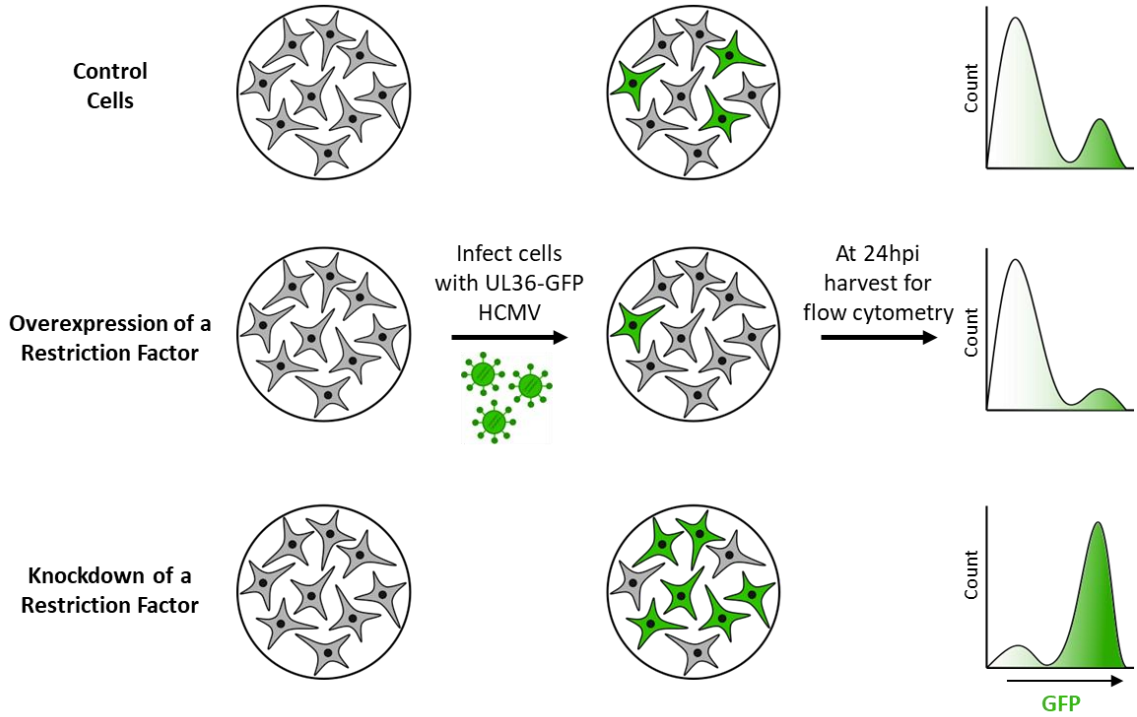


Figure 5.9 Schematic of HCMV restriction assay

Control, knockdown or overexpression cells were infected at low MOI with HCMV, where the immediate-early viral protein UL36 was fused to GFP. After 24 h the cells were harvested and analysed by flow cytometry. If the protein were a restriction factor, you would expect to see decreased infection of the overexpressing cells compared to controls, and therefore a reduced percent of GFP positive cells. Conversely, increased infection would be expected in the knockdown cells compared to the control, and therefore a higher percentage of GFP positive cells would be observed.

Low MOIs have traditionally been used for restriction assays (Adler *et al*, 2011; Nightingale *et al*, 2018). There is some previous evidence that restriction may be saturated or overcome at higher MOIs (Griffin & Goff, 2015; Nightingale *et al*, 2018; Tavalai *et al*, 2011). Therefore, a range of low MOIs were used for each infection.

The read out from these assays was detection of fluorescence from the GFP tagged immediate-early UL36 protein, by flow cytometry. Detection of GFP therefore suggests the virus has successfully entered the cells and begun expression of immediate-early viral genes. The assay therefore does not examine the efficiency of viral assembly or exit.

This assay was adapted from one developed by (Tavalai *et al*, 2011), examining the effect of Sp100 on HCMV infection. Positive controls used for these assays were the previously identified restriction factors, Sp100 and Daxx, both of which are components of ND10 bodies (Kim *et al*, 2011b; Adler *et al*, 2011; Woodhall *et al*, 2006).

5.6.2 Overexpression

Lentiviral transduction was used to produce cell lines overexpressing FLAG tagged ECE1. This was expressed from the pHAGE-P_{SFFV} plasmid, which had been generated using the gateway system to transfer the ECE1 sequence from a donor plasmid to the expression vector (see Methods 2.4.1). The control cells were transduced with the GAW control plasmid (GAW). Overexpression of ECE1 was validated by immunoblot (Figure 5.10). Some caution should be exerted in interpreting these results as the plasmid was overexpressing ECE-1b, a predominantly intracellular isoform of ECE1, and therefore possibly not the one detected in the PM proteomics.

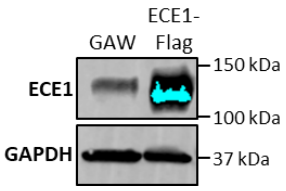


Figure 5.10 Immunoblot confirming overexpression of ECE1 in HFFF-TERTs

HFFF-TERTs expressing the GAW control plasmid or ECE1-Flag. Unfortunately, the gel was over-exposed on imaging resulting in the blue band, but still shows clear overexpression.

The restriction assay was performed as described in Figure 5.9. The assay was repeated four times on a range of low MOIs (Figure 5.11). Sp100 and Daxx were included as positive controls in three of the experiments. In almost all cases, overexpression of Sp100 and Daxx resulted in a significantly reduced percent infection compared to GAW control cells ($p < 0.05$); however, this reduction was much more dramatic with Daxx than Sp100. Overexpression of ECE1 also resulted in a clear reduction in HCMV infection, though more similar in extent to Sp100 than Daxx. This provided some initial support for ECE1 acting to restrict early stages of HCMV infection.

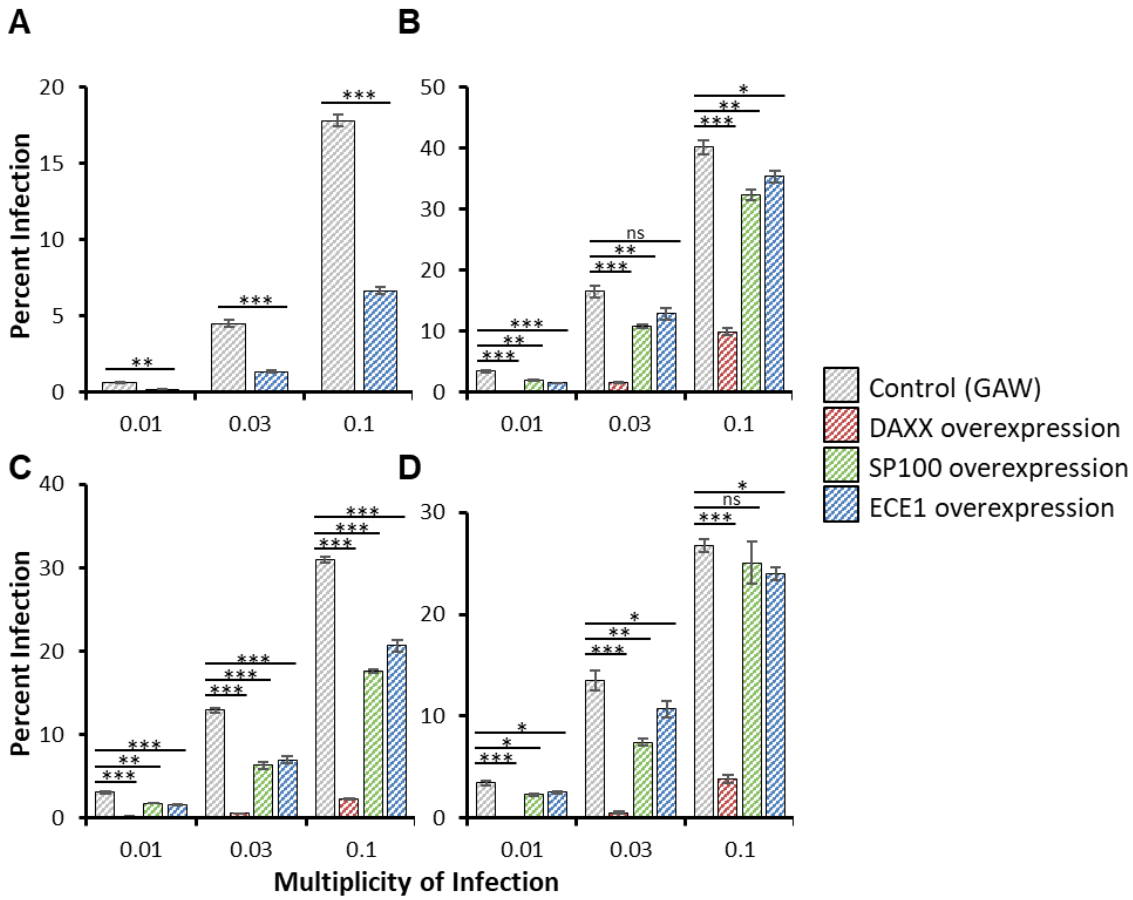


Figure 5.11 HCMV restriction assay on HFFF-TERTs overexpressing ECE1

(A-D) In all cases, cells were infected with UL36-GFP HCMV for 24 h before harvesting and analysis of the percent infection by flow cytometry. The control cell line was transduced with the GAW control plasmid. Sp100 and Daxx were included as positive controls for restriction factors. Overexpression of ECE1-Flag led to reduced percent infection in all cases, similar to that observed with Sp100 but to a much lesser extent than with Daxx. Data is represented as the mean of triplicate data ± SEM. P-values were estimated using a two-tailed t-test, with each sample compared to the GAW control cells; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ns = not significant.

5.6.3 shRNA knockdown

A complementary approach to the overexpression restriction assay examined two knockdown cell lines, Sh1 and Sh2. These were produced in HFFF-TERTs, each using an shRNA targeting a different sequence which was common to all isoforms of ECE1. As an antibody had not yet been identified for immunoblot at this stage, RT-qPCR was used to confirm a knockdown of 3.3 and 4.3 fold in ECE1 Sh1 and Sh2 cells respectively (Figure 5.12). Sp100 knockdown cells were produced by Dr Katie Nightingale (Weekes Lab) and validated by immunoblot (Nightingale *et al*, 2018).

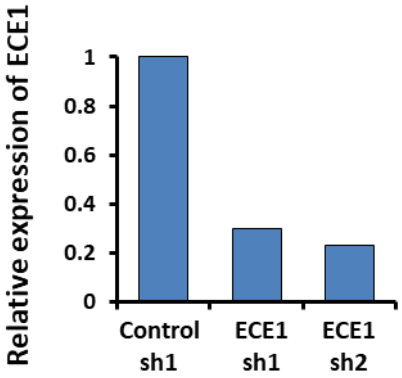


Figure 5.12 Validation of ECE1 knockdown in HFFF-TERTs by RT-qPCR

RT-qPCR was used to confirm the knockdown of ECE1 in HFFF-TERTs by shRNA. The two different shRNAs resulted in 3.3 and 4.3 fold knockdowns compared to the cell line expressing a control shRNA.

Restriction assays were performed as depicted in Figure 5.9. A range of low MOIs were used, and results from a subset of these in three different experiments are shown in Figure 5.13. Results here were inconclusive. The percent infection observed in the ECE1 shRNA knockdown cells was variable, with Figure 5.13C showing a clear restrictive effect of ECE1, and more variable results in Figure 5.13A and B, particularly with ECE1_sh1. Additionally, despite using the same Sp100_sh1 positive control cell line as was previously used in restriction assays (Nightingale *et al*, 2018), a reliable increase in the percent infection compared to the control cells was not observed. It was therefore difficult to draw definitive conclusions from this data.

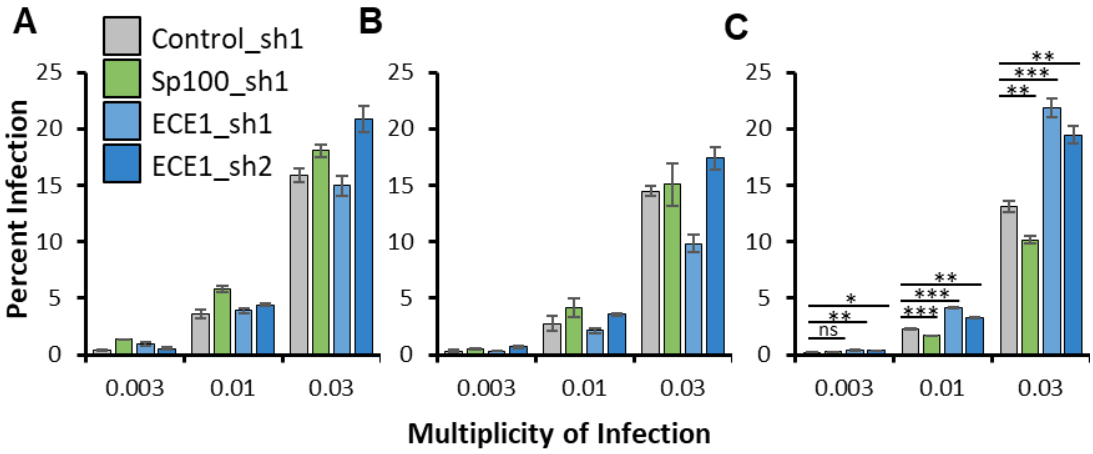


Figure 5.13 HCMV restriction assay on ECE1 shRNA knockdown HFFF-TERTs

In all cases, control shRNA or ECE1 knockdown cells were infected with UL36-GFP HCMV for 24 h before harvesting. The percent infection was assessed by flow cytometry. Sp100 was included as a positive control for a restriction factor.

(A) Bars represent the mean of duplicate data, and error bars display the range (except for the Sp100 data at an MOI of 0.003 where only one data point was available).

(B) Bars represent the mean of duplicate data, and error bars display the range.

(C) Bars represent mean of triplicate data ± SEM. P-values were estimated using a two-tailed t-test, with each samples compared to control_sh1; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ns = not significant.

There was some concern that the shRNA knockdown of ECE1 was being lost over time, resulting in variable effects depending on how long the cells had been passaged for prior to the assay. Following these restriction assays, an antibody had been identified for detection of ECE1. This was used to test the original lysates, and some lysates made from cells which had been passaged several times (passage ~5 following selection). It did appear that the level of knockdown was reduced, particularly in the sh2 cells, even after a low number of passages (Figure 5.14). This suggests either unreliable determination of the knockdown from the first set of lysates made immediately following selection, or the possibility that the knockdown might be lost completely following more time in culture, despite cells stably expressing the shRNA. Research forums suggest this is not an uncommon problem, and it has been reported before (Salazar *et al*, 2014). One possible reason for this phenomenon is that depending on where the shRNA integrates, it may not be highly expressed, or could later be silenced. It may also be possible that the cells were

not selected for long enough or should have been maintained in antibiotic selection to ensure only those expressing the shRNA survived. Although anecdotal evidence, other researchers on forums have seen maintenance of antibiotic resistance with loss of the shRNA knockdown. A possible explanation for this, is that the level of knockdown in different cells within the population was variable and those with a lower level of knockdown may have outcompeted the others over time (Sigma). It may therefore be important in future studies to only use very early passage cells, and to re-confirm the knockdown prior to experiments.

Figure 5.14 ECE1 knockdown in HFFF-TERTs was diminished following passaging

Immunoblot of lysates produced when the shRNA cells were originally made (no antibody available at the time) being probed for ECE1, alongside lysates produced from cells frozen after ~5 passages ('retest').

5.6.4 CRISPR knockdown

In order to improve the level of the knockdown and eliminate the concern regarding loss of the knockdown over passages, CRISPR/Cas9 was employed to generate knockout cell lines. The CRISPR/Cas9 technology has become commonly used in molecular biology. It relies on the Cas9 endonuclease being targeted to specific DNA sequences according to complementarity to a guide RNA (gRNA). Cas9 then generates double strand breaks, which when repaired by the error prone non-homologous end joining DNA repair mechanism, will often result in an insertion or deletion mutation (Sander & Joung, 2014; Sanjana *et al*, 2014; Shalem *et al*, 2014). Cells constitutively expressing Cas9 were transduced with lentiviruses expressing the gRNAs. Multiple cell lines were produced, expressing guides targeting different sequences in the gene of interest, or non-targeting

control sequences, and these were used in the restriction assay described previously (Figure 5.9).

Restriction assay with polyclonal CRISPR population

Polyclonal CRISPR populations were produced by transducing HFFF-TERTs constitutively expressing Cas9 with the gRNAs targeting ECE1, Sp100 or control guides. All the cell lines were made both in the presence or absence of raltegravir, which inhibited integration in order to limit off target effects. Five different guides were used for ECE1, and the level of knockdown was assessed by immunoblot (Figure 5.15). The control and Sp100 cells were produced together with Dr Katie Nightingale and Dr Ben Ravenhill (Weekes Lab), and knockout of Sp100 was also assessed by immunoblot (Nightingale *et al*, 2018).

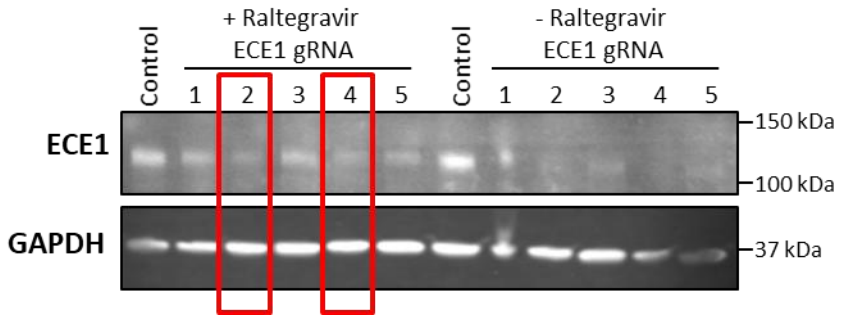


Figure 5.15 Immunoblot showing knockdown in polyclonal ECE1 CRISPR populations

Immunoblot of polyclonal ECE1 CRISPR cell populations using each of the five guides is shown for both the raltegravir treated and the untreated, integrated, cell lines. Guides 2 and 4 from the raltegravir treated cell lines were chosen for single cell cloning. This immunoblot was performed by Dr Ben Ravenhill (Weekes Lab).

All five of the integrated cell lines were tested in a restriction assay (Figure 5.16A). Guides 2 and 4, which showed the best knockdown in the polyclonal populations, also resulted in the greatest increase in infection. These were therefore used in a second restriction assay with the raltegravir treated cell lines, and subsequently for single cell cloning (Figure 5.16). In the restriction assay on the raltegravir treated cell lines, ECE1 appeared to restrict infection compared to the control_g1 cell line, however this showed

a significantly different percent infection to control_g2, so it was difficult to draw conclusions.

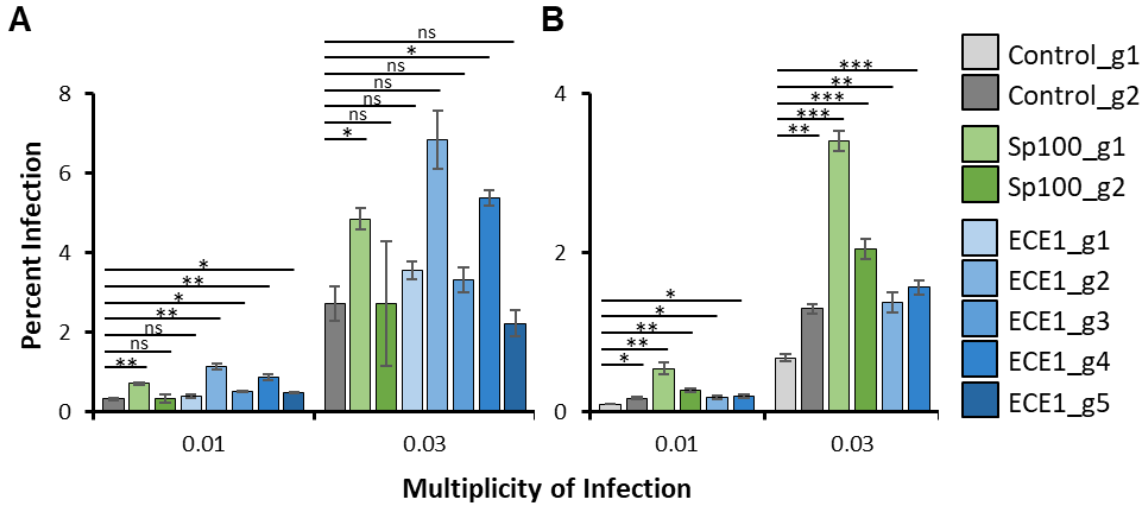


Figure 5.16 HCMV restriction assay on ECE1 polyclonal CRISPR populations

Five ECE1 CRISPR/Cas9 knockout cell lines, along with Sp100 knockout cell lines, were infected with UL36-GFP HCMV at an MOI of 0.01 and 0.03 for 24 h, and the percent infection measured by flow cytometry. Data is represented as the mean of triplicate data ± SEM. P- values were estimated using a two-tailed t-test, with each sample compared to control_g1; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ns = not significant.

(A) Five different ECE1 knockout cell lines, along with two Sp100 knockout cell lines were examined. Cells were generated in the absence of raltegravir.

(B) All cell lines were produced in the presence of raltegarvir, to prevent integration. Two ECE1 knockout cell lines (sg2, sg4) and two Sp100 knockout cell lines were examined.

Restriction Assay on Monoclonal CRISPR Populations

The raltegravir treated (not integrated) CRISPR cell lines with guides 2 and 4 were chosen for single cell cloning, due to having the greatest knockdown and largest effect on percent infection in the restriction assay. The cells were seeded at a density of 0.5 cells/well in a 96-well plate, so monoclonal populations could grow from a single cell. The resulting monoclonal populations were examined by immunoblot to identify those with an ECE1 knockout (Figure 5.17A) and five cell lines were selected (Figure 5.17B). Interestingly, four of these retained a band just under the expected ECE1 band, whilst ECE1_g4_G9 was missing this band.

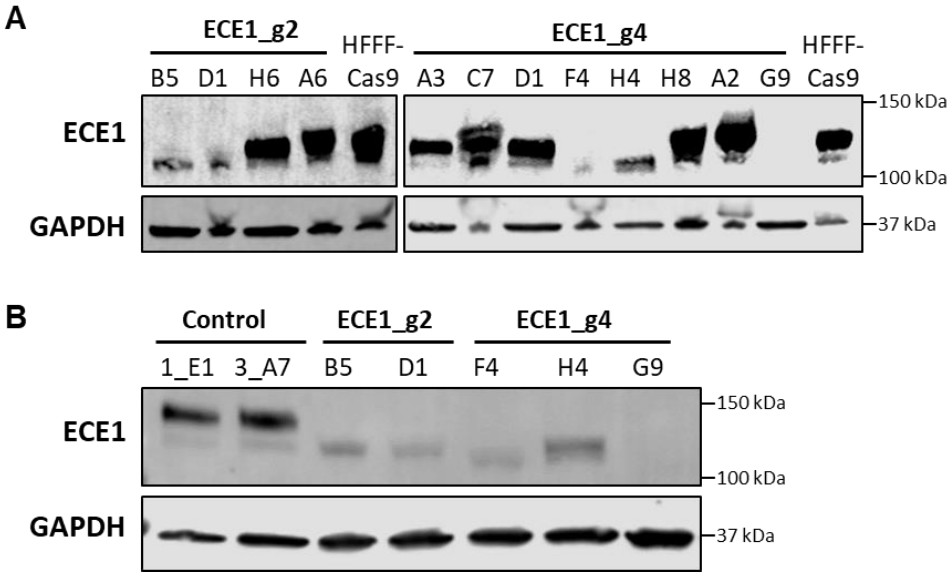


Figure 5.17 Immunoblot for ECE1 in monoclonal CRISPR populations

(A) Immunoblot of monoclonal CRISPR populations expressing ECE1 guides 2 and 4, compared to HFFF-TERT-Cas9 cell lines prior to the transduction of the gRNAs.

(B) Immunoblot of selected ECE1 knockout monoclonal cell lines from (A) compared to monoclonal populations transduced with control CRISPR guides.

Restriction assays were performed on the monoclonal cell lines identified in Figure 5.17B. The percent infection was most substantially increased in the ECE1_g4_G9, with mixed results from the other guides (Figure 5.18). It is interesting that G9 was the only monoclonal cell line where all bands on the immunoblot for ECE1 appear to be missing, and it would therefore be important for further investigation to determine what is being detected in the two bands. Furthermore, the two monoclonal control cell lines resulted in significant differences in infection in two of the three assays. It was therefore not possible to know what the expected level of infection should be, and what the ECE1 knockouts should be compared to.

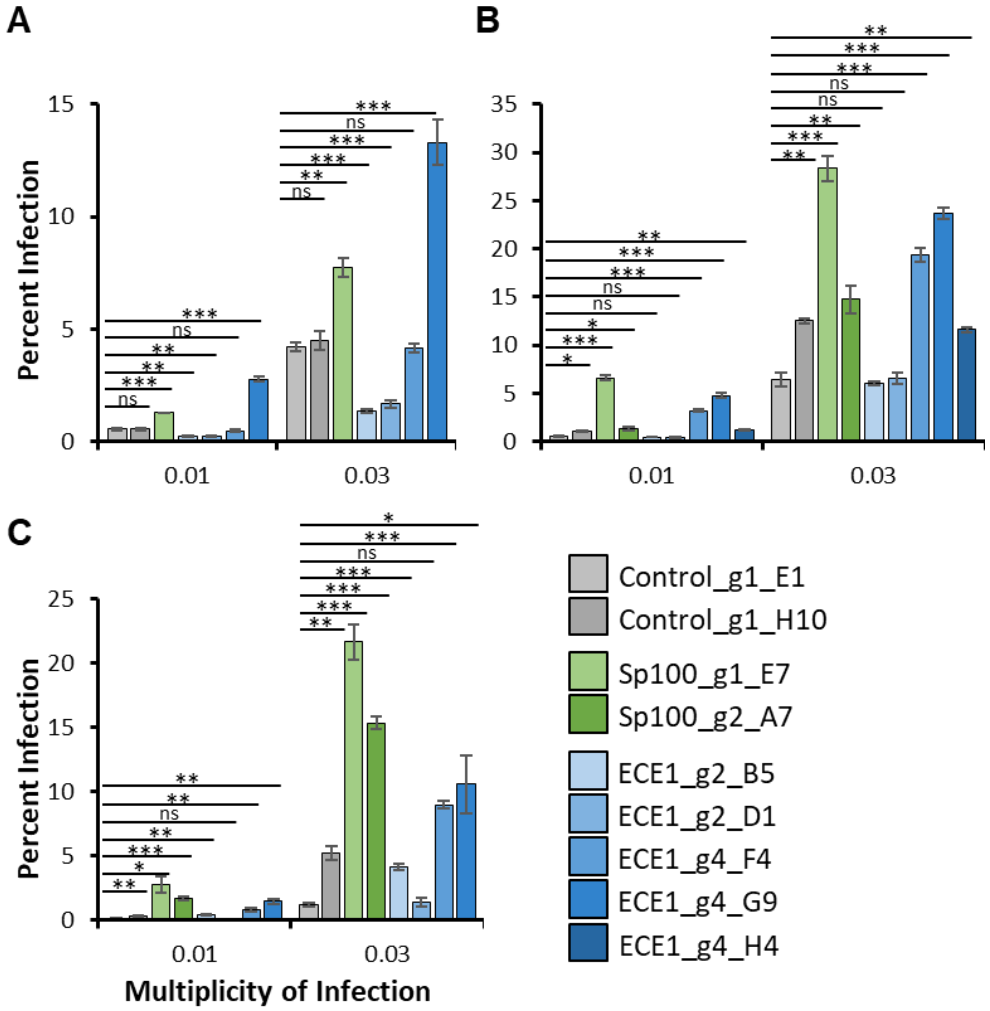


Figure 5.18 HCMV restriction assay on monoclonal ECE1 CRISPR populations

Monoclonal populations were produced from Control_g1, Sp100_g1, ECE1_g2 and ECE1_g4 polyclonal cell lines, all of which were transduced in the presence of raltegravir to prevent integration. They were infected with UL36-GFP HCMV at an MOI of 0.01 and 0.03 for 24 h, and the percent infection measured by flow cytometry. Data is represented as the mean of triplicate data ± SEM. P-values were estimated using a two-tailed t-test, with each sample compared to control_g1_E1; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ns = not significant.

(A) The two control cell lines were compared to Sp100_g1_E7, ECE1_g2_B5 and D1, and ECE1_g4_F4 and G9.

(B) The two control cell lines were compared to the two Sp100 and five ECE1 CRISPR knockout cell lines described.

(C) The two control cell lines were compared to the two Sp100 CRISPR cell lines, ECE1 sg2_B5 and D1, and ECE1 sg4_F4 and G9.

Comparison of Monoclonal CRISPR Control Cell Lines

As the percent infection differed between the control cell lines, further investigation was necessary to determine whether the variability was reproduced in other controls. Multiple monoclonal populations expressing the control CRISPR guides were therefore tested in a restriction assay (Figure 5.19). Four monoclonal cell lines were established from the single cell cloning of both the control_g1 and control_g3 populations, and a range of levels of infection were observed for these in the HCMV restriction assay (Figure 5.19A). In each case, the bar displays the mean of triplicate data for the given cell line, and error bars show the SEM. Each cell line was trypsinised, counted and seeded from a single flask.

To ensure there was no effect of inconsistent cell confluency due to differential counting and seeding, a repeat of the experiment included duplicates of some of the cell lines (Figure 5.19B). In these case three wells were seeded from each of two flasks of the stated cell line, which had been counted and seeded separately. In this way, the duplicate flasks were effectively treated as if they were for two different monoclonal control populations, so any effect of seeding could be examined. The percent infection observed between these pairs was very similar, suggesting the differences in infection were not simply due to seeding. This does not however, rule out difference in confluency due to different growth patterns. Additionally, the pattern of infection was similar between the two experiments, with control_g3_C9 resulting in the greatest level of infection, and control_g1_H3 and control_g1_H10 giving the least infection. A WT HFFF-TERT sample was also included in the second experiment, and gave a percent infection roughly in the middle of the various CRISPR cells. It is possible that off target effects of the control CRISPR guide may lead to knockouts of important antiviral genes, or that changes in growth patterns lead to differences in cell seeding between cell lines. As it was not possible to determine with confidence which of the control cell lines the ECE1 knockouts should be compared to, this approach was not used further.

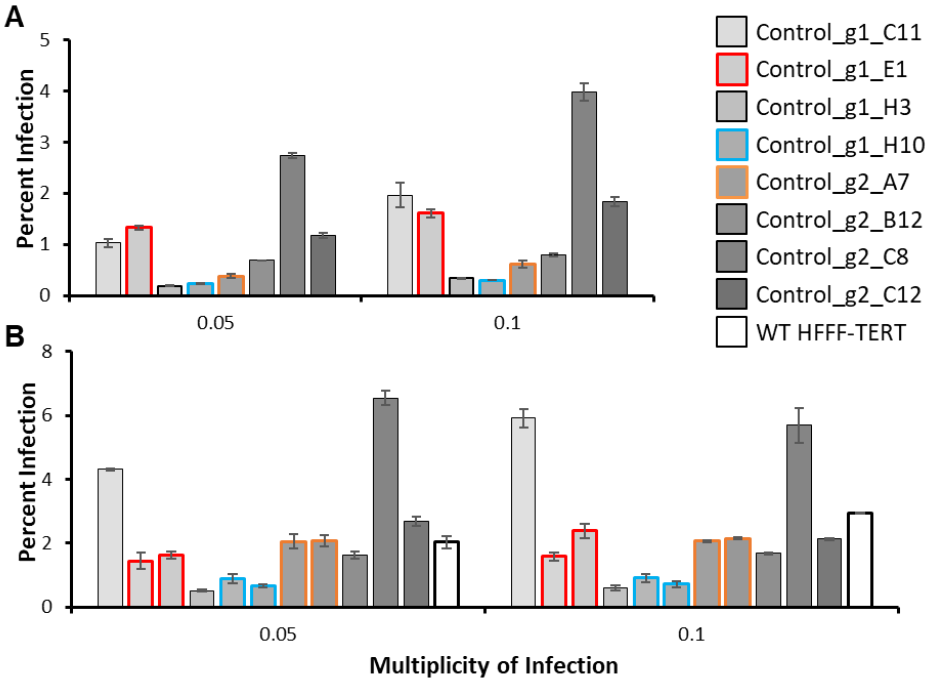


Figure 5.19 HCMV restriction assay comparing percent infection between different monoclonal CRISPR control guide populations

Monoclonal populations were produced from control_g1 and control_g2 polyclonal cell lines, all of which were transduced in the presence of raltegravir to prevent integration. They were infected with UL36-GFP HCMV at an MOI of 0.05 and 0.1 for 24 h, and the percentage infection measured by flow cytometry.

(A) Bars represent the average of triplicate data ± SEM.

(B) Bars represent the average of duplicate data, and the error bars display the range. Some cell lines have been included twice, counted and seeded separately each time, to examine the effects of cell seeding.

5.6.5 Two Colour Restriction Assays

The Two Colour Restriction Assay Methodology

Although the morphology and growth of the knockdown and overexpression cells did not obviously differ from control cells, there was some concern that cell seeding or small differences in growth of the different cell lines could affect the way the cells settle down in the well and their confluency prior to infection. Identical cell seeding is required between the control and test cells to ensure equal infection. Minor changes could affect the results of assays, both due to differences in the effective MOI if the same amount of virus is applied, and because cells are less readily infected at higher confluency.

In order to circumvent this potential issue, a 'two colour' restriction assay was developed. This system was developed and validated predominantly by Alice Fletcher-Etherington (Weekes Lab). She first produced HFFF-TERTs expressing either mCherry or iRFP. The mCherry cells were subsequently transduced with the control plasmid (control sh or GAW), whilst iRFP cells were transduced with an shRNA or overexpression plasmid for the protein of interest (ECE1, Sp100 or Daxx). The appropriate mCherry and iRFP cells were mixed in a 1:1 ratio and seeded in triplicate for each MOI. The cells were then infected for 24 h with UL36-GFP HCMV before harvesting and analysis by flow cytometry, as with the previous restriction assay. Each well therefore contained a mixed population of cells, and flow cytometry could determine the percent of GFP positive (and therefore infected) cells within both the mCherry and iRFP positive populations in each well. A 'restriction ratio' could then be calculated, comparing the percent infection in each population (Figure 5.20).

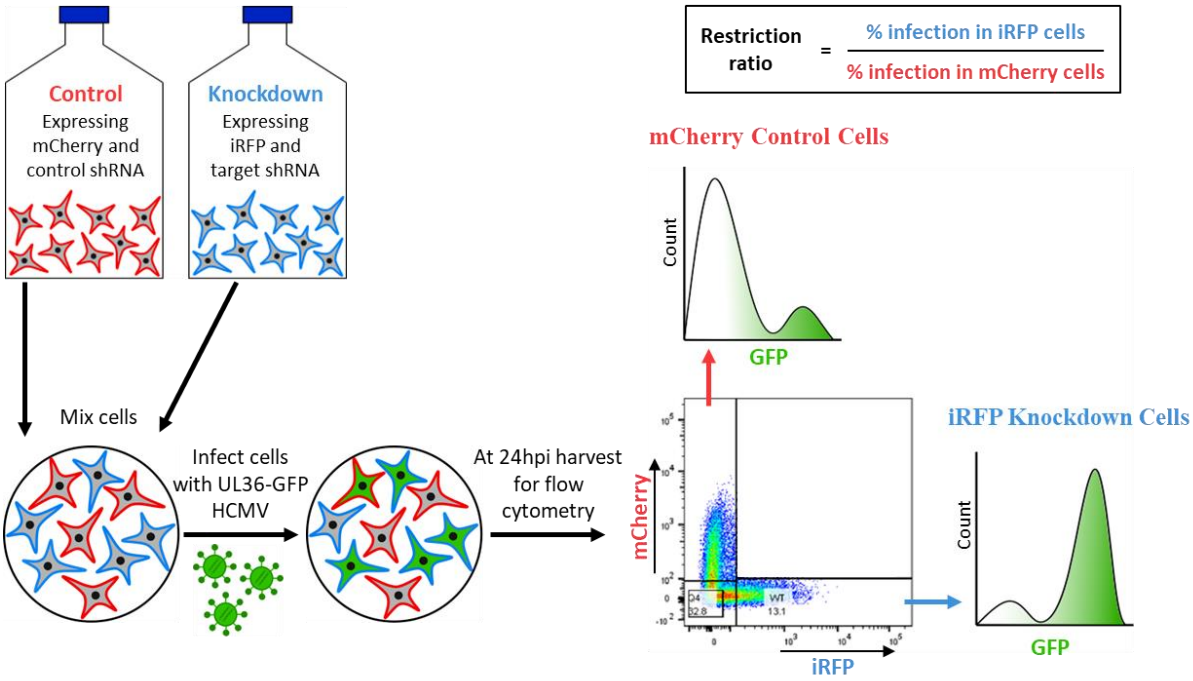


Figure 5.20 Schematic of two colour restriction assay for shRNA knockdown cells

Control cells expressing mCherry and a control shRNA were mixed in a 1:1 ratio with iRFP cells co-expressing the target shRNA. The mixed population was seeded in triplicate for each cell type, and infected with UL36-GFP HCMV, and infection proceeded for 24 h. The cells were then harvested and the % GFP+ cells in the mCherry population and in the iRFP population in each well was measured by flow cytometry.

Cell Lines

Cells expressing mCherry and iRFP were kindly provided by Alice Fletcher-Etherington (Weekes Lab). The plasmids for expression of mCherry and iRFP included a blasticidin resistance gene, enabling selection of cells expressing each fluorescent marker. They could then be transduced with constructs encoding an shRNA for the protein of interest or an overexpression plasmid, and subjected to a second round of antibiotic selection using hygromycin or puromycin respectively. Knockdown and overexpression of ECE1 was confirmed by immunoblot (Figure 5.21). Sp100 and Daxx cells were produced by Dr Katie Nightingale (Weekes Lab) and validated by immunoblot.

shRNAs were employed for production of knockdown cells due to the concerns regarding different levels of infection in the CRISPR control cells, as well as the technical difficulties of generating single cell CRISPR clones within this two colour system. Use of early passage cells ensured there would not be a loss of the knockdown affecting the results of the assays.

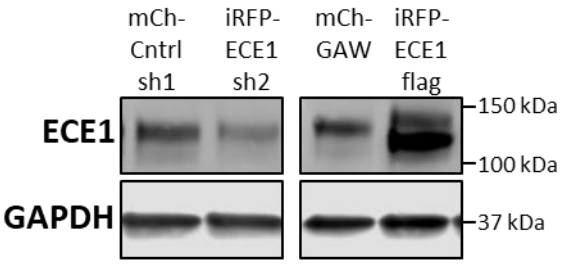


Figure 5.21 ECE1 cell lines for HCMV restriction assay in two colour system

Immunoblot confirming ECE1 knockdown and overexpression in iRFP cells compared to mCherry control cells expressing a control Sh or GAW plasmid.

Two Colour Restriction Assay: shRNA Knockdown

Four restriction assays were performed on the ECE1 knockdown cells using the two colour system (Figure 5.22). A restriction ratio greater than one (represented by the dotted line) would indicate increased infection with HCMV and therefore restrictive effects. Unlike in the previous shRNA based restriction assay, the Sp100_sh1 did work as a reliable positive control in all instances, showing restriction to various degrees.

Substantial restriction by ECE1 was observed in Figure 5.22C and D, similar in level to that observed with the Sp100_sh1 positive control; some of the lowest MOIs were included in these two assays. Restriction was also observed to a lesser extent in Figure 5.22B, however it was not seen in the first assay. The iRFP only control was included in Figure 5.22D to ensure that there was no difference in infection as a result of the cells expressing iRFP rather than mCherry. There was not much of a difference, though possibly a slight reduction in infection with the iRFP only cells, suggesting restrictive effects may be slightly compressed.

Two Colour Restriction Assay: Overexpression

Three assays were performed considering the effect of ECE1 overexpression, including Daxx as a positive control (Figure 5.23). In this case, a restriction ratio of less than one (represented by the dotted line) would indicate restriction, as overexpression of an ARF should reduce infection. ECE1 did not show any restrictive effects in these assays, in contrast to the results observed in the single colour overexpression restriction assays (Figure 5.11). In these two colour restriction assays, overexpression of ECE1 actually led to increased HCMV infection. The assay appears to technically work as Daxx shows a high level of restriction. However, the iRFP only control in Figure 5.23C also led to increased infection, to a similar extent to that seen with ECE1. The iRFP cells should show the same infection as mCherry control cells (expressing GAW); this result suggests there may be a difference between the mCherry and iRFP expressing cells in this instance, or an effect of the GAW plasmid.

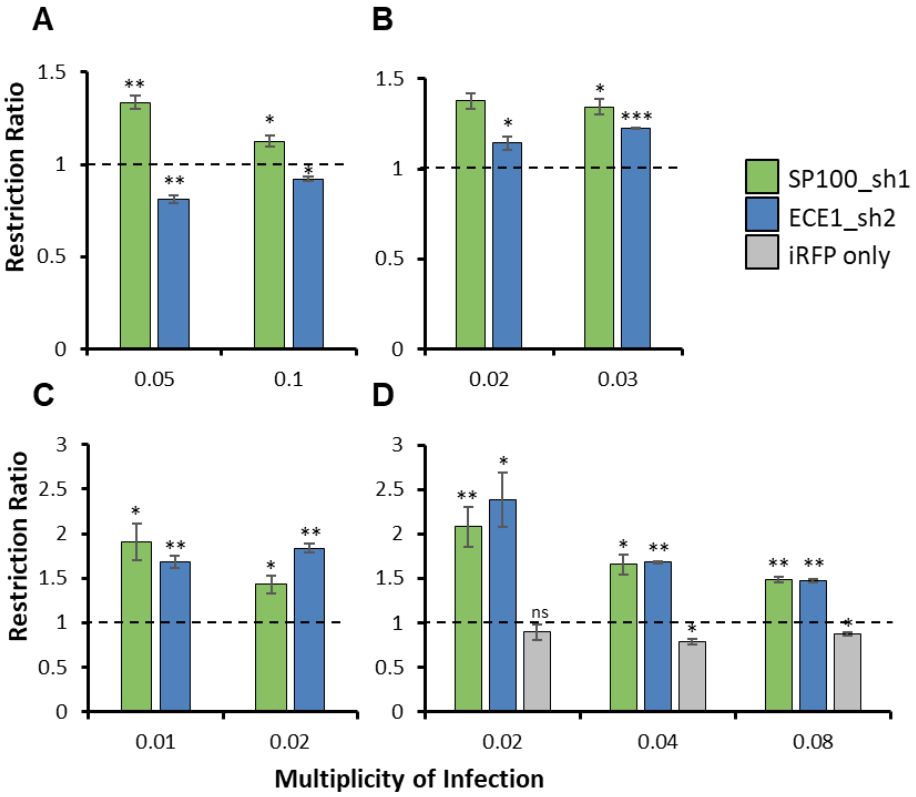


Figure 5.22 Two colour restriction assays on ECE1 shRNA knockdown cells

ECE1 restriction was examined using the two colours restriction assay. HFFF-TERTs expressing mCherry and control_sh1 were mixed in a 1:1 ratio with cells expressing iRFP and an shRNA targeting ECE1 or Sp100. The cells were infected with HCMV UL36-GFP for 24 h at the indicated MOI (0.01 - 0.1), and then the percent of GFP positive cells in both the mCherry and iRFP population in each well was measured by flow cytometry. A restriction ratio was calculated as described in Figure 5.20. Cell types were examined in triplicate except where specified, and the bars display the mean of the restriction ratios in the three wells ± SEM. The dotted line marks a restriction ratio of 1, equivalent to observing the same percent infection in the control and knockdown cells, and therefore no restriction; a restriction ratio > 1 indicates restriction. P-values were estimated using a paired two-tailed t-test, comparing the percent GFP+ cells in the iRFP population to the percent GFP+ cells in the mCherry population for each cell type. *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ns = not significant.

(A) Bars represent the mean of triplicate data ± SEM in all cases.

(B) Bars represent the mean of triplicate data ± SEM, with the exception of Sp100 at an MOI of 0.02, for which there were only two data points. For this sample, error bars display the range, and no p-value was determined.

(C) Bars represent the mean of triplicate data ± SEM in all cases.

(D) Bars represent the mean of triplicate data ± SEM in all cases. An additional 'iRFP only' sample was included, where the cells did not express any shRNA, to investigate effects of mCherry compared to iRFP expression.

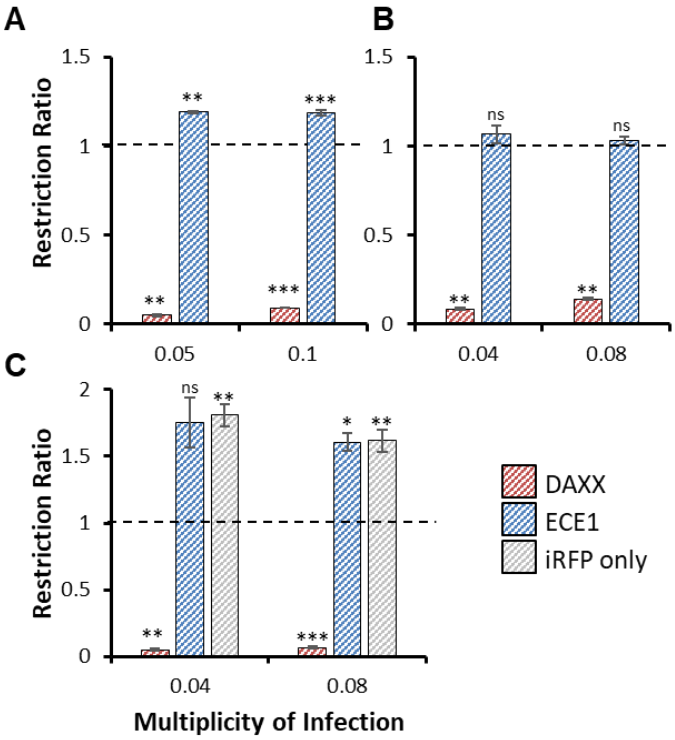


Figure 5.23 Two colour restriction assay on ECE1 overexpressing cells

(A-C) Each iRFP cell type (overexpressing ECE1 or Daxx) was mixed in a 1:1 ratio with mCherry GAW control cells and seeded in triplicate. Cells were infected for 24 h with HCMV UL36-GFP, and the percent of infection in the iRFP and mCherry cells in each well measured by flow cytometry. Data is represented as the mean restriction ratio (n=3) ± SEM. The dotted line marks a restriction ratio of 1, equivalent to observing the same percent infection in the control and overexpression cells, and therefore no restriction; a restriction ratio < 1 indicates restriction. P-values were estimated using a paired two tailed t-test comparing the percent GFP+ cells in the iRFP and mCherry population for each cell type. *p < 0.05, ** p < 0.01, *** p < 0.001, ns = not significant.

5.6.6 HCMV Assays in Novel ECE1 shRNA Cell Lines

Cell Lines

As the results from previous restriction assays were inconclusive, a final experiment was performed on a new set of cell lines. As the previously employed shRNA contstructs did not provide a reliable knockdown, four new shRNAs targeting ECE1 were used. Additionally, transductions for the controls, positive controls and ECE1 cell lines were

all performed in parallel on the same population of HFFF-TERTs. Immunoblot showed that of the four new shRNA expressing cell lines, ECE1_sh4, ECE1_sh5 and ECE1_sh6 all displayed a good level of knockdown. Additionally, control_sh1 seemed to have lower expression of ECE1 than control_sh2 or control_sh3 (Figure 5.24A). For this reason, ECE1_sh4, sh5 and sh6, and control_sh2 and sh3, were used in the shRNA based restriction and plaque assays. The Sp100 knockdown (Figure 5.24B) and overexpression of ECE1 and Daxx worked well (Figure 5.24A, C). These restriction assays used the 'single colour' system for simplicity and were only repeated once, in parallel with a plaque assay.

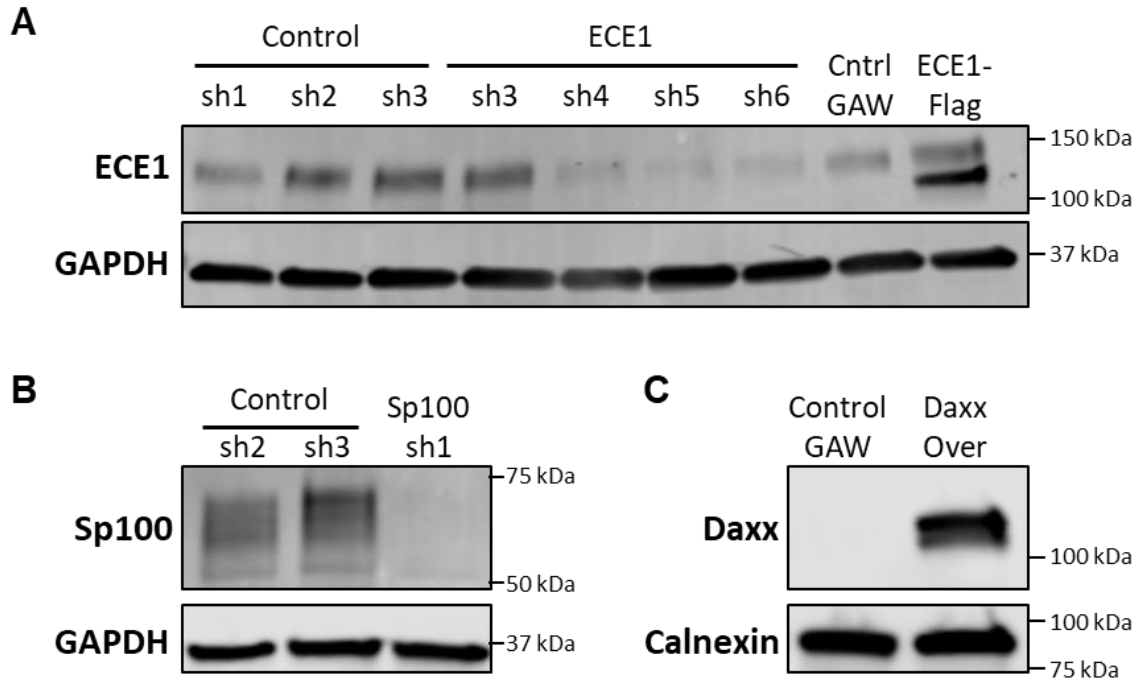


Figure 5.24 Immunoblot of shRNA and overexpressing cell lines

(A) WT HFFF-TERTs were transduced with control shRNAs or a GAW plasmid to generate control cells. Four different ECE1 shRNAs were transduced, with sh4, sh5 and sh6 producing the best knockdown. An ECE1 overexpressing line showed increased levels of ECE1 compared to GAW.

(B) WT HFFF-TERTs were transduced with either control shRNAs or an shRNA targeting Sp100, and knockdown was confirmed by immunoblot.

(C) WT HFFF-TERTs were transduced with a control GAW plasmid, or one overexpressing Daxx, and overexpression was confirmed by immunoblot.

shRNA Restriction Assay

The infections at an MOI of 0.01 resulted in generally greater infection in the ECE1 knockdown cells compared to the controls. However, all infections other than that of the Sp100 knockdown cells resulted in less than 0.15 % infection; therefore, there were too few infected cells to draw any substantial conclusions. The 0.03 MOI resulted in 0.8 and 1.3 % infection in control cells, whilst ECE1 knockdowns had between 1.0 and 2.2 % infection. ECE1 sh5 and sh6 showed 2 fold and 1.6 fold more infection in the knockdown compared to the sh3 control (and a greater FC than this when compared to the sh1 control). This does suggest some level of restriction, however it was far less than observed with Sp100, which worked as a strong positive control in this assay (Figure 5.25).

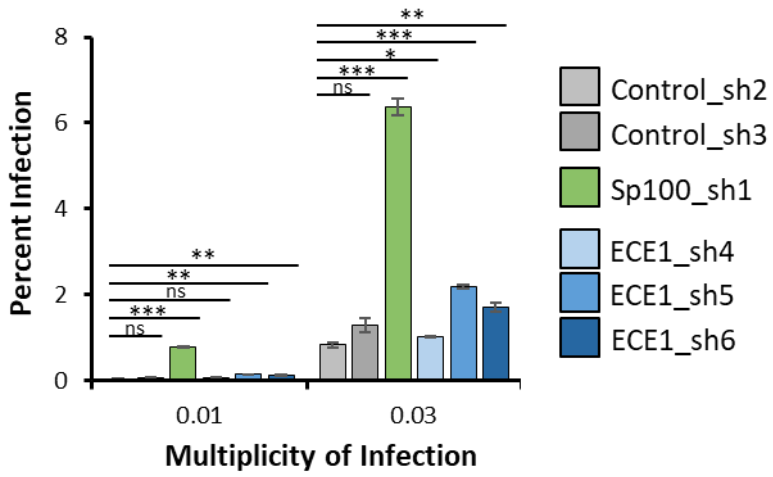


Figure 5.25 Repeat HCMV restriction assay on shRNA cell lines

A new set of control, Sp100 and ECE1 shRNA cell lines were produced. They were infected with UL36-GFP HCMV for 24 h prior to harvesting, and the percent infection was assessed by flow cytometry. Data is presented as the mean ± SEM (n=3). P-values were estimated using a two-tailed t-test, and were compared to the percent infection in the control_sh2 cell line; *p < 0.05, ** p < 0.01, *** p < 0.001, ns = not significant.

Overexpression Restriction Assay

The overexpression restriction assay showed no significant effect of ECE1 overexpression. Although it should be noted that Daxx had a far smaller effect in this assay compared to the previous overexpression experiments, though still significant (Figure 5.26).

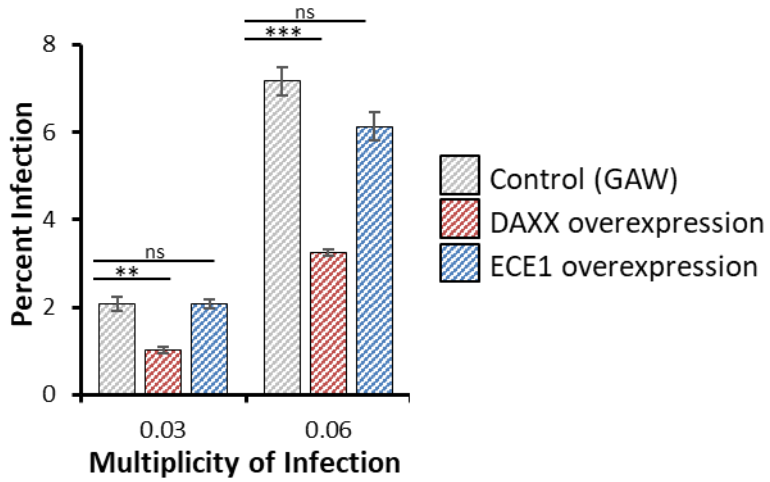


Figure 5.26 Repeat HCMV restriction assay on overexpression cell lines

A new set of control, Daxx and ECE1 overexpressing cell lines were produced. They were infected with UL36-GFP HCMV for 24 h prior to harvesting and analysis by flow cytometry. Data is presented as the mean ± SEM (n=3). P-values were estimated using a two-tailed t-test, and were compared to the percent infection in the control GAW cell line; *p < 0.05, ** p < 0.01, *** p < 0.001, ns = not significant.

Plaque Assay

A plaque assay was also performed in parallel with the restriction assays. The UL36-GFP restriction assay only considers any restriction of the virus in the first stages of a single round of infection. In contrast, the plaque assay examines any effect over multiple replication cycles and on cell-to-cell spread. The number of plaques and the size of the plaques can be observed.

At 24 h after seeding, the cells were infected with AD169-GFP HCMV. This strain of the virus was used as infection with the Merlin strain of HCMV (as used in the restriction assays) does not form plaques that are clearly distinct and easy to count. Cells were

infected with virus at multiple dilutions in order to obtain wells with a suitable number of plaques to count. Infection was allowed to proceed on a rocker for 2 h at 37ºC, and then the inoculum was replaced with a 1:1 mix of 2x DMEM culture media, and Avicel. After two weeks at 37ºC, the Avicel was removed, the wells fixed, and the fluorescent plaques examined and counted under the microscope. Of the four MOIs tested, an MOI of 0.02 and 0.06 resulted in too many plaques to allow accurate counting, but an MOI of 0.002 or 0.006 generated between 5-100 plaques/well.

The plaque assay examining ECE1 and Daxx overexpressing cells provided remarkably similar results to the restriction assay on the same cells (Figure 5.26, Figure 5.27B). Again, Daxx significantly reduced the number of plaques observed, and ECE1 had some effect though to a much lesser extent.

The plaque assay on the shRNA cells showed quite different results (Figure 5.27A). There was a significant difference between the number of plaques in the two control cell lines, and the Sp100 and ECE1 knockdowns led to fewer plaques than the control_sh2. This is the opposite of what would be expected for a restriction factor, which was particularly surprising for the positive control Sp100, and is in contrast to previously published data (Adler *et al*, 2011). It is possible that the control_sh3 is a better baseline, showing a more similar number of plaques to the GAW control. However, even if the other cell lines are compared to this control, not much change was detected apart from a large reduction in the number of plaques with ECE1_sh4. Without any replicates of the experiment it is not possible to determine if these effects are reproducible, or indeed an effect of differences in cell seeding. As the cells remained in culture for two weeks, this could have a substantial impact.

Another interesting aspect of the plaque assay data could be the size of the plaques, indicating a difference in viral spread. In this instance there was no obvious difference in plaque size by eye, therefore the plaques were not imaged and measured. By eye, the plaques formed by the Sp100 knockdown cells appeared slightly smaller than others. This was again the opposite of what might be expected, though has previously been observed by other members of the Weekes lab. Along with no increase in the number of plaques, this may indicate an artefactual effect of the cell line employed and therefore represents a significant caveat in this experiment.

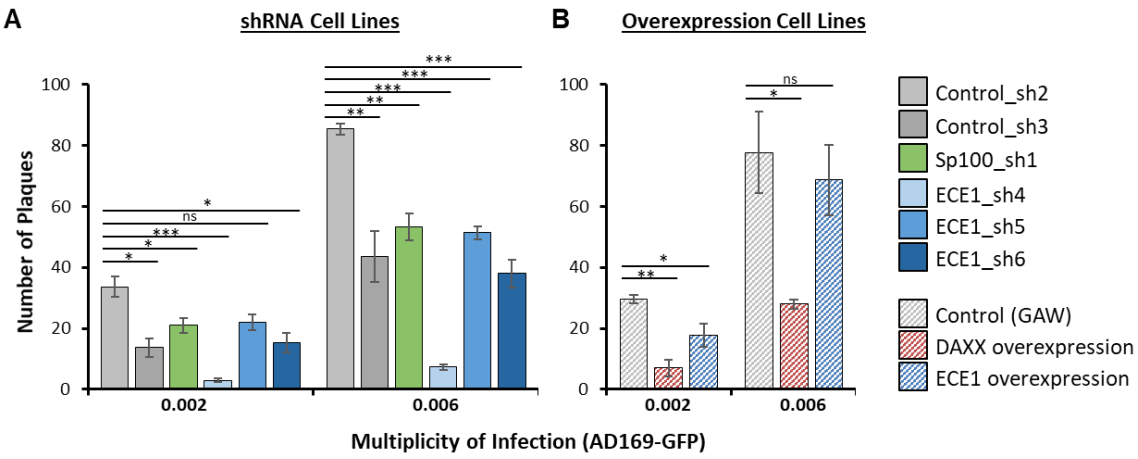


Figure 5.27 Plaque assay on ECE1 shRNA and overexpression cells

Cells were infected with AD169-GFP for 2 h, prior to replacing the inoculum with 1:1 media and Avicel. Plaques were counted after two weeks. Data is presented as the mean ± SEM (n=3). P-values were estimated using a two-tailed t-test; *p < 0.05, ** p < 0.01, *** p < 0.001, ns = not significant.

(A) ECE1 shRNA cells were infected, along with control and Sp100 cell lines. T-tests compared each cell line to control_sh2.

(B) ECE1 overexpressing cells were infected, along with the GAW control and Daxx cell lines. T-tests compared each cell line to GAW cells.

5.7 Summary of ECE1 Results

Table 5.4 Table summarising results of ECE1 restriction and plaque assays

Experiment	ECE1 restriction?	Summary of Results	Effectiveness of Controls
Overexpression	Yes	All assays showed some level of restriction by ECE1, though to a lesser extent than Daxx	Worked well. Greater effect of Daxx than Sp100
shRNA Knockdown	Mixed results	One of the three assays showed ECE1 restricting infection	Sp100 did not work as an effective control
CRISPR Polyclonal Population	Mixed Results	The two polyclonal populations with the best knockdown showed restriction in integrated cell lines, but not the raltegravir treated ones	One of the Sp100 guides worked well
CRISPR Monoclonal Population	Yes?	Restriction by some cell lines, (particularly in 'cleanest' knockout) but interpretation of this depends on what is detected in the two bands on the immunoblot	Sp100 control worked well, however significant differences in the percent infection was observed between control cell lines
shRNA Knockdown Two Colour	Yes	ECE1 knockdown led to a small increase in infection	Sp100 control worked well
Overpxression Two Colour	No	ECE1 overexpression led to a small increase in infection, not decrease as expected	Daxx control worked well. 'iRFP only' led to a small increase in infection
shRNA Knockdown Repeat sh4-6	Small effect	ECE1 knockdown led to a small increase in infection, though much less than Sp100	Sp100 control worked very well
Overexpression Repeat	No	ECE1 overexpression led to a small, not significant decrease in infection	Daxx control worked, though less substantial effect than seen previously
shRNA Plaque Assay	No	ECE1 knockdown led to reduced percent infection	Sp100 knockdown also led to reduced percent infection
Overexpression Plaque Assay	No	ECE1 led to a small, not significant decrease in infection	Daxx control worked well

5.8 Discussion

A proteomic screen was used to identify changes at the surface of primary monocytes and T cells upon IFN stimulation. The resulting data was overlapped with proteomic analyses of viral infection, in order to identify candidate restriction factors which were both stimulated by IFN and downregulated by viral factors. TMEM123 may present an interesting candidate for a restriction factor. If it does have a role in triggering oncotic cell death (Ma *et al*, 2001), it may be beneficial for the virus to inhibit this. However, due to a lack of reagents this was not further investigated. ECE1 was identified as a candidate restriction factor.

5.8.1 How Does HCMV Downregulate ECE1?

ECE1 was IFN stimulated at the surface of primary T cells in all five donors, and was also downregulated by HCMV and to a lesser extent HIV. Fortunately, previous literature was available to demonstrate specific downregulation of ECE1 by the HCMV US2 protein, answering the first question about how the virus targets ECE1. Infection with a HCMV US2 deletion mutant demonstrated that US2 is necessary for the downregulation of ECE1. Other US2 targets include MHC molecules, integrins and members of the immunoglobulin superfamily. Targeting of the integrins as well as some other US2 targets was found to be dependent on the cellular E3 ligase TRC8, with shRNA knockdown of TRC8 diminishing the downregulation. HCMV US2 recruits TRC8 to these proteins, resulting in their ubiquitination and proteasomal degradation (Hsu *et al*, 2015). Immunoblot demonstrated that single-gene expression of US2 was also sufficient to induce the downregulation of ECE1. It would be interesting to examine the expression of ECE1 during infection of TRC8 knockout cells to determine if the same mechanism is responsible here.

5.8.2 Does ECE1 Restrict Viral Infection?

Despite fulfilling the criteria for a candidate restriction factor, it proved challenging to obtain a conclusive result on the role, if any, of ECE1 in HCMV infection (Table 5.4).

Whilst shRNA knockdown of ECE1 led to a significant increase in infection in many of the restriction assays, the effect was not consistent. Overexpression of ECE1 limited infection in the WT HFFF-TERTs providing initially promising results, but did not show this in the two colour system. In the additional repeat of the restriction assay on newly generated cells, ECE1 overexpression did not show a substantial effect. Caution should be exercised in interpreting the results of the overexpression assays, as overexpression was of ECE-1b, a predominantly intracellular isoform. ECE1 can also form heterodimers, and in this case the localisation is dominated by isoform B. As there were no antibodies which definitely worked for cell surface flow cytometry, overexpression of ECE1 at the cell surface was not tested, but would be important to know for drawing more definitive conclusions from this data.

Some of the initial shRNA knockdown experiments showed an increase in infection with the ECE1 shRNA cell lines, and three of the four two colour shRNA assays showed equivalent restriction by ECE1 and Sp100, though the effect of Sp100 was not as big as might be expected of an ARF. However, there was no effect seen in some of the other experiments, and it varied between the two shRNA cell lines. In the initial experiments, this may have been due to a reduced level of knockdown from passaging of the cells. However, subsequently produced knockdowns for the final experiment presented were frozen when the lysates were produced, and used rapidly after thawing, so should not have had this complication.

Unfortunately no reliable results could be obtained from the monoclonal CRISPR system due to differences in the percent infection between control cell lines, however ECE1_g4_G9 and in some cases ECE1_g4_F4 did show increased infection. If further experiment were to be performed, it may be beneficial to increase the number of control cell lines used. The monoclonal CRISPR assay and the final restriction and plaque assays presented included multiple controls, and showed there were often differences in the infection of these cells. This was a particularly obvious effect in the CRISPR cells. This could be the result of different off-target effects between the monoclonal populations. CRISPR is known to result in some off-target effects, with up to five mismatches being tolerated in order to generate a double strand break (Fu *et al,* 2013). Another possibility is that because the HFFF-TERTs transduced with Cas9 were not single cell cloned prior to transduction with the guides, the parent population of cells was not homogeneous, and

may have resulted in differential effects of the CRISPR guides. Perhaps proteomics experiments could be used to investigate any substantial differences between the cells that give the highest and lowest rates of infection, though most likely there will be many differences and discerning the important proteomic changes could be challenging.

Despite this variation however, when the positive controls worked effectively, a seven fold increase in infection was seen with the Sp100 knockout (Figure 5.25), and a > 10 fold decrease was seen with Daxx (Figure 5.11). ECE1 did not reliably show effects of this magnitude, and had it done so, some variation in the controls would have been less critical. This is not to say it has no role in infection, but it is unlikely to be a major restriction factor acting early during HCMV infection, or if it is it may act in concert with other proteins. To establish a more conclusive result from these restriction assays, another possible experiment would be to complement the CRISPR knockouts with ECE1, in particular the ECE1_g4_G9 clone. This clone showed the cleanest knockout by immunoblot, and a substantial increase in infection in two of the three experiments. The Abcam antibody for ECE1 often detected two bands, and this was the only clone for which both were missing. The two bands have been seen previously with a different antibody (Lambert *et al*, 2008). In the Lambert *et al* study, a targeting siRNA only appears to reduce the upper band. However, overexpression of ECE1 also increased the signal in the lower band, as was observed in experiments here (Figure 5.21, Figure 5.24). It would therefore be important to sequence the clones to determine the effect of the CRISPR knockouts, and perhaps screen additional monoclonal populations for more complete knockout cell lines.

It may be valuable to repeat these assays with a HCMV US2 deletion mutant, which is unable to degrade ECE1; in the current assays, infection with wild type virus may result in a comparable level of ECE1 in the knockdown and control cells. The deletion mutant would therefore increase the window for observing a restrictive effect. It is also possible that ECE1 acts at a later point in infection than would be detected with the UL36-GFP restriction assay. The HCMV tegument protein UL32 is expressed later during infection, and a UL32-GFP virus is also available, so a restriction assay with this may be more informative if ECE1 acts later during infection. The results of the plaque assay would cover more of the viral replication cycle, and also did not show an obvious phenotype relating to a restricting effect. However, this was a single replicate of the experiment,

and Sp100 did not work as an effective positive control. To fully establish that there was no effect in the plaque assay, it would be necessary to repeat it and identify more consistent control cell lines. In both the restriction assay and the plaque assay, ECE1_sh4 showed a reduced percent infection and number of plaques, whilst this was increased with ECE1_sh5. The consistency between the two different types of assay, and the different viruses (UL36-GFP Merlin Strain HCMV and AD169-GFP) may be interesting, however this could also be an artefact of cell seeding, as cells for both assays were seeded together.

As well as downregulation by HCMV, ECE1 was also downregulated to a lesser extent at the surface of HIV infected cells (Matheson *et al*, 2015). With CD4+ T cells being both a major site of HIV infection, and the cell type ECE1 IFN stimulation was observed in, a HIV restriction assay may be relevant. Assays to investigate this could involve infecting cells with ΔEnv-GFP-HIV and using flow cytometry to assess the percent infection, as with HCMV. This would examine any restrictive effects preventing viral entry or integration. Additionally, the supernatants can be harvested and the amount of p24, a HIV capsid protein, measured by enzyme-linked immunosorbent assay (ELIZA).

5.8.3 Other Considerations from the Data

Although ECE1 was selected for follow up, other proteins in the dataset may be candidate restriction factors. One limitation of the current overlap between the IFN and viral data, is that many of the proteins identified in the primary monocytes were not quantified in the viral datasets. Of the 41 IFN-stimulated proteins, just 15 were quantified in any of the HCMV infection datasets, as these were in fibroblast cell lines. In addition to the experiments presented in this thesis, I have been using multiplexed proteomics to examine HCMV infection of primary DCs in collaboration with Dr Richard Stanton (Cardiff University), both at the whole cell and PM level. This dataset may provide a better comparator for the monocyte data, particularly if there are leukocyte specific proteins restricting infection. Additionally, the investigation presented here shortlisted IFN-stimulated proteins based on reasonably stringent criteria, requiring the proteins to be upregulated in all five donors, despite evidence that differences between donors may be expected. Therefore, some interesting candidate restriction factors may have been filtered out.

CHAPTER 6: DISCUSSION AND FUTURE DIRECTIONS

6.1 IFN Screen

6.1.1 Summary

Multiplexed quantitated proteomics was used to investigate the effects of IFNα2a stimulation on the cell surface of primary CD14+ monocytes and CD4+ T cells, across five donors. This was complemented with the study of a cultured monocyte cell line, THP-1s, and pan-monocyte populations from two further donors. This resulted in the quantification of 607 PM proteins in monocytes, and 584 in T cells, which could be used to examine the cell surface proteome, and to investigate the IFN-induced changes.

Methodology based on iBAQ was used to determine the abundance of different proteins at the cell surface. In both monocytes and T cells, a large proportion of the surface proteome consisted of a relatively small number of proteins, with some commonalities between the cell types. Additionally, the cell surface proteomes were highly similar between donors.

To examine IFN stimulation, criteria were determined for identifying proteins that were consistently IFN stimulated across multiple donors. In the monocytes, 41 proteins met these criteria, and the observation of similar proteomic changes in the pan-monocyte populations provided confidence in the findings. Just 13 proteins were consistently IFN stimulated in T cells. There was also a large amount of variability in the IFN response between donors, as has been observed in the literature. Comparison of the IFN stimulated genes in both cell types identified TMEM123 as the only protein upregulated in all cell types, other than HLAs and the known restriction factor tetherin. IFN stimulation of TMEM123 was validated in THP-1 cells by RT-qPCR, but unfortunately no antibodies have yet been identified for validation at the proteomic level or for further investigations.

6.1.2 Discussion and Future Directions

This is a large dataset, with many potential avenues for further investigation. The focus in this thesis has been on identifying the most consistent results, in order to be able to provide some validation, and to identify candidate ARFs. However, there are a variety of ways in which this investigation could be extended, some of which are summarised here:

1) ***Comparison of the primary and cultured cell data –*** It would be interesting to perform a comprehensive analysis of differences in the cell surface proteome and IFN response of the primary monocytes compared to the cultured THP-1s, determining the validity of this cell line as a model for primary cells in research. The THP-1 data is currently only from a 'single shot' analysis, used for validation of the workflow and a very brief comparison. Upon fractionation, and repetition of the THP-1 experiment to ensure reliability, it may be valuable to do a more thorough comparison. Cultured cell lines thought to represent their equivalent primary cells are routinely used in laboratory research, despite observations of differences in their proteome and cytokine response (Bosshart & Heinzelmann, 2016). This approach provides a method for systematic analysis of the proteome and IFN response in these cells. Investigation of equivalent T cell lines such as Jurkats, CEM-T4 or SUPT1s in the same way would be similarly interesting. Additionally, it would assist in choosing appropriate cell lines for testing antibodies and validation of the data.

2) ***Investigating proteins which are IFN stimulated in a subset of donors -*** This investigation only considered proteins as IFN stimulated if they were upregulated to some extent in all five donors, filtering for the most consistent results. Given the variability between donors, it may be interesting to consider proteins which were only upregulated in a subset of donors, as these may still be true IFN responsive proteins. The extent of variability in the IFN response between donors was initially surprising, though upon investigation had been observed previously at the transcriptomic level (Schlaak *et al*, 2002). In order to identify true IFN stimulated proteins which are not observed in all donors, it may be necessary to have a bigger sample. It would also be challenging to validate with orthogonal

techniques, requiring assessment of multiple donors. For the purposes of this investigation, interest was predominantly in the most consistently IFN stimulated proteins, so this was not necessary, but it is worth noting that some important results may have been overlooked. A follow up experiment could involve repeating the IFN stimulation multiple times for a single donor to see if the same level of variability is observed. This would however, be challenging in the context of the monocytes, as these samples were obtained from anonymous donors through the NHSBT, so it would not be possible to obtain repeated donations. It would additionally be beneficial to repeat the experiment on the cultured cell line to ensure none of the inconsistency is generated from the experimental workflow, and does represent true variation between donors.

3) ***Investigating proteins which are not annotated as being at the PM*** **-** This dataset was filtered for proteins with PM related annotations. It is highly likely that this has led to removal of some genuine cell surface proteins from the data due to poor annotation, leading to more potentially interesting data being bypassed.

In this investigation, the data was filtered for annotated PM proteins which were consistently modulated, in order to identify the highest confidence targets. These points are therefore all given as suggestions for expanding the results rather than necessities.

Another consideration with this data, is that post-translational modifications were not examined. As PTMs apart from oxidation of methionine were not included in the database when searching the proteomic data, peptides modified upon IFN stimulation would not have been quantified in these samples. In the specific scenario where a protein was only quantified by a small number of peptides, the majority of which were modified, and the modified version of the protein accounted for a substantial portion of the protein pool, then the protein would appear downregulated. This may be relevant here as several components of the IFN response pathway, including the IFN receptor, are known to be phosphorylated upon signalling (Zheng et al., 2011). However, it is likely that only a minority of the peptides from any given protein would be modified, so this should not affect quantitation too greatly. It would be interesting to do further proteomic investigations with prior enrichment for post-translationally modified proteins,

particularly as this may modify localisation of proteins as in the case of ECE1 (Kuruppu and Smith, 2012), and this may alter abundance at the plasma membrane.

TMEM123 was the only protein other than the positive controls which was IFN stimulated in all cell types examined. It may represent a pan-leukocyte marker of IFN stimulation, and have as yet uncharacterised roles in immunity. Little is currently known about this protein, aside from its ability to induce oncotic cell death upon the application of an antibody. It may even represent another candidate restriction factor, as it is downregulated by HCMV. Unfortunately, reagents were not yet identified for validation at the proteomic level, or for further investigation of TMEM123. A first step in continuing this process would be to generate a cell line overexpressing tagged TMEM123, so the size of it could be determined, as this is variable between cell lines. This would help to identify a suitable antibody for validation and further studies. A proteomic experiment considering overexpression and knockdown of TMEM123 could yield information on its function.

6.2 VACV Screen

6.2.1 Summary

A quantitative temporal proteomic analysis of a single round of VACV infection was conducted. Samples were harvested at three mock time points, and six infection time points up to 18 hpi. The experiment was performed in triplicate, and provided robust and interesting data. The majority of changes occurred linearly, with the greatest modulation of the host proteome at 18 hpi. During infection, 265 proteins were downregulated by VACV, and these were enriched for cell surface proteins, innate immune proteins, collagens and cadherins amongst other classes; this was highly similar to classes of proteins modulated by HCMV. A further proteomic investigation examined the addition of MG132 during infection, a proteasome inhibitor, identifying that around two thirds of the downregulated proteins were specifically targeted for proteasomal degradation. This included all the canonical IFITs. From this data, HDAC5 was identified as being downregulated by both VACV and HCMV, and was proteasomally targeted by VACV

C6; this was validated by proteomics and immunoblot using a C6 deletion mutant. Further examination showed that overexpression of HDAC5 reduced viral titre, whilst knockdown resulted in increased infection with both VACV and HSV-1, suggesting HDAC5 acts to restrict viral infection.

6.2.2 Discussion and Future Directions

This dataset has provided a comprehensive and valuable resource for the poxvirus field, and has already been used to identify HDAC5 as a novel restriction factor for VACV and HSV-1, as well as C6 as its viral antagonist. It is also informative in terms of guiding future investigations. Of the 265 proteins downregulated by VACV, the majority were rescued by proteasomal inhibition. However a subset of these, including most of the collagens, were not. Samples were harvested for RNA sequencing alongside those harvested for proteomic analysis. It would therefore be interesting to determine if the collagens and other proteins not proteasomally degraded are modulated at the transcriptional level.

Enrichment analysis on the downregulated proteins highlighted terms referencing the cell surface, such as extracellular, cell attachment site and immunoglobulin. This suggests that the cell surface is a key site of modulation for VACV. Further experiments are now being conducted using PM enrichment of infected cells, so the effects on the cell surface can be more thoroughly analysed. This may highlight additional candidate restriction factors, as well as an improved dataset for identification of immune ligands. It is also being used for investigation of viral proteins present at the cell surface.

In this experiment, C6 was validated as the viral protein responsible for targeting HDAC5 for degradation. The Smith lab has a variety of deletion mutant viruses available, and further proteomic experiments have been commenced to assess the role of other viral proteins in infection. In these experiments, multiplexed proteomics has been used to analyse a series of mutants, so the host proteins targeted by each can be determined. This will complement the current dataset well in terms of identifying additional host-virus interactions.

6.3 ECE1

Overlap of the IFN screen with the VACV, HCMV and HIV proteomic screens led to the identification of ECE1 as a candidate restriction factor. The effect in HCMV infection are as yet inconclusive, with mixed results from the various restriction assays, and challenges ensuring appropriate controls. It would be interesting to investigate this further using overexpression of an isoform of ECE1 known to be localised at the cell surface. Some additional single cell CRISPR lines have been screened, and it would be useful to examine any others that have a complete knockout. There is substantial data to suggest there may be some effect of ECE1 on HCMV infection, however this was not completely reproducible, and is unlikely to be a particularly large effect. It would also be extremely interesting to investigate ECE1 in the context of HIV infection, with CD4+ T cells being the primary targets of viral infection and the cells in which IFN stimulation was observed. Some cell lines have been prepared for this, however the experiments have not yet been performed.

6.4 Identification of ARFs

The data regarding IFN stimulation was overlapped with various datasets on viral infection to identify candidate ARFs which were both upregulated by IFN and downregulated by viral infection. There are several ways in which this analysis could be expanded to identify more potential ARFs.

The current comparison of data to identify candidate restriction factors utilised the relatively stringent criteria that the protein should be IFN stimulated in all five donors. More candidates may have been identified if a lower threshold was employed. For example, epherin receptor B2 (EPHB2) did not fulfil this criteria, however it was upregulated in four monocyte donors and was upregulated 2.4 fold on average. It was additionally upregulated more than 2 fold in both pan-monocyte samples, and 1.9 fold in THP-1s. It was downregulated at the WCL and PM level during HCMV infection, and also by VACV. EPHB2 is a receptor tyrosine kinase which forms heterotetrameric receptor-ligand complexes with Ephrin ligands, and triggers bidirectional signalling, with

roles in cell migration and development (Park & Lee, 2015; Bryson & Bhandoola, 2015). The family of Epherin receptors and ligands has various roles in immunity, and have been previously identified as viral entry receptors for KSHV and EBV (Darling & Lamb, 2019).

Additionally, overlap with the IFN and VACV data so far only included WCL data for viral infection. As the IFN data focusses on the cell surface, it will be interesting to see if more of the proteins quantified in the IFN data are identified in the PM VACV dataset. Another limitation of this so far, is that many of the viral infections were in fibroblasts, so a number of the proteins identified in the primary leukocytes were not quantified in the viral data. Some additional proteins were seen in the CEMT4 HIV infection dataset. Additional experiments are currently being performed in collaboration with Dr Richard Stanton (Cardiff University), investigating the effect of HCMV infection of primary DCs. This will be especially interesting for overlap with the IFN data, as there are likely to be more commonly quantified proteins.

This project attempted to identify restriction factors on the basis of IFN stimulation and viral downregulation. Another approach discussed briefly in this thesis was the overlap of data from multiple viruses to identify proteins which are commonly targeted and therefore likely to be important in infection. In this investigation, the HCMV and VACV infection datasets were overlapped to identify 85 commonly downregulated proteins. HDAC5 was one such protein, and was later found to restrict VACV and HSV-1 infection. I previously performed a similar comparison between HCMV and EBV infection to identify commonly downregulated proteins (Ersing *et al*, 2017). This highlighted multiple poly(A) RNA binding proteins, and several proteins involved in synapse organisation including three neuroligins. It is of course not always the case that a restriction factor is able to limit infection with all viruses. BST2 for example is a potent HIV restriction factor, but HCMV may actually co-opt it, using a ‘reverse tethering’ mechanism to facilitate entry (Viswanathan *et al*, 2011). However, a comprehensive overlap of the accumulating datasets in the Weekes lab, investigating different viral infections, would be interesting.

6.5 Concluding Remarks

This work used proteomic screens exploring two characteristic features of ARFs: stimulation by IFN and downregulation by viral infection. As well as providing a basis for identifying candidate ARFs, these screens have both provided interesting and expansive datasets in their own right, which are useful resources in the field and a basis for many further investigations.

BIBLIOGRAPHY

Abdel-Mohsen M, Deng X, Liegler T, Guatelli JC, Salama MS, Ghanem HE -d. A, Rauch A, Ledergerber B, Deeks SG, Gunthard HF, Wong JK & Pillai SK (2014) Effects of Alpha Interferon Treatment on Intrinsic Anti-HIV-1 Immunity In Vivo. J. Virol. 88: 763–767

Adler M, Tavalai N, Müller R & Stamminger T (2011) Human cytomegalovirus immediate-early gene expression is restricted by the nuclear domain 10 component Sp100. J. Gen. Virol. 92: 1532–1538

Agalioti T, Lomvardas S, Parekh B, Yie J, Maniatis T & Thanos D (2000) Ordered recruitment of chromatin modifying and general transcription factors to the IFN-β promoter. Cell 103: 667–678

Alberts B (2015) Molecular biology of the cell

Altfeld M & Gale MJ (2015) Innate immunity against HIV-1 infection. Nat. Immunol. 16: 554–562

Amiri-Dashatan N, Koushki M, Abbaszadeh HA, Rostami-Nejad M & Rezaei-Tavirani M (2018) Proteomics applications in health: Biomarker and drug discovery and food industry. Iran. J. Pharm. Res. 17: 1523–1536

Anderson JS, Javien J, Nolta JA & Bauer G (2009) Preintegration HIV-1 inhibition by a combination lentiviral vector containing a chimeric TRIM5α protein, a CCR5 shRNA, and a TAR decoy. Mol. Ther. 17: 2103–2114

Arend WP, Malyak M, Guthridge CJ & Gabay C (1998) INTERLEUKIN-1 RECEPTOR ANTAGONIST: Role in Biology. Annu. Rev. Immunol. 16: 27–55

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM & Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25: 25–29

Ashley CL, Glass MS, Abendroth A, McSharry BP & Slobedman B (2017) Nuclear domain 10 components upregulated via interferon during human cytomegalovirus infection potently regulate viral infection. J. Gen. Virol. 98: 1795–1805

Aso H, Ito J, Koyanagi Y & Sato K (2019) Comparative description of the expression profile of interferon-stimulated genes in multiple cell lineages targeted by HIV-1 infection. Front. Microbiol. 10:

Assarsson E, Greenbaum JA, Sundstrom M, Schaffer L, Hammond JA, Pasquetto V, Oseroff C, Hendrickson RC, Lefkowitz EJ, Tscharke DC, Sidney J, Grey HM, Head SR,

Peters B & Sette A (2008) Kinetic analysis of a complete poxvirus transcriptome reveals an immediate-early class of genes. Proc Natl Acad Sci U S A 105: 2140–2145

Azevedo LS, Pierrotti LC, Abdala E, Costa SF, Strabelli TMV, Campos SV, Ramos JF, Latif AZA, Litvinov N, Maluf NZ, Caiaffa Filho HH, Pannuti CS, Lopes MH, dos Santos VA, Linardi C da CG, Yasuda MAS & Marques HH de S (2015) Cytomegalovirus infection in transplant recipients. Clin. (São Paulo, Brazil) 70: 515–23

Bacsa S, Karasneh G, Dosa S, Liu J, Valyi-Nagy T & Shukla D (2011) Syndecan-1 and syndecan-2 play key roles in herpes simplex virus type-1 infection. J. Gen. Virol. 92: 733–743

Baldick Jr. CJ, Moss B, Baldick CJ & Moss B (1993) Characterization and temporal regulation of mRNAs encoded by vaccinia virus intermediate-stage genes. J. Virol. 67: 3515–3527

Barnes K, Brown C & Turner AJ (1998) Endothelin-Converting Enzyme:Ultrastructural Localization and Its Recycling From the Cell Surface. Hypertension 31: 3–9

Bartel S, Doellinger J, Darsow K, Bourquain D, Buchholz R, Nitsche A & Lange HA (2011) Proteome analysis of vaccinia virus IHD-W-infected HEK 293 cells with 2-dimensional gel electrophoresis and MALDI-PSD-TOF MS of on solid phase support N-terminally sulfonated peptides. Virol J 8: 380

Bausch-Fluck D, Hofmann A, Bock T, Frei AP, Cerciello F, Jacobs A, Moest H, Omasits U, Gundry RL, Yoon C, Schiess R, Schmidt A, Mirkowska P, Härtlová A, Van Eyk JE, Bourquin JP, Aebersold R, Boheler KR, Zandstra P & Wollscheid B (2015) A mass spectrometric-derived cell surface protein atlas. PLoS One 10: 1–22

Bazhin A V, von Ahn K, Fritz J, Werner J & Karakhanova S (2018) Interferon-α Up-Regulates the Expression of PD-L1 Molecules on Immune Cells Through STAT3 and p38 Signaling. Front. Immunol. 9: 1–13

Beltran PMJ & Cristea IM (2015) The lifecycle and pathogenesis of human cytomegalovirus infection: lessons from proteomics. Expert Rev Proteomics 11: 697–711

Bergantz L, Subra F, Deprez E, Delelis O & Richetta C (2019) Interplay between Intrinsic and Innate Immunity during HIV Infection. Cells 8: 922

Biolatti M, Gugliesi F, Dell'Oste V & Landolfo S (2018) Modulation of the innate immune response by human cytomegalovirus. Infect. Genet. Evol. 64: 105–114

Biron CA, Byron KS & Sullivan JL (1989) Severe Herpesvirus Infections in an Adolescent without Natural Killer Cells. N. Engl. J. Med. 320: 1731–1735

Boehme KW, Guerrero M & Compton T (2006) Human cytomegalovirus envelope glycoproteins B and H are necessary for TLR2 activation in permissive cells. J Immunol 177: 7094–7102

Bosshart H & Heinzelmann M (2016) THP-1 cells as a model for human monocytes. Ann. Transl. Med. 4: 4–7

van Boxel-Dezaire AHH, Rani MRS & Stark GR (2006) Complex Modulation of Cell Type-Specific Signaling in Response to Type I Interferons. Immunity 25: 361–372

van Boxel-Dezaire AHH, Zula JA, Xu Y, Ransohoff RM, Jacobberger JW & Stark GR (2010) Major Differences in the Responses of Primary Human Leukocyte Subsets to IFN-β. J. Immunol. 185: 5888–5899

Boyette LB, Macedo C, Hadi K, Elinoff BD, Walters JT, Ramaswami B, Chalasani G, Taboas JM, Lakkis FG & Metes DM (2017) Phenotype, function, and differentiation potential of human monocyte subsets. PLoS One 12: e0176460

Brass AL, Huang IC, Benita Y, John SP, Krishnan MN, Feeley EM, Ryan BJ, Weyer JL, van der Weyden L, Fikrig E, Adams DJ, Xavier RJ, Farzan M & Elledge SJ (2009) The IFITM Proteins Mediate Cellular Resistance to Influenza A H1N1 Virus, West Nile Virus, and Dengue Virus. Cell 139: 1243–1254

Broyles SS & Fesler BS (1990) Vaccinia virus gene encoding a component of the viral early transcription factor. J Virol 64: 1523–1529

Bryson JL & Bhandoola A (2015) Editorial: Ephs, ephrins, and early T cell development. J. Leukoc. Biol. 98: 877–9

Burshtyn DN (2013) NK cells and poxvirus infection. Front. Immunol. 4: 7

Businger R, Deutschmann J, Gruska I, Milbradt J, Wiebusch L, Gramberg T & Schindler M (2019) Human cytomegalovirus overcomes SAMHD1 restriction in macrophages via pUL97. Nat. Microbiol. 4: 2260–2272

Caller LG, Davies CTR, Antrobus R, Lehner PJ, Weekes MP & Crump CM (2019) Temporal Proteomic Analysis of BK Polyomavirus Infection Reveals Virus-Induced G2 Arrest and Highly Effective Evasion of Innate Immune Sensing. J Virol 93:

Campbell JH, Hearps AC, Martin GE, Williams KC & Crowe SM (2014) The importance of monocytes and macrophages in HIV pathogenesis, treatment, and cure. AIDS 28: 2175–2187

Carter GC, Rodger G, Murphy BJ, Law M, Krauss O, Hollinshead M & Smith GL (2003) Vaccinia virus cores are transported on microtubules. J Gen Virol 84: 2443–2458

Cattaruzza F, Poole DP & Bunnett NW (2013) Arresting inflammation: contributions of plasma membrane and endosomal signalling to neuropeptide-driven inflammatory disease. Biochem Soc Trans 41: 137–143

Caza T & Landas S (2015) Functional and Phenotypic Plasticity of CD4+ T Cell Subsets. Biomed Res. Int. 2015:

Chan KC & Issaq HJ (2013) Fractionation of Peptides by Strong Cation-Exchange Liquid Chromatography BT - Proteomics for Biomarker Discovery. In, Zhou M & Veenstra T (eds) pp 311–315. Totowa, NJ: Humana Press

Chang HW, Watson JC & Jacobs BL (1992) The E3L gene of vaccinia virus encodes an inhibitor of the interferon-induced, double-stranded RNA-dependent protein kinase. Proc. Natl. Acad. Sci. U. S. A. 89: 4825–4829

Chauhan A, Hahn S, Gartner S, Pardo CA, Netesan SK, McArthur J & Nath A (2007) Molecular programming of endothelin-1 in HIV-infected brain: Role of Tat in up-regulation of ET-1 and its inhibition by statins. FASEB J. 21: 777–789

Chou W, Ngo T & Gershon PD (2012) An overview of the vaccinia virus infectome: a survey of the proteins of the poxvirus-infected cell. J Virol 86: 1487–1499

Cobbold M, Khan N, Pourgheysari B, Tauro S, McDonald D, Osman H, Assenmacher M, Billingham L, Steward C, Crawley C, Olavarria E, Goldman J, Chakraverty R, Mahendra P, Craddock C & Moss PAH (2005) Adoptive transfer of cytomegalovirus-specific CTL to stem cell transplant patients after selection by HLA-peptide tetramers. J. Exp. Med. 202: 379–386

Colamonici OR, Domanski P, Sweitzer SM, Larner A & Buller RM (1995) Vaccinia virus B18R gene encodes a type I interferon-binding protein that blocks interferon alpha transmembrane signaling. J Biol Chem 270: 15974–15978

Colomer-Lluch M, Ruiz A, Moris A & Prado JG (2018) Restriction Factors: From Intrinsic Viral Restriction to Shaping Cellular Immunity Against HIV-1. Front Immunol 9: 2876

Cornberg M, Jaroszewicz J, Manns MP & Wedemeyer H (2010) Treatment of chronic hepatitis B. Minerva Gastroenterol. Dietol. 56: 451–465

Cottrell GS, Padilla BE, Amadesi S, Poole DP, Murphy JE, Hardt M, Roosterman D, Steinhoff M & Bunnett NW (2009) Endosomal endothelin-converting enzyme-1. A regulator of β-arrestin-dependent ERK signaling. J. Biol. Chem. 284: 22411–22425

Cox J & Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26: 1367–1372

Croft NP, de Verteuil DA, Smith SA, Wong YC, Schittenhelm RB, Tscharke DC & Purcell AW (2015) Simultaneous Quantification of Viral Antigen Expression Kinetics Using Data-Independent (DIA) Mass Spectrometry. Mol Cell Proteomics 14: 1361–1372

Daffis S, Szretter KJ, Schriewer J, Li J, Youn S, Errett J, Lin TY, Schneller S, Zust R, Dong H, Thiel V, Sen GC, Fensterl V, Klimstra WB, Pierson TC, Buller RM, Gale Jr. M, Shi PY & Diamond MS (2010) 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature 468: 452–456

Van Damme N, Goff D, Katsura C, Jorgenson RL, Mitchell R, Johnson MC, Stephens EB & Guatelli J (2008) The Interferon-Induced Protein BST-2 Restricts HIV-1 Release and Is Downregulated from the Cell Surface by the Viral Vpu Protein. Cell Host Microbe 3: 245–252

Darling TK & Lamb TJ (2019) Emerging roles for Eph receptors and ephrin ligands in immunity. Front. Immunol. 10: 1–15

Dephoure N & Gygi SP (2011) A solid phase extraction-based platform for rapid phosphoproteomic analysis. Methods 54: 379–386

Dho SH, Lee K, Jeong D, Kim C-J, Chung K-S, Kim JY, Park B-C, Park SS, Kim S & Kwon K (2016) GPR171 expression enhances proliferation and metastasis of lung cancer cells. Oncotarget 7:

Diamond MS & Farzan M (2013) The broad-spectrum antiviral functions of IFIT and IFITM proteins. Nat Rev Immunol 13: 46–57

Dolan A, Cunningham C, Hector RD, Hassan-Walker AF, Lee L, Addison C, Dargan DJ, McGeoch DJ, Gatherer D, Emery VC, Griffiths PD, Sinzger C, McSharry BP, Wilkinson GWG & Davison AJ (2004) Genetic content of wild-type human cytomegalovirus. J. Gen. Virol. 85: 1301–1312

Dollard SC, Grosse SD & Ross DS (2007) New estimates of the prevalence of neurological and sensory sequelae and mortality associated with congenital cytomegalovirus infection. Rev. Med. Virol. 17: 355–63

Doyle T, Goujon C & Malim MH (2015) HIV-1 and interferons: who's interfering with whom? Nat. Rev. Microbiol. 13: 403–413

Duggal NK & Emerman M (2012) Evolutionary conflicts between viruses and restriction factors shape immunity. Nat. Rev. Immunol. 12: 687–695

Eckman EA, Watson M, Marlow L, Sambamurti K & Eckman CB (2003) Alzheimer's disease β-amyloid peptide is increased in mice deficient in endothelin-converting enzyme. J. Biol. Chem. 278: 2081–2084

Ehrenreich H, Anderson RW, Fox CH, Rieckmann P, Hoffman GS, Travis WD, Coligan JE, Kehrl JH & Fauci AS (1990) Endothelins, peptides with potent vasoactive properties, are produced by human macrophages. J. Exp. Med. 172: 1741–8

Ehrenreich H, Rieckmann P, Sinowatz F, Weih KA, Arthur LO, Goebel FD, Burd PR, Coligan JE & Clouse KA (1993) Potent stimulation of monocytic endothelin-1 production by HIV-1 glycoprotein 120. J. Immunol. 150: 4601 LP – 4609

Elias JE & Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods 4: 207–214

Elias JE & Gygi SP (2010) Target-decoy search strategy for mass spectrometry-based proteomics. Methods Mol. Biol. 604: 55–71

Emery VC (2001) Investigation of CMV disease in immunocompromised patients. J. Clin. Pathol. 54: 84–88

Engelman A & Cherepanov P (2012) The structural biology of HIV-1: mechanistic and therapeutic insights. Nat. Rev. Microbiol. 10: 279–290

Ersing I, Nobre L, Wang LW, Soday L, Ma Y, Paulo JA, Narita Y, Ashbaugh CW, Jiang C, Grayson NE, Kieff E, Gygi SP, Weekes MP & Gewurz BE (2017) A Temporal Proteomic Map of Epstein-Barr Virus Lytic Replication in B Cells. Cell Rep 19: 1479–1493

Fahnoe DC, Knapp J, Johnson GD & Ahn K (2000) Inhibitor Potencies and Substrate Preference for Endothelin-Converting Enzyme-1 are Dramatically Affected by PH. J. Cardiovasc. Pharmacol. 36:

Farina G, York M, Collins C & Lafyatis R (2011) dsRNA activation of endothelin-1 and markers of vascular activation in endothelial cells and fibroblasts. Ann. Rheum. Dis. 70: 544 LP – 550

Fenner F, Henderson DA, Arita I, Jezek Z & Ladnyi ID (1988) Smallpox and its eradication. World Heal. Organ. Geneva

Ferguson BJ, Mansur DS, Peters NE, Ren H & Smith GL (2012) DNA-PK is a DNA sensor for IRF-3-dependent innate immunity. Elife 1: e00047

Fields BN, Knipe DM & Howley PM (2013) Fields virology Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins

Fitzpatrick K, Skasko M, Deerinck TJ, Crum J, Ellisman MH & Guatelli J (2010) Direct restriction of virus release and incorporation of the interferon-induced protein BST-2 into HIV-1 particles. PLoS Pathog. 6:

Frasca L, Fedele G, Deaglio S, Capuano C, Palazzo R, Vaisitti T, Malavasi F & Ausiello CM (2006) CD38 orchestrates migration, survival, and Th1 immune response of human mature dendritic cells. Blood 107: 2392–2399

de Freitas LFD, Oliveira RP, Miranda MCG, Rocha RP, Barbosa-Stancioli EF, Faria AMC & da Fonseca FG (2018) The Virulence of Different Vaccinia Virus Strains Is Directly Proportional to Their Ability To Downmodulate Specific Cell-Mediated Immune Compartments In Vivo. J. Virol. 93: 1–16

Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK & Sander JD (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat. Biotechnol. 31: 822–6

Fu Y, Maianu L, Melbert BR & Garvey WT (2004) Facilitative glucose transporter gene expression in human lymphocytes, monocytes, and macrophages: A role for GLUT isoforms 1, 3, and 5 in the immune response and foam cell formation. Blood Cells, Mol. Dis. 32: 182–190

Furukawa K, Saleh D, Bayan F, Emoto N, Kaw S, Yanagisawa M & Giaid A (1996) Co-expression of endothelin-1 and endothelin-converting enzyme-1 in patients with chronic rhinitis. Am. J. Respir. Cell Mol. Biol. 14: 248–253

Ganser-Pornillos BK & Pornillos O (2019) Restriction of HIV-1 and other retroviruses by TRIM5. Nat Rev Microbiol 17: 546–556

Gariano GR, Dell'Oste V, Bronzini M, Gatti D, Luganini A, de Andrea M, Gribaudo G, Gariglio M & Landolfo S (2012) The intracellular DNA sensor IFI16 gene acts as restriction factor for human cytomegalovirus replication. PLoS Pathog 8: e1002498

Gazit R, Garty B-Z, Monselise Y, Hoffer V, Finkelstein Y, Markel G, Katz G, Hanna J, Achdout H, Gruda R, Gonen-Gross T & Mandelboim O (2004) Expression of KIR2DL1 on the entire NK cell population: a possible novel immunodeficiency syndrome. Blood 103: 1965 LP – 1966

Gerlini G, Mariotti G, Chiarugi A, Di Gennaro P, Caporale R, Parenti A, Cavone L, Tun-Kyi A, Prignano F, Saccardi R, Borgognoni L & Pimpinelli N (2008) Induction of CD83 + CD14 + Nondendritic Antigen-Presenting Cells by Exposure of Monocytes to IFN-α. J. Immunol. 181: 2999–3008

Gibbert K, Schlaak JF, Yang D & Dittmer U (2013) IFN-α subtypes: Distinct biological activities in anti-viral therapy. Br. J. Pharmacol. 168: 1048–1058

Gomes I, Aryal DK, Wardman JH, Gupta A, Gagnidze K, Rodriguiz RM, Kumar S, Wetsel WC, Pintar JE, Fricker LD & Devi LA (2013) GPR171 is a hypothalamic G protein-coupled receptor for BigLEN, a neuropeptide involved in feeding. Proc. Natl. Acad. Sci. 110: 16211–16216

Gonzalez-Enriquez GV, Escoto-Delgadillo M, Vazquez-Valls E & Torres-Mendoza BM (2017) SERINC as a Restriction Factor to Inhibit Viral Infectivity and the Interaction with HIV. J. Immunol. Res. 2017:

Graessel A, Hauck SM, Von Toerne C, Kloppmann E, Goldberg T, Koppensteiner H, Schindler M, Knapp B, Krause L, Dietz K, Schmidt-Weber CB & Suttner K (2015) A combined omics approach to generate the surface atlas of human naive CD4+ T cells during early T-cell receptor activation. Mol. Cell. Proteomics 14: 2085–2102

Graves PR & Haystead TAJ (2002) Molecular biologist's guide to proteomics. Microbiol. Mol. Biol. Rev. 66: 39–63; table of contents

Greenwood EJ, Matheson NJ, Wals K, van den Boomen DJ, Antrobus R, Williamson JC & Lehner PJ (2016) Temporal proteomic analysis of HIV infection reveals remodelling of the host phosphoproteome by lentiviral Vif variants. Elife 5:

Griffin DO & Goff SP (2015) HIV-1 Is Restricted prior to Integration of Viral DNA in Primary Cord-Derived Human CD34 + Cells. J. Virol. 89: 8096–8100

Gruffat H, Manetr E & Sergeant A (2002) MEF2-mediated recruitment of class II HDAC at the EBV immediate early gene BZLF1 links latency and chromatin remodeling. EMBO Rep. 3: 141–146

Guerra S, Cáceres A, Knobeloch KP, Horak I & Esteban M (2008) Vaccinia virus E3 protein prevents the antiviral action of ISG15. PLoS Pathog. 4:

Guo ZS, Lu B, Guo Z, Giehl E, Feist M, Dai E, Liu W, Storkus WJ, He Y, Liu Z & Bartlett DL (2019) Vaccinia virus-mediated cancer immunotherapy: Cancer vaccines and oncolytics. J. Immunother. Cancer 7: 1–21

Haas W, Faherty BK, Gerber SA, Elias JE, Beausoleil SA, Bakalarski CE, Li X, Villen J & Gygi SP (2006) Optimization and use of peptide mass measurement accuracy in shotgun proteomics. Mol Cell Proteomics 5: 1326–1337

Häggström M (2014) Medical gallery of Mikael Häggström 2014. WikiJournal Med. 1:

Haider S & Pal R (2013) Integrated Analysis of Transcriptomic and Proteomic Data. Curr. Genomics 14: 91–110

Hallett M, Human R a T & Physiology C (1992) A Loss-of-Function Mutation in the EndothelinConverting Enzyme 1 (ECE-1) Associated with Hirschsprung Disease, Cardiac Defects, and Autonomic Dysfunction. Carbon N. Y. 15: 1106–1107

Hammonds J, Wang JJ, Yi H & Spearman P (2010) Immunoelectron microscopic evidence for tetherin/BST2 as the physical bridge between HIV-1 virions and the plasma membrane. PLoS Pathog. 6:

Helset E, Sildnes T & Konopski ZS (1994) Endothelin-1 Stimulates Monocytes in vitro to Release Chemotactic Activity Identified as Interleukin-8 and Monocyte Chemotactic Protein-1. Mediators Inflamm. 3: 155–160

Henig N, Avidan N, Mandel I, Staun-Ram E, Ginzburg E, Paperna T, Pinter RY & Miller A (2013) Interferon-Beta Induces Distinct Gene Expression Response Patterns in Human Monocytes versus T cells. PLoS One 8:

Hill-Batorski L, Halfmann P, Marzi A, Lopes TJS, Neumann G, Feldmann H & Kawaoka Y (2015) Loss of Interleukin 1 Receptor Antagonist Enhances Susceptibility to Ebola Virus Infection. J. Infect. Dis. 212: S329–S335

Hollinshead M, Johns HL, Sayers CL, Gonzalez-Lopez C, Smith GL & Elliott G (2012) Endocytic tubules regulated by Rab GTPases 5 and 11 are used for envelopment of herpes simplex virus. EMBO J 31: 4204–4220

Holmes KK, Bertozzi S, Bloom BR, Jha P, Gelband H, DeMaria LM & Horton S (2017) Major Infectious Diseases: Key Messages from Disease Control Priorities, Third Edition. In Disease Control Priorities, Third Edition (Volume 6): Major Infectious Diseases, Holmes KK Bertozzi S Bloom BR & Jha P (eds) pp 45–66. Washington, DC: The World Bank

Honda K, Takaoka A & Taniguchi T (2006) Type I Inteferon Gene Induction by the Interferon Regulatory Factor Family of Transcription Factors. Immunity 25: 349–360

Hong MeeAe SJ, Lim Chrissie KA, Jarret Abigail, Pangallo Joseph, Loo Yueh-Ming, Liu Shuanghu & Hagedorn Curt H, Gale Michael SR (2016) Interferon lambda 4 expression is suppressed by the host during viral infection. J. Exp. Med.: 1–18

Hong Y, MacNab S, Lambert LA, Turner AJ, Whitehouse A & Usmani BA (2011) Herpesvirus saimiri-based endothelin-converting enzyme-1 shRNA expression decreases prostate cancer cell invasion and migration. Int. J. Cancer 129: 586–598

Horgan RP & Kenny LC (2011) 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. Obstet. Gynaecol. 13: 189–195

Hosseinzadeh Z, Warsi J, Elvira B, Almilaji A, Shumilina E & Lang F (2015) Up-regulation of Kv1.3 Channels by Janus Kinase 2. J. Membr. Biol. 248: 309–317

Hotter D, Sauter D & Kirchhoff F (2013) Emerging role of the host restriction factor tetherin in viral immune sensing. J. Mol. Biol. 425: 4956–4964

Houde M, Desbiens L & D'Orléans-Juste P (2016) Endothelin-1: Biosynthesis, Signaling and Vasoreactivity. In Endothelium, Pharmacology RAKBT-A in (ed) pp 143–175. Academic Press

Hrecka K, Hao C, Gierszewska M, Swanson SK, Kesik-Brodacka M, Srivastava S, Florens L, Washburn MP & Skowronski J (2011) Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. Nature 474: 658–661

Hsu JL, van den Boomen DJH, Tomasec P, Weekes MP, Antrobus R, Stanton RJ, Ruckova E, Sugrue D, Wilkie GS, Davison AJ, Wilkinson GWG & Lehner PJ (2015) Plasma Membrane Profiling Defines an Expanded Class of Cell Surface Proteins Selectively Targeted for Degradation by HCMV US2 in Cooperation with UL141. PLoS Pathog. 11: 1–24

Huang DW, Sherman BT, Lempicki RA, (2009a) Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37: 1–13

Huang DW, Sherman BT, Lempicki RA (2009b) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4: 44–57

Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S, de Castro E, Coggill P, Corbett M, Das U, Daugherty L, Duquenne L, Finn RD, Fraser M, Gough J, Haft D, et al (2012) InterPro in 2011: new developments in the family and domain prediction database. Nucleic Acids Res 40: D306-12

Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, Beausoleil SA, Villen J, Haas W, Sowa ME & Gygi SP (2010) A tissue-specific atlas of mouse protein phosphorylation and expression. Cell 143: 1174–1189

Isaacs A & Lindenmann J (1957) Virus Interference. I. The Interferon. Proc. R. Soc. London. Ser. B - Biol. Sci. 147: 258 LP – 267

Itzhak DN, Davies C, Tyanova S, Mishra A, Williamson J, Antrobus R, Cox J, Weekes MP & Borner GHH (2017) A Mass Spectrometry-Based Approach for Mapping Protein Subcellular Localization Reveals the Spatial Proteome of Mouse Primary Neurons. Cell Rep 20: 2706–2718

Ivashkiv LB & Donlin LT (2014) Regulation of type I interferon responses. Nat. Rev. Immunol. 14: 36–49

Jacobs N, Bartlett NW, Clark RH & Smith GL (2008) Vaccinia virus lacking the Bcl-2-like protein N1 induces a stronger natural killer cell response to infection. J Gen Virol 89: 2877–2881

Jafri F & Ergul A (2003) Nuclear Localization of Endothelin-Converting Enzyme-1: Subisoform Specificity. Arterioscler. Thromb. Vasc. Biol. 23: 2192–2196

Ji X, Cheung R, Cooper S, Li Q, Greenberg HB & He XS (2003) Interferon alfa regulated gene expression in patients initiating interferon treatment for chronic hepatitis C. Hepatology 37: 610–621

Johnson GD, Stevenson T & Ahn K (1999) Hydrolysis of Peptide Hormones by Endothelin-converting. Biochemistry 274: 4053–4058

Julian MW, Shao G, VanGundy ZC, Papenfuss TL & Crouser ED (2013) Mitochondrial Transcription Factor A, an Endogenous Danger Signal, Promotes TNFα Release via RAGE- and TLR9-Responsive Plasmacytoid Dendritic Cells. PLoS One 8: e72354

Kall L, Canterbury JD, Weston J, Noble WS & MacCoss MJ (2007) Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nat Methods 4: 923–925

Kar UK, Simonian M & Whitelegge JP (2017) Integral membrane proteins: bottom-up, top-down and structural proteomics. Expert Rev. Proteomics 14: 715–723

Keck JG, Baldick Jr. CJ & Moss B (1990) Role of DNA replication in vaccinia virus gene expression: a naked template is required for transcription of three late trans-activator genes. Cell 61: 801–809

Kenneson A & Cannon MJ (2007) Review and meta-analysis of the epidemiology of congenital cytomegalovirus (CMV) infection. Rev. Med. Virol. 17: 253–76

Keskinen P, Ronni T, Matikainen S, Lehtonen A & Julkunen I (1997) Regulation of HLA class I and II expression by interferons and influenza A virus in human peripheral blood mononuclear cells. Immunology 91: 421–429

Kestler 3rd HW, Ringler DJ, Mori K, Panicali DL, Sehgal PK, Daniel MD, Desrosiers RC, Kestier HW, Ringler DJ, Mori K, Panicali DL, Sehgal PK, Daniel MD & Desrosiers RC (1991) Importance of the nef gene for maintenance of high virus loads and for development of AIDS. Cell 65: 651–662

Kim DD & Song WC (2006) Membrane complement regulatory proteins. Clin. Immunol. 118: 127–136

Kim ET, White TE, Brandariz-Núñez A, Diaz-Griffero F & Weitzman MD (2013) SAMHD1 restricts herpes simplex virus 1 in macrophages by limiting DNA replication. J. Virol. 87: 12949–56

Kim W, Bennett EJ, Huttlin EL, Guo A, Li J, Possemato A, Sowa ME, Rad R, Rush J, Comb MJ, Harper JW & Gygi SP (2011a) Systematic and quantitative assessment of the ubiquitin-modified proteome. Mol Cell 44: 325–340

Kim YE, Lee JH, Kim ET, Shin HJ, Gu SY, Seol HS, Ling PD, Lee CH & Ahn JH (2011b) Human cytomegalovirus infection causes degradation of Sp100 proteins that suppress viral gene expression. J. Virol. 85: 11928–11937

Kirschner MW (2005) The meaning of systems biology. Cell 121: 503–504

Kisselev AF, van der Linden WA & Overkleeft HS (2012) Proteasome Inhibitors: An Expanding Army Attacking a Unique Target. Chem. Biol. 19: 99–115

Kline JN, Geist LJ, Monick MM, Stinski MF & Hunninghake GW (1994) Regulation of expression of the IL-1 receptor antagonist (IL-1ra) gene by products of the human cytomegalovirus immediate early genes. J. Immunol. 152: 2351 LP – 2357

Kluge SF, Sauter D & Kirchhoff F (2015) SnapShot: Antiviral Restriction Factors. Cell 163: 774-774e1

Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera Mdel C & Yusa K (2014) Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat Biotechnol 32: 267–273

Korthals M, Safaian N, Kronenwett R, Maihöfer D, Schott M, Papewalis C, Diaz Blanco E, Winter M, Czibere A, Haas R, Kobbe G & Fenk R (2007) Monocyte derived dendritic cells generated by IFN-α acquire mature dendritic and natural killer cell properties as shown by gene expression analysis. J. Transl. Med. 5: 1–11

Kotton CN (2019) Updates on antiviral drugs for cytomegalovirus prevention and treatment. Curr. Opin. Organ Transplant. 24: 469–475

Kummer S, Flöttmann M, Schwanhäusser B, Sieben C, Veit M, Selbach M, Klipp E & Herrmann A (2014) Alteration of protein levels during influenza virus H1N1 infection in host cells: A proteomic survey of host and virus reveals differential dynamics. PLoS One 9: 1–11

Kuruppu S, Reeve S & Ian Smith A (2007) Characterisation of endothelin converting enzyme-1 shedding from endothelial cells. FEBS Lett. 581: 4501–4506

Kuruppu S & Smith AI (2012) Endothelin converting enzyme-1 phosphorylation and trafficking. FEBS Lett. 586: 2212–2217

Kuruppu S, Tochon-Danguy N & Smith AI (2012) Protein Kinase C recognition sites in the cytoplasmic domain of Endothelin Converting Enzyme-1c. Biochem. Biophys. Res. Commun. 427: 606–610

Kyoizumi S, Ohara T, Kusunoki Y, Hayashi T, Koyama K & Tsuyama N (2004) Expression Characteristics and Stimulatory Functions of CD43 in Human CD4 + Memory T Cells: Analysis Using a Monoclonal Antibody to CD43 That Has a Novel Lineage Specificity. J. Immunol. 172: 7246–7253

Lahouassa H, Daddacha W, Hofmann H, Ayinde D, Eric C, Dragin L, Bloch N, Maudet C, Bertrand M, Pancino G, Priet S, Canard B, Laguette N, Benkirane M, Transy C,

Landau NR & Kim B (2013) SAMHD1 restricts HIV-1 by reducing the intracellular pool of deoxynucleotide triphosphates. Nat. Immunol. 13: 223–228

Lambert LA, Whyteside AR, Turner AJ & Usmani BA (2008) Isoforms of endothelin-converting enzyme-1 (ECE-1) have opposing effects on prostate cancer cell invasion. Br. J. Cancer 99: 1114–1120

Landini MP, Rossier E & Schmitz H (1988) Antibodies to human cytomegalovirus structural polypeptides during primary infection. J. Virol. Methods 22: 309–317

Lanzavecchia A & Sallusto F (2000) Dynamics of T lymphocyte responses: Intermediates, effectors, and memory cells. Science 290: 92–97

Lebbink RJ, de Ruiter T, Adelmeijer J, Brenkman AB, van Helvoort JM, Koch M, Farndale RW, Lisman T, Sonnenberg A, Lenting PJ & Meyaard L (2006) Collagens are functional, high affinity ligands for the inhibitory immune receptor LAIR-1. J. Exp. Med. 203: 1419–1425

Lemercier C, Verdel A, Galloo B, Curtet S, Brocard MP & Khochbin S (2000) mHDA1/HDAC5 histone deacetylase interacts with and represses MEF2A transcriptional activity. J. Biol. Chem. 275: 15594–15599

Lens MB & Dawes M (2002) Interferon alfa therapy for malignant melanoma: a systematic review of randomized controlled trials. J. Clin. Oncol. 20: 1818–1825

Létourneau M, Roby P, Tremblay F, Carette J, Fournier A, Letourneau M, Roby P, Tremblay F, Carette J FA, Létourneau M, Roby P, Tremblay F, Carette J, Fournier A & Letourneau M, Roby P, Tremblay F, Carette J FA (2000) New Short Peptide Substrates of Endothelin-Converting Enzyme and Characterization of the Enzyme. J. Cardiovasc. Pharmacol. 36: S28–S29

Li C, Du S, Tian M, Wang Y, Bai J, Tan P, Liu W, Yin R, Wang M, Jiang Y, Li Y, Zhu N, Zhu Y, Li T, Wu S, Jin N & He F (2018) The Host Restriction Factor Interferon-Inducible Transmembrane Protein 3 Inhibits Vaccinia Virus Infection. Front Immunol 9: 228

Li Y, Qin H & Ye M (2020) An overview on enrichment methods for cell surface proteome profiling. J. Sep. Sci. 43: 292–312

Lietzén N, Öhman T, Rintahaka J, Julkunen I, Aittokallio T, Matikainen S & Nyman TA (2011) Quantitative Subcellular Proteome and Secretome Profiling of Influenza A Virus-Infected Human Primary Macrophages. PLoS Pathog. 7:

Liu B, Li NL, Wang J, Shi PY, Wang T, Miller MA & Li K (2014) Overlapping and distinct molecular determinants dictating the antiviral activities of TRIM56 against flaviviruses and coronavirus. J Virol 88: 13821–13835

Liu JSH, Amaral TD, Brosnan CF & Lee SC (1998) IFNs are critical regulators of IL-1 receptor antagonist and IL-1 expression in human microglia. J. Immunol. 161: 1989–1996

Liu R, Olano LR, Mirzakhanyan Y, Gershon PD, Moss B, Ifits TI, Liu R, Olano LR, Mirzakhanyan Y, Gershon PD, Liu R, Olano LR, Mirzakhanyan Y, Gershon PD & Moss B (2019) Vaccinia Virus Ankyrin-Repeat / F-Box Protein Article Vaccinia Virus Ankyrin-Repeat / F-Box Protein Targets Interferon-Induced IFITs for Proteasomal Degradation. CellReports 29: 816-828.e6

Liu W, Li T, Wang P, Liu W, Liu F, Mo X, Liu Z, Song Q, Lv P, Ruan G & Han W (2018) LRRC25 plays a key role in all-trans retinoic acid-induced granulocytic differentiation as a novel potential leukocyte differentiation antigen. Protein Cell 9: 785–798

Liu Y, Beyer A & Aebersold R (2016) On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell 165: 535–550

Lomonte P, Thomas J, Texier P, Caron C, Khochbin S & Epstein AL (2004) Functional interaction between class II histone deacetylases and ICP0 of herpes simplex virus type 1. J Virol 78: 6744–6757

Lu J & Willars GB (2019) Endothelin-converting enzyme-1 regulates glucagon-like peptide-1 receptor signalling and resensitisation. Biochem. J. 476: 513–533

Lu Y, Stuart JH, Talbot-Cooper C, Agrawal-Singh S, Huntly B, Smid AI, Snowden JS, Dupont L & Smith GL (2019) Histone deacetylase 4 promotes type I interferon signaling, restricts DNA viruses, and is degraded via vaccinia virus protein C6. Proc. Natl. Acad. Sci. U. S. A. 116: 11997–12006

Ma F, Zhang C, Prasad KVS, Freeman GJ & Schlossman SF (2001) Molecular cloning of Porimin, a novel cell surface receptor mediating oncotic cell death. Proc. Natl. Acad. Sci. U. S. A. 98: 9778–9783

Maartens G, Celum C & Lewin SR (2014) HIV infection: Epidemiology, pathogenesis, treatment, and prevention. Lancet 384: 258–271

Mack EA, Kallal LE, Demers DA & Biron CA (2011) Type 1 interferon induction of natural killer cell gamma interferon production for defense during lymphocytic choriomeningitis virus infection. MBio 2: 13–15

Macours N, Poels J, Hens K, Francis C & Huybrechts R (2004) Structure, Evolutionary Conservation, and Functions of Angiotensin- and Endothelin-Converting Enzymes. Int Rev Cytol. Volume 239: 47–97

Madhav N, Oppenheim B, Gallivan M, Mulembakani P, Rubin E & Wolfe N (2017) Pandemics: Risks, Impacts, and Mitigation. In Disease Control Priorities, Third Edition (Volume 9): Improving Health and Reducing Poverty, Jamison DT Gelband H Horton S Jha P Laxminarayan R Mock CN & Nugent R (eds) pp 315–345. Washington, DC: The World Bank

de Magalhaes JC, Andrade AA, Silva PN, Sousa LP, Ropert C, Ferreira PC, Kroon EG, Gazzinelli RT & Bonjardim CA (2001) A mitogenic signal triggered at an early stage of vaccinia virus infection: implication of MEK/ERK and protein kinase A in virus multiplication. J Biol Chem 276: 38353–38360

Makarov A & Denisov E (2009) Dynamics of ions of intact proteins in the Orbitrap mass analyzer. J. Am. Soc. Mass Spectrom. 20: 1486–1495

Malim MH & Bieniasz PD (2012) HIV restriction factors and mechanisms of evasion. Cold Spring Harb. Perspect. Med. 2: 1–16

Malim MH & Emerman M (2008) HIV-1 Accessory Proteins-Ensuring Viral Survival in a Hostile Environment. Cell Host Microbe 3: 388–398

Malm G & Engman M-L (2007) Congenital cytomegalovirus infections. Semin. Fetal Neonatal Med. 12: 154–159

Marshall ASJ, Willmen JA, Lin HH, Williams DL, Gordon S & Brown GD (2004) Identification and Characterization of a Novel Human Myeloid Inhibitory C-type Lectin-like Receptor (MICL) That Is Predominantly Expressed on Granulocytes and Monocytes. J. Biol. Chem. 279: 14792–14802

Martínez-Miguel P, Medrano-Andrés D, Griera-Merino M, Ortiz A, Rodríguez-Puyol M, Rodríguez-Puyol D & López-Ongil S (2017) Tweak up-regulates endothelin-1 system in mouse and human endothelial cells. Cardiovasc. Res. 113: 207–221

Martínez-Miguel P, Raoch V, Zaragoza C, Valdivielso JM, Rodríguez-Puyol M, Rodríguez-Puyol D & López-Ongil S (2009) Endothelin-converting enzyme-1 increases in atherosclerotic mice: potential role of oxidized low density lipoproteins. J. Lipid Res. 50: 364–75

Matheson NJ, Sumner J, Wals K, Rapiteanu R, Weekes MP, Vigan R, Weinelt J, Schindler M, Antrobus R, Costa ASH, Frezza C, Clish CB, Neil SJD & Lehner PJ (2015) Cell Surface Proteomic Map of HIV Infection Reveals Antagonism of Amino Acid Metabolism by Vpu and Nef. Cell Host Microbe 18: 409–423

Matthews SJ & McCoy C (2004) Peginterferon alfa-2a: A review of approved and investigational uses. Clin. Ther. 26: 991–1025

McAlister GC, Huttlin EL, Haas W, Ting L, Jedrychowski MP, Rogers JC, Kuhn K, Pike I, Grothe RA, Blethrow JD & Gygi SP (2012) Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. Anal. Chem. 84: 7469–7478

McAlister GC, Nusinow DP, Jedrychowski MP, Wuhr M, Huttlin EL, Erickson BK, Rad R, Haas W & Gygi SP (2014) MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. Anal Chem 86: 7150–7158

McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N & Haynes BF (2010) The immune response during acute HIV-1 infection: clues for vaccine development. Nat. Rev. Drug Discov. 10: 11–23

McNab F, Mayer-Barber K, Sher A, Wack A & O'Garra A (2015) Type I interferons in infectious disease. Nat. Rev. Immunol. 15: 87–103

McSharry BP, Jones CJ, Skinner JW, Kipling D & Wilkinson GW (2001) Human telomerase reverse transcriptase-immortalized MRC-5 and HCA2 human fibroblasts are fully permissive for human cytomegalovirus. J Gen Virol 82: 855–863

Mercer J, Snijder B, Sacher R, Burkard C, Bleck CK, Stahlberg H, Pelkmans L & Helenius A (2012) RNAi screening reveals proteasome- and Cullin3-dependent stages in vaccinia virus infection. Cell Rep 2: 1036–1047

Michel N, Allespach I, Venzke S, Fackler OT & Keppler OT (2005) The nef protein of human immunodeficiency virus establishes superinfection immunity by a dual strategy to downregulate cell-surface CCR5 and CD4. Curr. Biol. 15: 714–723

Mitchell RS, Katsura C, Skasko MA, Fitzpatrick K, Lau D, Ruiz A, Stephens EB, Margottin-Goguet F, Benarous R & Guatelli JC (2009) Vpu antagonizes BST-2-mediated restriction of HIV-1 release via β-TrCP and endo-lysosomal trafficking. PLoS Pathog. 5:

Mogensen TH, Melchjorsen J, Larsen CS & Paludan SR (2010) Innate immune recognition and activation during HIV infection. Retrovirology 7: 54

Molet S, Furukawa K, Maghazechi A, Hamid Q & Giaid A (2000) Chemokine- and cytokine-induced expression of endothelin 1 and endothelin-converting enzyme 1 in endothelial cells. J. Allergy Clin. Immunol. 105: 333–338

Monferran S, Muller C, Mourey L, Frit P & Salles B (2004) The Membrane-associated Form of the DNA Repair Protein Ku is Involved in Cell Adhesion to Fibronectin. J. Mol. Biol. 337: 503–511

Morchikh M, Cribier A, Raffel R, Amraoui S, Cau J, Severac D, Dubois E, Schwartz O, Bennasser Y & Benkirane M (2017) HEXIM1 and NEAT1 Long Non-coding RNA Form a Multi-subunit Complex that Regulates DNA-Mediated Innate Immune Response. Mol. Cell 67: 387-399.e5

Moss B (1968) Inhibition of HeLa cell protein synthesis by the vaccinia virion. J Virol 2: 1028–1037

Moss B, Knipe DM & Howley PM (2013) Poxviridae. In Fields Virology, Knipe DM & Howley PM (eds) pp 2129–2159. Philadelphia: Lippincott Williams & Wilkins

Moss B & Salzman NP (1968) Sequential Protein Synthesis Following Vaccinia Virus Infection. J. Virol. 2: 1016–1027

Muller L, Barret A, Etienne E, Meidan R, Valdenaire O, Corvol P & Tougard C (2003) Heterodimerization of endothelin-converting enzyme-1 isoforms regulates the subcellular distribution of this metalloprotease. J. Biol. Chem. 278: 545–555

Müller U, Steinhoff U, Reis LFL, Hemmi S, Pavlovic J, Zinkernagel RM & Aguet M (1994) Functional role of type I and type II interferons in antiviral defense. Science 264: 1918–1921

Murrell B, Vollbrecht T, Guatelli J & Wertheim JO (2016) The Evolutionary Histories of Antiretroviral Proteins SERINC3 and SERINC5 Do Not Support an Evolutionary Arms Race in Primates. J. Virol. 90: JVI.00972-16

Naamati A, Williamson JC, Greenwood EJD, Marelli S, Lehner PJ & Matheson NJ (2019) Functional proteomic atlas of HIV infection in primary human CD4+ T cells. Elife 8: 1–27

Najarro P, Traktman P & Lewis JA (2001) Vaccinia virus blocks gamma interferon signal transduction: viral VH1 phosphatase reverses Stat1 activation. J Virol 75: 3185–3196

Nathans R, Cao H, Sharova N, Ali A, Sharkey M, Stranska R, Stevenson M & Rana TM (2008) Small-molecule inhibition of HIV-1 Vif. Nat. Biotechnol. 26: 1187–1192

Neagu MR, Ziegler P, Pertel T, Strambio-De-Castillia C, Grütter C, Martinetti G, Mazzucchelli L, Grütter M, Manz MG & Luban J (2009) Potent inhibition of HIV-1 by TRIM5-cyclophilin fusion proteins engineered from human components. J. Clin. Invest. 119: 3035–3047

Neil SJD, Zang T & Bieniasz PD (2008) Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature 451: 425–430

Nightingale K, Lin KM, Ravenhill BJ, Davies C, Nobre L, Fielding CA, Ruckova E, Fletcher-Etherington A, Soday L, Nichols H, Sugrue D, Wang ECYY, Moreno P, Umrania Y, Huttlin EL, Antrobus R, Davison AJ, Wilkinson GWGG, Stanton RJ, Tomasec P, et al (2018) High-Definition Analysis of Host Protein Stability during Human Cytomegalovirus Infection Reveals Antiviral Factors and Viral Evasion Mechanisms. Cell Host Microbe 24: 447-460 e11

Nobre L, Nightingale K, Ravenhill BJ, Antrobus R, Soday L, Nichols J, Davies J, Seirafian S, Wang ECY, Davison AJ, Wilkinson GWG, Stanton RJ, Huttlin EL & Weekes MP (2019) Human cytomegalovirus interactome analysis identifies degradation hubs, domain associations and viral protein functions. Elife 8: 1–35

Noor Z, Ahn SB, Baker MS, Ranganathan S & Mohamedali A (2020) Mass spectrometry–based protein identification in proteomics—a review. Brief. Bioinform.

Okada R, Kondo T, Matsuki F, Takata H & Takiguchi M (2008) Phenotypic classification of human CD4+ T cell subsets and their differentiation. Int. Immunol. 20: 1189–1199

Olgun NS, Patel HJ, Stephani R, Lengyel I & Reznik SE (2010) Blockade of endothelin-1 with a novel series of 1,3,6-trisubstituted-2- carboxy-quinol-4-ones controls infection-associated preterm birth. Am. J. Pathol. 177: 1929–1935

Ono K, Matsumori A, Shioi T, Furukawa Y, Sasayama S & Kloner RA (1999) Contribution of endothelin-1 to myocardial injury in a murine model of myocarditis: Acute effects of Bosentan, an endothelin receptor antagonist. Circulation 100: 1823–1829

Padilla BE, Cottrell GS, Roosterman D, Pikios S, Muller L, Steinhoff M & Bunnett NW (2007) Endothelin-converting enzyme-1 regulates endosomal sorting of calcitonin receptor-like receptor and β-arrestins. J. Cell Biol. 179: 981–997

Park I & Lee H-S (2015) EphB/ephrinB signaling in cell adhesion and migration. Mol. Cells 38: 14–9

Parkinson JE & Smith GL (1994) Vaccinia virus gene A36R encodes a M(r) 43-50 K protein on the surface of extracellular enveloped virus. Virology 204: 376–390

Parrish S & Moss B (2007) Characterization of a second vaccinia virus mRNA-decapping enzyme conserved in poxviruses. J Virol 81: 12973–12978

Parrish S, Resch W & Moss B (2007) Vaccinia virus D10 protein has mRNA decapping activity, providing a mechanism for control of host and viral gene expression. Proc Natl Acad Sci U S A 104: 2139–2144

Pascovici D, Handler DCL, Wu JX & Haynes PA (2016) Multiple testing corrections in quantitative proteomics: A useful but blunt tool. Proteomics 16: 2448–2453

Paul F, Pellegrini S & Uzé G (2015) IFNA2: The prototypic human alpha interferon. Gene 567: 132–137

Paulus C, Krauss S & Nevels M (2006) A human cytomegalovirus antagonist of type I IFN-dependent signal transducer and activator of transcription signaling. Proc. Natl. Acad. Sci. U. S. A. 103: 3840–5

Pellett PE & Roizman B (2013) Herpesviridae. In Fields Virology, Knipe DM & Howley PM (eds) pp 1802–1822. Lipincott Williams and Wilkins, Philadelphia

Perez-Caballero D, Zang T, Ebrahimi A, McNatt MW, Gregory DA, Johnson MC & Bieniasz PD (2009) Tetherin Inhibits HIV-1 Release by Directly Tethering Virions to Cells. Cell 139: 499–511

Pérez-Moreno P, Indo S, Niechi I, Huerta H, Cabello P, Jara L, Aguayo F, Varas-Godoy M, Burzio VA & Tapia JC (2019) Endothelin-converting enzyme-1c promotes stem cell traits and aggressiveness in colorectal cancer cells. Mol. Oncol.: 1–16

Perng YC & Lenschow DJ (2018) ISG15 in antiviral immunity and beyond. Nat. Rev. Microbiol. 16: 423–439

Perreira JM, Chin CR, Feeley EM & Brass AL (2013) IFITMs restrict the replication of multiple pathogenic viruses. J. Mol. Biol. 425: 4937–4955

Peters NE, Ferguson BJ, Mazzon M, Fahy AS, Krysztofinska E, Arribas-Bosacoma R, Pearl LH, Ren H & Smith GL (2013) A Mechanism for the Inhibition of DNA-PK-Mediated DNA Sensing by a Virus. PLoS Pathog. 9:

Pichlmair A, Lassnig C, Eberle CA, Gorna MW, Baumann CL, Burkard TR, Burckstummer T, Stefanovic A, Krieger S, Bennett KL, Rulicke T, Weber F, Colinge J, Muller M & Superti-Furga G (2011) IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. Nat Immunol 12: 624–630

Pincus T, Rowe WP & Lilly F (1971) A major genetic locus affecting resistance to infection with murine leukemia viruses. II. Apparent identity to a major locus described for resistance to friend murine leukemia virus. J. Exp. Med. 133: 1234–1241

Poole E, Juss JK, Krishna B, Herre J, Chilvers ER & Sinclair J (2015) Alveolar macrophages isolated directly from human cytomegalovirus (HCMV)-seropositive individuals are sites of HCMV reactivation in vivo. J. Infect. Dis. 211: 1936–1942

Poole E & Sinclair J (2015) Sleepless latency of human cytomegalovirus. Med. Microbiol. Immunol. 204: 421–429

Poole E, Wills M & Sinclair J (2014) Human Cytomegalovirus Latency: Targeting Differences in the Latently Infected Cell with a View to Clearing Latent Infection. New J. Sci. 2014: e313761

Portal D, Rosendorff A & Kieff E (2006) Epstein-Barr nuclear antigen leader protein coactivates transcription through interaction with histone deacetylase 4. Proc Natl Acad Sci U S A 103: 19278–19283

Pott J, Mahlakoiv T, Mordstein M, Duerr CU, Michiels T, Stockinger S, Staeheli P & Hornef MW (2011) IFN-lambda determines the intestinal epithelial antiviral host defense. Proc Natl Acad Sci U S A 108: 7944–7949

R Core Team (2018) R: A language and environment for statistical computing.

Rahm N & Telenti A (2012) The role of tripartite motif family members in mediating susceptibility to HIV-1 infection. Curr Opin HIV AIDS 7: 180–186

Ramirez PW, Famiglietti M, Sowrirajan B, DePaula-Silva A, Rodesch C, Barker E, Bosque A & Planelles V (2014) Downmodulation of CCR7 by HIV-1 Vpu Results in Impaired Migration and Chemotactic Signaling within CD4+ T Cells. Cell Rep. 7: 2019–2030

Ramirez, Sharma, Singh, Stoneham, Vollbrecht & Guatelli (2019) Plasma Membrane-Associated Restriction Factors and Their Counteraction by HIV-1 Accessory Proteins. Cells 8: 1020

Rappsilber J, Mann M & Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. 2: 1896–1906

Ravenhill BJ, Soday L, Houghton J, Antrobus R & Weekes MP (2020) Comprehensive cell surface proteomics defines markers of classical, intermediate and non-classical monocytes. Sci. Rep. 10: 1–11

Reeves MB & Sinclair JH (2013) Circulating dendritic cells isolated from healthy seropositive donors are sites of human cytomegalovirus reactivation in vivo. J. Virol. 87: 10660–10667

Reijerkerk A, Lakeman KAM, Drexhage JAR, Van Het Hof B, Van Wijck Y, Van Der Pol SMA, Kooij G, Geerts D & De Vries HE (2012) Brain endothelial barrier passage by monocytes is controlled by the endothelin system. J. Neurochem. 121: 730–737

Rice AP & Roberts BE (1983) Vaccinia virus induces cellular mRNA degradation. J Virol 47: 529–539

Rice GI, Bond J, Asipu A, Brunette RL, Manfield IW, Carr IM, Fuller JC, Jackson RM, Lamb T, Briggs TA, Ali M, Gornall H, Couthard LR, Aeby A, Attard-Montalto SP, Bertini E, Bodemer C, Brockmann K, Brueton LA, Corry PC, et al (2009) Mutations involved in Aicardi-Goutières syndrome implicate SAMHD1 as regulator of the innate immune response. Nat. Genet. 41: 829–832

Roberts KL & Smith GL (2008) Vaccinia virus morphogenesis and dissemination. Trends Microbiol. 16: 472–479

Roosterman D, Cottrell GS, Padilla BE, Muller L, Eckman CB, Bunnett NW & Steinhoff M (2007) Endothelin-converting enzyme 1 degrades neuropeptides in endosomes to control receptor recycling. Proc. Natl. Acad. Sci. U. S. A. 104: 11838–11843

Rosa A, Chande A, Ziglio S, De Sanctis V, Bertorelli R, Goh SL, McCauley SM, Nowosielska A, Antonarakis SE, Luban J, Santoni FA & Pizzato M (2015) HIV-1 Nef promotes infection by excluding SERINC5 from virion incorporation. Nature 526: 212–7

Rosa C La & Diamond DJ (2012) The immune response to human CMV. Futur. Virol 7: 279–293

Rosengren AT, Nyman TA, Syyrakki S, Matikainen S & Lahesmaa R (2005) Proteomic and transcriptomic characterization of interferon-α-induced human primary T helper cells. Proteomics 5: 371–379

Rosenstein Y, Santana A & Pedraza-Alva G (1999) CD43, a molecule with multiple functions. Immunol. Res. 20: 89–99

Rossi L, Lemoli RM & Goodell MA (2013) Gpr171, a putative P2Y-like receptor, negatively regulates myeloid differentiation in murine hematopoietic progenitors. Exp. Hematol. 41: 102–112

Rusinova I, Forster S, Yu S, Kannan A, Masse M, Cumming H, Chapman R & Hertzog PJ (2013) INTERFEROME v2.0: An updated database of annotated interferon-regulated genes. Nucleic Acids Res. 41: 1040–1046

Saffert RT & Kalejta RF (2006) Inactivating a cellular intrinsic immune defense mediated by Daxx is the mechanism through which the human cytomegalovirus pp71 protein stimulates viral immediate-early gene expression. J Virol 80: 3863–3871

Salazar N, Muñoz D, Hoy J & Lokeshwar BL (2014) Use of shRNA for stable suppression of chemokine receptor expression and function in human cancer cell lines. Methods Mol. Biol. 1172: 209–218

Sander JD & Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. 32: 347–350

Sanjana NE, Shalem O & Zhang F (2014) Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods 11: 783–784

Santini SM, Lapenta C, Logozzi M, Parlato S, Spada M, Di Pucchio T & Belardelli F (2000) Type I interferon as a powerful adjuvant for monocyte-derived dendritic cell development and activity in vitro and in Hu-PBL-SCID mice. J. Exp. Med. 191: 1777–1788

Satheshkumar PS, Anton LC, Sanz P & Moss B (2009) Inhibition of the ubiquitin-proteasome system prevents vaccinia virus DNA replication and expression of intermediate and late genes. J Virol 83: 2469–2479

Savas JN, Stein BD, Wu CC & Yates JR (2011) Mass spectrometry accelerates membrane protein analysis. Trends Biochem. Sci. 36: 388–96

Schlaak JF, Hilkens CMU, Costa-Pereira AP, Strobl B, Aberger F, Frischauf AM & Kerr IM (2002) Cell-type and donor-specific transcriptional responses to interferon-α: Use of customized gene arrays. J. Biol. Chem. 277: 49428–49437

Schneider WM, Chevillotte MD & Rice CM (2014) Interferon-Stimulated Genes: A Complex Web of Host Defenses. Annu. Rev. Immunol. 32: 513–545

Schoggins JJW, Wilson SJS, Panis M, Murphy MMY, Jones CTCT, Bieniasz P, Rice CM & CM R (2011) A diverse array of gene products are effectors of the type I interferon antiviral response. Nature 472: 481–485

Schreiner S & Wodrich H (2013) Virion factors that target Daxx to overcome intrinsic immunity. J. Virol. 87: 10412–10422

Schwabe M, Brini AT, Bosco MC, Rubboli F, Egawa M, Zhao J, Princler GL & Kung HF (1994) Disruption by interferon-α of an autocrine interleukin-6 growth loop in IL-6-dependent U266 myeloma cells by homologous and heterologous down- regulation of the IL-6 receptor α- and β-chains. J. Clin. Invest. 94: 2317–2325

Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M, Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W & Selbach M (2011) Global quantification of mammalian gene expression control. Nature 473: 337–342

Schweizer A, Valdenaire O, Nelbock P, Deuschle U, Dumas Milne Edwards J-B, Stumpf JG & Loffler B-M (1997) Human endothelin-converting enzyme (ECE-1): three isoforms with distinct subcellular localizations. Biochem. J. 328: 871–877

Seirafian S, Prod'Homme V, Sugrue D, Davies J, Fielding C, Tomasec P & Wilkinson GWG (2014) Human cytomegalovirus suppresses Fas expression and function. J. Gen. Virol. 95: 933–939

Senbanjo LT & Chellaiah MA (2017) CD44: A multifunctional cell surface adhesion receptor is a regulator of progression and metastasis of cancer cells. Front. Cell Dev. Biol. 5:

Seo GJ, Kim C, Shin WJ, Sklan EH, Eoh H & Jung JU (2018) TRIM56-mediated monoubiquitination of cGAS for cytosolic DNA sensing. Nat Commun 9: 613

Shah K, Tom Blake J, Huang C, Fischer P & Koo GC (2003) Immunosuppressive effects of a Kv1.3 inhibitor. Cell. Immunol. 221: 100–106

Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Heckl D, Ebert BL, Root DE, Doench JG, Mikkelson T, Heckl D, Ebert BL, Root DE, Doench JG & Zhang F (2014) Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Science 343: 84–87

Sharif MN, Šošić D, Rothlin C V, Kelly E, Lemke G, Olson EN & Ivashkiv LB (2006) Twist mediates suppression of inflammation by type I IFNs and Axl. J. Exp. Med. 203: 1891–1901

Shi Q, Jiang J & Luo G (2013) Syndecan-1 Serves as the Major Receptor for Attachment of Hepatitis C Virus to the Surfaces of Hepatocytes. J. Virol. 87: 6866–6875

Shiow LR, Rosen DB, Brdičková N, Xu Y, An J, Lanier LL, Cyster JG & Matloubian M (2006) CD69 acts downstream of interferon-α/β to inhibit S1P 1 and lymphocyte egress from lymphoid organs. Nature 440: 540–544

Shisler J, Yang C, Walter B, Ware CF & Gooding LR (1997) The adenovirus E3-10.4K/14.5K complex mediates loss of cell surface Fas (CD95) and resistance to Fas-induced apoptosis. J. Virol. 71: 8299–306

Shnayder M, Nachshon A, Krishna B, Poole E, Boshkov A, Binyamin A, Maza I, Sinclair J, Schwartz M & Stern-Ginossar N (2018) Defining the Transcriptional Landscape during Cytomegalovirus Latency with Single-Cell RNA Sequencing. MBio 9: 1–17

Sigma Duration of shRNA knockdown. Available at: https://www.sigmaaldrich.com/life-science/functional-genomics-and-rnai/shrna/learning-center/mission-faqs/set-up-experiments.html#duration_of_knockdown

Silva PN, Soares JA, Brasil BS, Nogueira S V, Andrade AA, de Magalhaes JC, Bonjardim MB, Ferreira PC, Kroon EG, Bruna-Romero O & Bonjardim CA (2006) Differential role played by the MEK/ERK/EGR-1 pathway in orthopoxviruses vaccinia and cowpox biology. Biochem J 398: 83–95

Silvin A & Manel N (2015) Innate immune sensing of HIV infection. Curr. Opin. Immunol. 32: 54–60

Sinzger C, Grefte A, Plachter B, Gouw AS, The TH & Jahn G (1995) Fibroblasts, epithelial cells, endothelial cells and smooth muscle cells are major targets of human cytomegalovirus infection in lung and gastrointestinal tissues. J. Gen. Virol. 76 (Pt 4): 741–750

Sleijfer S, Bannink M, Van Gool AR, Kruit WHJ & Stoter G (2005) Side effects of interferon-α therapy. Pharm. World Sci. 27: 423–431

Smiljanovic B, Grün JR, Biesen R, Schulte-Wrede U, Baumgrass R, Stuhlmüller B, Maslinski W, Hiepe F, Burmester GR, Radbruch A, Häupl T & Grützkau A (2012) The multifaceted balance of TNF-α and type I/II interferon responses in SLE and RA: How monocytes manage the impact of cytokines. J. Mol. Med. 90: 1295–1309

Smith GL, Benfield CTO, Maluquer de Motes C, Mazzon M, Ember SWJ, Ferguson BJ & Sumner RP (2013a) Vaccinia virus immune evasion: Mechanisms, virulence and immunogenicity. J. Gen. Virol. 94: 2367–2392

Smith GL & Law M (2004) The exit of Vaccinia virus from infected cells. Virus Res. 106: 189–197

Smith GL, Talbot-Cooper C & Lu Y (2018) How Does Vaccinia Virus Interfere With Interferon? Adv Virus Res 100: 355–378

Smith S, Weston S, Kellam P & Marsh M (2014) IFITM proteins—cellular inhibitors of viral entry. Curr. Opin. Virol. 4: 71–77

Smith W, Tomasec P, Aicheler R, Loewendorf A, Nemcovicova I, Wang EC, Stanton RJ, Macauley M, Norris P, Willen L, Ruckova E, Nomoto A, Schneider P, Hahn G, Zajonc DM, Ware CF, Wilkinson GW & Benedict CA (2013b) Human cytomegalovirus glycoprotein UL141 targets the TRAIL death receptors to thwart host innate antiviral defenses. Cell Host Microbe 13: 324–335

Smollich M, Götte M, Yip GW, Yong ES, Kersting C, Fischgräbe J, Radke I, Kiesel L & Wülfing P (2007) On the role of endothelin-converting enzyme-1 (ECE-1) and neprilysin in human breast cancer. Breast Cancer Res. Treat. 106: 361–369

Soday L, Lu Y, Albarnaz JD, Davies CTR, Antrobus R, Smith GL & Weekes MP (2019) Quantitative Temporal Proteomic Analysis of Vaccinia Virus Infection Reveals Regulation of Histone Deacetylases by an Interferon Antagonist. Cell Rep. 27: 1920-1933.e7

Soh TK, Davies CTR, Muenzner J, Connor V, Bouton CR, Barrow HG, Smith C, Emmott E, Antrobus R, Graham SC, Weekes MP & Crump CM (2019) Herpes simplex virus-1 pUL56 degrades GOPC to alter the plasma membrane proteome. bioRxiv: 729343

Sondel PM, Buhtoiarov IN & DeSantes K (2003) Pleasant memories: Remembering immune protection while forgetting about graft-versus-host disease. J. Clin. Invest. 112: 25–27

Spencer RH, Sokolov Y, Li H, Takenaka B, Milici AJ, Aiyar J, Nguyen A, Park H, Jap BK, Hall JE, Gutman GA & Chandy KG (1997) Purification, Visualization, and Biophysical Characterization of Kv1.3 Tetramers. J. Biol. Chem. 272: 2389–2395

Stanford SM, Rapini N & Bottini N (2012) Regulation of TCR signalling by tyrosine phosphatases: From immune homeostasis to autoimmunity. Immunology 137: 1–19

Stanton RJ, Baluchova K, Dargan DJ, Cunningham C, Sheehy O, Seirafian S, McSharry BP, Neale ML, Davies JA, Tomasec P, Davison AJ & Wilkinson GWG (2010) Reconstruction of the complete human cytomegalovirus genome in a BAC reveals RL13 to be a potent inhibitor of replication. J. Clin. Invest. 120: 3191–3208

Stoddart CA, Keir ME & McCune JM (2010) IFN-α-induced upregulation of CCR5 leads to expanded HIV tropism in vivo. PLoS Pathog. 6:

Strnadova P, Ren H, Valentine R, Mazzon M, Sweeney TR, Brierley I & Smith GL (2015) Inhibition of Translation Initiation by Protein 169: A Vaccinia Virus Strategy to Suppress Innate and Adaptive Immunity and Alter Virus Virulence. PLoS Pathog 11: e1005151

Stuart JH, Sumner RP, Lu Y, Snowden JS & Smith GL (2016) Vaccinia Virus Protein C6 Inhibits Type I IFN Signalling in the Nucleus and Binds to the Transactivation Domain of STAT2. PLoS Pathog 12: e1005955

Swaney DL & Villén J (2016) Proteomic analysis of protein posttranslational modifications by mass spectrometry. Cold Spring Harb. Protoc. 2016: 207–209

Sylwester A, Mitchell B, Edgar J, Taormina C, Pelte C, Ruchti F, Sleath P, Grabstein K, Hosken N, Kern F, Nelson J & Picker L (2005) Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects. J. Exp. Med. 202: 673–685

Symons JA, Alcami A & Smith GL (1995) Vaccinia virus encodes a soluble type I interferon receptor of novel structure and broad species specificity. Cell 81: 551–560

Tan JT & Surh CD (2006) T cell memory. Curr. Top. Microbiol. Immunol. 311: 85–115

Tapia JC & Niechi I (2019) Endothelin-converting enzyme-1 in cancer aggressiveness. Cancer Lett. 452: 152–157

Tavalai N, Adler M, Scherer M, Riedl Y & Stamminger T (2011) Evidence for a dual antiviral role of the major nuclear domain 10 component Sp100 during the immediate-early and late phases of the human cytomegalovirus replication cycle. J. Virol. 85: 9447–9458

Tavalai N, Papior P, Rechter S & Stamminger T (2008) Nuclear domain 10 components promyelocytic leukemia protein and hDaxx independently contribute to an intrinsic antiviral defense against human cytomegalovirus infection. J Virol 82: 126–137

Taylor-Wiedeman J, Sissons JG, Borysiewicz LK & Sinclair JH (1991) Monocytes are a major site of persistence of human cytomegalovirus in peripheral blood mononuclear cells. J. Gen. Virol. 72: 2059–2064

Taylor RT & Bresnahan WA (2005) Human cytomegalovirus immediate-early 2 gene expression blocks virus-induced beta interferon production. J. Virol. 79: 3873–7

Teale A, Campbell S, Van Buuren N, Magee WC, Watmough K, Couturier B, Shipclark R & Barry M (2009) Orthopoxviruses require a functional ubiquitin-proteasome system for productive replication. J Virol 83: 2099–2108

Terenzi F, Saikia P & Sen GC (2008) Interferon-inducible protein, P56, inhibits HPV DNA replication by binding to the viral protein E1. EMBO J. 27: 3311–3321

The Gene Ontology Consortium (2019) The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. 47: D330–D338

Thompson A, Schäfer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T & Hamon C (2003) Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal. Chem. 75: 1895–1904

Ting L, Rad R, Gygi SP & Haas W (2011) MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. Nat Methods 8: 937–940

le Tortorec A, Willey S & Neil SJD (2011) Antiviral inhibition of enveloped virus release by Tetherin/BST-2: Action and counteraction. Viruses 3: 520–540

Tscharke DC, Reading PC & Smith GL (2002) Dermal infection with vaccinia virus reveals roles for virus proteins not seen using other inoculation routes. J Gen Virol 83: 1977–1986

Tudorache IF, Trusca VG & Gafencu AV (2017) Apolipoprotein E - A Multifunctional Protein with Implications in Various Pathologies as a Result of Its Structural Features. Comput. Struct. Biotechnol. J. 15: 359–365

Turnbull ML, Wise HM, Nicol MQ, Smith N, Dunfee RL, Beard PM, Jagger BW, Ligertwood Y, Hardisty GR, Xiao H, Benton DJ, Coburn AM, Paulo JA, Gygi SP, McCauley JW, Taubenberger JK, Lycett SJ, Weekes MP, Dutia BM & Digard P (2016) Role of the B Allele of Influenza A Virus Segment 8 in Setting Mammalian Host Range and Pathogenicity. J Virol 90: 9263–9284

Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M & Cox J (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods

UNAIDS (2019) Global HIV Statistics. : 6 Available at: https://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf

Unterholzner L, Sumner RP, Baran M, Ren H, Mansur DS, Bourke NM, Randow F, Smith GL & Bowie AG (2011) Vaccinia virus protein C6 is a virulence factor that binds TBK-1 adaptor proteins and inhibits activation of IRF3 and IRF7. PLoS Pathog 7: e1002247

Urban M, Klein M, Britt WJ, Haßfurther E & Mach M (1996) Glycoprotein H of human cytomegalovirus is a major antigen for the neutralizing humoral immune response. J. Gen. Virol. 77: 1537–1547

Valdenaire O, Lepailleur-Enouf D, Egidy G, Thouard A, Barret A, Vranckx R, Tougard C & Michel J (1999) A fourth isoform of endothelin-converting enzyme (ECE-1) is generated from an additional promoter. Molecular cloning and characterization. Eur. J. Biochem. 264: 341–349

Verardi PH, Titong A & Hagen CJ (2012) A vaccinia virus renaissance: New vaccine and immunotherapeutic uses after smallpox eradication. Hum. Vaccines Immunother. 8: 961–970

Verdin E, Dequiedt F & Kasler HG (2003) Class II histone deacetylases: Versatile regulators. Trends Genet. 19: 286–293

Viswanathan K, Smith MS, Malouli D, Mansouri M, Nelson JA & Fruh K (2011) BST2/Tetherin enhances entry of human cytomegalovirus. PLoS Pathog 7: e1002332

Vivier E, Tomasello E, Baratin M, Walzer T & Ugolini S (2008) Functions of natural killer cells. Nat Immunol 9: 503–510

Vladimer GI, Gorna MW & Superti-Furga G (2014) IFITs: Emerging Roles as Key Anti-Viral Proteins. Front Immunol 5: 94

Wali A & Strayer DS (1999) Infection with vaccinia virus alters regulation of cell cycle progression. DNA Cell Biol 18: 837–843

Walker JE, Chen RX, McGee J, Nacey C, Pollard RB, Abedi M, Bauer G, Nolta JA & Anderson JS (2012) Generation of an HIV-1-Resistant Immune System with CD34+ Hematopoietic Stem Cells Transduced with a Triple-Combination Anti-HIV Lentiviral Vector. J. Virol. 86: 5719–5729

Wang B, Kang W, Zuo J, Kang W & Sun Y (2017) The significance of type-I interferons in the pathogenesis and therapy of human immunodeficiency virus 1 infection. Front. Immunol. 8: 1–12

Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES & Sabatini DM (2015) Identification and characterization of essential genes in the human genome. Science 350: 1096–1101

Wang W, Yen H, Chen CH, Jasani N, Soni R, Koscica K & Reznik SE (2010) Prevention of inflammation-associated preterm birth by knockdown of the endothelin-1-matrix metalloproteinase-1 pathway. Mol. Med. 16: 505–512

Wathen MW & Stinski MF (1982) Temporal Patterns of Human Cytomegalovirus Transcription: Mapping the Viral RNAs Synthesized at Immediate Early, Early, and Late Times After Infection. J. Virol. 41: 462–477

Weekes MP, Antrobus R, Lill JR, Duncan LM, Hor S, Lehner PJ, Hör S & Lehner PJ (2010) Comparative analysis of techniques to purify plasma membrane proteins. J.Biomol.Tech. 21: 108–115

Weekes MP, Antrobus R, Talbot S, Hor S, Simecek N, Smith DL, Bloor S, Randow F, Lehner PJ, Hör S, Simecek N, Smith DL, Bloor S, Randow F & Lehner PJ (2012) Proteomic plasma membrane profiling reveals an essential role for gp96 in the cell surface expression of LDLR family members, including the LDL receptor and LRP6. J. Proteome Res. 11: 1475–1484

Weekes MP, Tomasec P, Huttlin EL, Fielding CA, Nusinow D, Stanton RJ, Wang ECY, Aicheler R, Murrell I, Wilkinson GWG, Lehner PJ, Gygi SP, E.C.Y. W, Aicheler R, Murrell I, Wilkinson GWG, Lehner PJ & Gygi SP (2014) Quantitative temporal viromics: An approach to investigate host-pathogen interaction. Cell 157: 1460–1472

De Weerd NA & Nguyen T (2012) The interferons and their receptors--distribution and regulation. Immunol. Cell Biol. 90: 483–491

Whitwell HJ, Worthington J, Blyuss O, Gentry-Maharaj A, Ryan A, Gunu R, Kalsi J, Menon U, Jacobs I, Zaikin A & Timms JF (2020) Improved early detection of ovarian cancer using longitudinal multimarker models. Br. J. Cancer

Wilcock D, Duncan SA, Traktman P, Zhang WH & Smith GL (1999) The vaccinia virus A4OR gene product is a nonstructural, type II membrane glycoprotein that is expressed at the cell surface. J Gen Virol 80: 2137–2148

Wilson SJ, Schoggins JW, Zang T, Kutluay SB, Jouvenet N, Alim MA, Bitzegeio J, Rice CM & Bieniasz PD (2012) Inhibition of HIV-1 particle assembly by 2',3'-cyclic-nucleotide 3'-phosphodiesterase. Cell Host Microbe 12: 585–97

Woodhall DL, Groves IJ, Reeves MB, Wilkinson G & Sinclair JH (2006) Human Daxx-mediated repression of human cytomegalovirus gene expression correlates with a repressive chromatin structure around the major immediate early promoter. J. Biol. Chem. 281: 37652–37660

World Health Organisation WHO fact sheets. Available at: https://www.who.int/news-room/fact-sheets

Wu R, Dephoure N, Haas W, Huttlin EL, Zhai B, Sowa ME & Gygi SP (2011) Correct interpretation of comprehensive phosphorylation dynamics requires normalization by protein expression changes. Mol Cell Proteomics 10: M111 009654

Xu D, Emoto N, Giaid A, Slaughter C, Kaw S, deWit D & Yanagisawa M (1994) ECE-1: A membrane-bound metalloprotease that catalyzes the proteolytic activation of big endothelin-1. Cell 78: 473–485

Yang Z & Bjorkman PJ (2008) Structure of UL18, a peptide-binding viral MHC mimic, bound to a host inhibitory receptor. Proc. Natl. Acad. Sci. 105: 10095–10100

Yang Z, Bruno DP, Martens CA, Porcella SF & Moss B (2010) Simultaneous high-resolution analysis of vaccinia virus and host cell transcriptomes by deep RNA sequencing. Proc Natl Acad Sci U S A 107: 11513–11518

Yang Z, Cao S, Martens CA, Porcella SF, Xie Z, Ma M, Shen B & Moss B (2015) Deciphering Poxvirus Gene Expression by RNA Sequencing and Ribosome Profiling. J. Virol. 89: 6874–6886

Yang Z, Maruri-Avidal L, Sisler J, Stuart CA & Moss B (2013) Cascade regulation of vaccinia virus gene expression is modulated by multistage promoters. Virology 447: 213–220

Yang Z, Reynolds SE, Martens CA, Bruno DP, Porcella SF & Moss B (2011) Expression profiling of the intermediate and late stages of poxvirus replication. J Virol 85: 9899–9908

Yap MW, Nisole S & Stoye JP (2005) A Single Amino Acid Change in the SPRY Domain of Human Trim5α Leads to HIV-1 Restriction. Curr. Biol. 15: 73–78

Yoo NK, Pyo CW, Kim Y, Ahn BY & Choi SY (2008) Vaccinia virus-mediated cell cycle alteration involves inactivation of tumour suppressors associated with Brf1 and TBP. Cell Microbiol 10: 583–592

Yoon HJ, Zhu Z, Gwaltney JM & Elias JA (1999) Rhinovirus regulation of IL-1 receptor antagonist in vivo and in vitro: a potential mechanism of symptom resolution. J. Immunol. 162: 7461–9

Zhang C, Xu Y, Gu J & Schlossman SF (1998) A cell surface receptor defined by a mAb mediates a unique type of cell death similar to oncosis. Proc. Natl. Acad. Sci. U. S. A. 95: 6290–6295

Zhang Q, Liu Z, Mi Z, Li X, Jia P, Zhou J, Yin X, You X, Yu L, Guo F, Ma J, Liang C & Cen S (2011) High-throughput assay to identify inhibitors of Vpu-mediated down-regulation of cell surface BST-2. Antiviral Res. 91: 321–329

Zhang Y, Fonslow BR, Shan B, Baek M-C & Yates JR (2013) Protein analysis by shotgun/bottom-up proteomics. Chem. Rev. 113: 2343–94

Zhao X, Li J, Winkler CA, An P & Guo JT (2019) IFITM genes, variants, and their roles in the control and pathogenesis of viral infections. Front. Microbiol. 10: 1–12

Zheng H, Qian J, Varghese B, Baker DP & Fuchs S (2011) Ligand-Stimulated Downregulation of the Alpha Interferon Receptor: Role of Protein Kinase D2. Mol. Cell. Biol. 31: 710–720

Zheng Q, Hou J, Zhou Y, Yang Y, Xie B & Cao X (2015) Siglec1 suppresses antiviral innate immune response by inducing TBK1 degradation via the ubiquitin ligase TRIM27. Cell Res. 25: 1121–1136

Zhu J, Yan J & Thornhill WB (2012) N-glycosylation promotes the cell surface expression of Kv1.3 potassium channels. FEBS J. 279: 2632–2644

Ziegler-Heitbrock L, Ancuta P, Crowe S, Dalod M, Grau V, Hart DN, Leenen PJM, Liu YJ, MacPherson G, Randolph GJ, Scherberich J, Schmitz J, Shortman K, Sozzani S, Strobl H, Zembala M, Austyn JM & Lutz MB (2010) Nomenclature of monocytes and dendritic cells in blood. Blood 116: 5–7

Zuhair M, Smit GSA, Wallis G, Jabbar F, Smith C, Devleesschauwer B & Griffiths P (2019) Estimation of the worldwide seroprevalence of cytomegalovirus: A systematic review and meta-analysis. Rev. Med. Virol. 29: 1–6

APPENDIX

Appendix Table 1. Oligonucleotides for generation of cell lines

	Name	Target sequence	Oligonucleotides	
			Forward	Reverse
shRNA Knockdown	Control_sh1	GTTATAGGCTCGCAAAAGG	gatccGTTATAGGCTCGCAAAAGGTTCAAGAGACCTTTTGCGAGCCTATAACTTTTTg	aattcAAAAAAGTTATAGGCTCGCAAAAGGTCTCTTGAACCTTTTGCGAGCCTATAACg
	Control_sh2	GGCATATAACTATTTAGGTAT	gatccGGCATATAACTATTTAGGTATTTCAAGAGAATACCTAAATAGTTATATGCCTTTTTTg	aattcAAAAAAGGCATATAACTATTTAGGTATTCTCTTGAAATACCTAAATAGTTATATGCCg
	Control_sh3	CGTGATCTTCACCGACAAGAT	gatccCGTGATCTTCACCGACAAGATTTCAAGAGAATCTTGTCGGTGAAGATCACGTTTTTTg	aattcAAAAAACGTGATCTTCACCGACAAGATTCTCTTGAAATCTTGTCGGTGAAGATCACGg
	Sp100_sh1	CGCTAGGAAGCCAACAAACAA	gatccCGCTAGGAAGCCAACAAACAATTCAAGAGATTGTTTGTTGGCTTCCTAGCGTTTTTTg	aattcAAAAAACGCTAGGAAGCCAACAAACAATCTCTTGAATTGTTTGTTGGCTTCCTAGCGg
	TMEM123_sh1	CCACACAACTCCAGTGCTAAC	gatccCCACACAACTCCAGTGCTAACTTCAAGAGAGTTAGCACTGGAGTTGTGTGGTTTTTTg	aattcAAAAAACCACACAACTCCAGTGCTAACTCTCTTGAAGTTAGCACTGGAGTTGTGTGGg
	TMEM123_sh2	GGGATGGTCTCAACAAATATG	gatccGGGATGGTCTCAACAAATATGTTCAAGAGACATATTTGTTGAGACCATCCCTTTTTTg	aattcAAAAAAGGGATGGTCTCAACAAATATGTCTCTTGAACATATTTGTTGAGACCATCCCg
	ECE1_sh1	GCAGTTCCAGACCTCTACTTT	gatccGCAGTTCCAGACCTCTACTTTTTCAAGAGAAAAGTAGAGGTCTGGAACTGCTTTTTTg	aattcAAAAAAGCAGTTCCAGACCTCTACTTTTCTCTTGAAAAAGTAGAGGTCTGGAACTGCg
	ECE1_sh2	GCCGATGAGAAGTTCATGGAA	gatccGCCGATGAGAAGTTCATGGAATTCAAGAGATTCCATGAACTTCTCATCGGCTTTTTTg	aattcAAAAAAGCCGATGAGAAGTTCATGGAATCTCTTGAATTCCATGAACTTCTCATCGGCg
	ECE1_sh3	CTTCCACAGCCCCCGGAGT	gatccCTTCCACAGCCCCCGGAGTTTCAAGAGAACTCCGGGGGGCTGTGGAAGTTTTTTg	aattcAAAAAACTTCCACAGCCCCCGGAGTTCTCTTGAAACTCCGGGGGCTGTGGAAGg
	ECE1_sh4	GCCTTAAACTTTGGTGGCATA	gatccGCCTTAAACTTTGGTGGCATATTCAAGAGATATGCCACCAAAGTTTAAGGCTTTTTTg	aattcAAAAAAGCCTTAAACTTTGGTGGCATATCTCTTGAATATGCCACCAAAGTTTAAGGCg
	ECE1_sh5	TGTCTATGTCAGTGCCGATTC	gatccTGTCTATGTCAGTGCCGATTCTTCAAGAGAGAATCGGCACTGACATAGACATTTTTTg	aattcAAAAAATGTCTATGTCAGTGCCGATTCTCTCTTGAAGAATCGGCACTGACATAGACAg
	ECE1_sh6	GATCAATGAATCCGAGCCTAT	gatccGATCAATGAATCCGAGCCTATTTCAAGAGAATAGGCTCGGATTCATTGATCTTTTTTg	aattcAAAAAAGATCAATGAATCCGAGCCTATTCTCTTGAAATAGGCTCGGATTCATTGATCg
CRISPR knockout	Control_g1	GTACAGCTAAGTTAAACTCG	caccgGTACAGCTAAGTTAAACTCG	aaacCGAGTTTAACTTAGCTGTACc
	Control_g2	GATGTCCGTTGTAGTCCTCG	caccgGATGTCCGTTGTAGTCCTCG	aaacCGAGGACTACAACGGACATCc
	Sp100_g1	GCAGCCTGTCATCTACACCC	caccgGCAGCCTGTCATCTACACCC	aaacGGGTGTAGATGACAGGCTGCc
	Sp100_g2	TGAGATGGGGAACCCGAAGG	caccgTGAGATGGGGAACCCGAAGG	aaacCCTTCGGGTTCCCCATCTCAc
	ECE1_g1	GGAAACCCGAAAATCAGCCA	caccgGGAAACCCGAAAATCAGCCA	aaacTGGCTGATTTTCGGGTTTCCc
	ECE1_g2	GAACTGGGTGAAGAAGAACG	caccgGAACTGGGTGAAGAAGAACG	aaacCGTTCTTCTTCACCCAGTTCc
	ECE1_g3	CATGGAGCTCAAGATGGAGC	caccgCATGGAGCTCAAGATGGAGC	aaacGCTCCATCTTGAGCTCCATGc
	ECE1_g4	GAGCTCCATGGACCCCACAG	caccgGAGCTCCATGGACCCCACAG	aaacCTGTGGGGTCCATGGAGCTCc
	ECE1_g5	CTGGGGAAGCTGCTGGGCGG	caccgCTGGGGAAGCTGCTGGGCGG	aaacCCGCCCAGCAGCTTCCCCAGc

Appendix Table 2. Proteins downregulated during VACV infection

The 265 proteins which met the 'sensitive' criteria for downregulation are given. This requires them to be downregulated > 2 fold on average across the three replicates, at any time point compared to the 18 h mock ('sensitive' criteria). P-values are given for proteins that additionally met the 'stringent' criteria of being quantified in all three replicates of the time-course, and $p < 0.05$ at the point of maximal downregulation (estimated using a BH-corrected two-tailed t-test).

Gene Symbol	Description	Maximal fold downregulation	P-value for stringently downregulated proteins
PTOV1	Prostate tumor-overexpressed gene 1 protein	0.04	
MID1IP1	Mid1-interacting protein 1	0.10	
FBXO22	F-box only protein 22	0.12	0.008
CITED2	Cbp/p300-interacting transactivator 2	0.13	
CNTLN	Isoform 2 of Centlein	0.13	
CST3	Cystatin-C	0.15	
ZFAND3	AN1-type zinc finger protein 3	0.17	0.009
COL1A1	Collagen alpha-1(I) chain	0.17	0.034
CREB3L2	Cyclic AMP-responsive element-binding protein 3-like protein 2	0.17	0.008
IFIT5	Interferon-induced protein with tetratricopeptide repeats 5	0.17	0.010
SCD	Acyl-CoA desaturase	0.18	
TWIST2	Twist-related protein 2	0.19	0.005
NDFIP1	NEDD4 family-interacting protein 1	0.20	
SCAND1	SCAN domain-containing protein 1	0.22	0.006
SULF2	Extracellular sulfatase Sulf-2	0.22	
FRMD6	FERM domain-containing protein 6	0.22	0.006
TP53	Cellular tumor antigen p53	0.23	0.009
EPB41L5	Band 4.1-like protein 5	0.23	0.011
COL1A2	Collagen alpha-2(I) chain	0.23	0.030
PPDPF	Pancreatic progenitor cell differentiation and proliferation factor	0.23	0.032
AR	Androgen receptor	0.23	0.050
DAZAP2	DAZ-associated protein 2	0.23	0.009
SPARC	SPARC	0.24	0.009
TMEM59	Transmembrane protein 59	0.25	0.018
CDC25B	Isoform 4 of M-phase inducer phosphatase 2	0.25	
TIMP1	Metalloproteinase inhibitor 1	0.25	0.011
FBLN1	Isoform C of Fibulin-1	0.26	
COL3A1	Collagen alpha-1(III) chain	0.26	0.030
ZNF703	Zinc finger protein 703	0.27	0.023
CDKN1A	Cyclin-dependent kinase inhibitor 1	0.27	
SOX4	Transcription factor SOX-4	0.23	
FAM127A	Protein FAM127A	0.28	0.028

Gene Symbol	Description	Maximal fold downregulation	P-value for stringently downregulated proteins
PRRX1	Isoform 2 of Paired mesoderm homeobox protein 1	0.28	
CYBA	Cytochrome b-245 light chain	0.29	
MAP4K2	Mitogen-activated protein kinase kinase kinase kinase 2	0.29	0.013
HMGCR	Isoform 3 of 3-hydroxy-3-methylglutaryl-coenzyme A reductase	0.22	
COL5A2	Collagen alpha-2(V) chain	0.30	0.013
IFIT2	Interferon-induced protein with tetratricopeptide repeats 2	0.30	0.023
ITM2C	Integral membrane protein 2C	0.30	0.021
LUM	Lumican	0.30	0.031
MXRA8	Matrix-remodeling-associated protein 8	0.30	0.008
TUSC1	Tumor suppressor candidate gene 1 protein	0.30	
LDB2	LIM domain-binding protein 2	0.31	
CKS1B	Cyclin-dependent kinases regulatory subunit 1	0.31	0.012
GPC1	Glypican-1	0.31	0.017
MT1X	Metallothionein-1X	0.31	
APP	Amyloid beta A4 protein	0.31	0.012
IGFBP4	Insulin-like growth factor-binding protein 4	0.31	0.027
GINM1	Glycoprotein integral membrane protein 1	0.31	0.008
IFIT1	Interferon-induced protein with tetratricopeptide repeats 1	0.31	0.030
RHOBTB3	Rho-related BTB domain-containing protein 3	0.32	
RARA	Retinoic acid receptor alpha	0.32	
CKS2	Cyclin-dependent kinases regulatory subunit 2	0.32	
ZER1	Protein zer-1 homolog	0.32	0.016
MASP1	Isoform 2 of Mannan-binding lectin serine protease 1	0.32	0.015
MMADHC	Methylmalonic aciduria and homocystinuria type D protein, mitochondrial	0.32	
CXXC5	CXXC-type zinc finger protein 5	0.32	0.008
MRFAP1	MORF4 family-associated protein 1	0.32	0.025
ADAMTSL3	ADAMTS-like protein 3	0.32	
COL5A1	Collagen alpha-1(V) chain	0.33	0.016
SERF2	Small EDRK-rich factor 2	0.33	0.017
FSTL1	Follistatin-related protein 1	0.33	0.008
HDAC5	Isoform 3 of Histone deacetylase 5	0.33	0.013
TRPC4AP	Short transient receptor potential channel 4-associated protein	0.33	
TERT	Telomerase reverse transcriptase	0.33	0.043
COA5	Cytochrome c oxidase assembly factor 5	0.34	0.029
TPGS2	Tubulin polyglutamylase complex subunit 2	0.34	
CREB3L1	Cyclic AMP-responsive element-binding protein 3-like protein 1	0.32	0.023
LRP11	Low-density lipoprotein receptor-related protein 11	0.34	
EPHA4	Ephrin type-A receptor 4	0.34	0.008

Gene Symbol	Description	Maximal fold downregulation	P-value for stringently downregulated proteins
MAGED1	Isoform 2 of Melanoma-associated antigen D1	0.34	0.008
DKK3	Dickkopf-related protein 3	0.35	
AHR	Aryl hydrocarbon receptor	0.35	0.009
EFR3B	Protein EFR3 homolog B	0.31	
CDC42EP3	Cdc42 effector protein 3	0.35	0.008
IFITM3	Interferon-induced transmembrane protein 3	0.35	0.017
RNF4	E3 ubiquitin-protein ligase RNF4	0.26	
PCDHB8	Protocadherin beta-8	0.35	
SEMA5A	Semaphorin-5A	0.36	
PCDHGA6	Protocadherin gamma-A6	0.36	
HLA-C	HLA class I histocompatibility antigen, Cw-4 alpha chain	0.36	
IGFBP7	Insulin-like growth factor-binding protein 7	0.36	0.024
TCF4	Isoform SEF2-1D of Transcription factor 4	0.36	0.017
LAPTM4A	Lysosomal-associated transmembrane protein 4A	0.36	
TCP11L2	T-complex protein 11-like protein 2	0.36	
ITM2B	Integral membrane protein 2B	0.36	0.026
IFFO1	Isoform 5 of Intermediate filament family orphan 1	0.36	0.026
TRIM13	Isoform 3 of E3 ubiquitin-protein ligase TRIM13	0.36	
OLFML3	Olfactomedin-like protein 3	0.36	0.011
FAM168B	Myelin-associated neurite-outgrowth inhibitor	0.37	0.050
COL11A1	Collagen alpha-1(XI) chain	0.37	0.013
FADS2	Fatty acid desaturase 2	0.37	0.046
TCF4	Transcription factor 4	0.36	
XPA	DNA repair protein complementing XP-A cells	0.37	0.009
SCRIB	Isoform 3 of Protein scribble homolog	0.37	0.009
COL6A2	Collagen alpha-2(VI) chain	0.37	0.011
CDKN3	Cyclin-dependent kinase inhibitor 3	0.37	
CALCOCO2	Isoform 4 of Calcium-binding and coiled-coil domain-containing protein 2	0.38	0.009
DNAJC15	DnaJ homolog subfamily C member 15	0.38	
SDC4	Syndecan-4	0.38	0.009
GAS1	Growth arrest-specific protein 1	0.38	0.009
ECM1	Isoform 4 of Extracellular matrix protein 1	0.38	0.033
TNFRSF1A	Tumor necrosis factor receptor superfamily member 1A	0.38	0.029
IFFO2	Intermediate filament family orphan 2	0.37	0.033
PCOLCE	Procollagen C-endopeptidase enhancer 1	0.38	0.048
CHCHD2	Coiled-coil-helix-coiled-coil-helix domain-containing protein 2	0.39	0.009
HERPUD1	Homocysteine-responsive endoplasmic reticulum-resident ubiquitin-like domain member 1 protein	0.39	0.028
COL5A3	Collagen alpha-3(V) chain	0.39	
TCHP	Trichoplein keratin filament-binding protein	0.39	0.014

Gene Symbol	Description	Maximal fold downregulation	P-value for stringently downregulated proteins
RFWD3	E3 ubiquitin-protein ligase RFWD3	0.39	
NR3C1	Isoform Alpha-2 of Glucocorticoid receptor	0.40	0.008
STC2	Stanniocalcin-2	0.40	0.017
TNRC18	Trinucleotide repeat-containing gene 18 protein	0.38	0.010
RASA2	Ras GTPase-activating protein 2	0.40	0.008
AXL	Tyrosine-protein kinase receptor UFO	0.40	0.026
CLSTN1	Calsyntenin-1	0.40	
ZFAND5	AN1-type zinc finger protein 5	0.41	
DCN	Decorin	0.41	0.024
EPHB4	Ephrin type-B receptor 4	0.41	0.012
SIX2	Homeobox protein SIX2	0.41	
SIPA1L2	Signal-induced proliferation-associated 1-like protein 2	0.41	
COL2A1	Collagen alpha-1(II) chain	0.41	
DENND4C	DENN domain-containing protein 4C	0.37	
LOXL1	Lysyl oxidase homolog 1	0.41	0.020
EPHB2	Ephrin type-B receptor 2	0.41	0.017
TXNIP	Thioredoxin-interacting protein	0.41	0.015
PDGFRL	Platelet-derived growth factor receptor-like protein	0.41	
PJA2	E3 ubiquitin-protein ligase Praja-2	0.41	0.021
KIAA0922	Isoform 4 of Transmembrane protein 131-like	0.41	0.032
COL6A2	Isoform 2C2A of Collagen alpha-2(VI) chain	0.42	0.032
HSD17B14	17-beta-hydroxysteroid dehydrogenase 14	0.26	
PJA1	E3 ubiquitin-protein ligase Praja-1	0.42	
TRAIP	E3 ubiquitin-protein ligase TRAIP	0.42	
TRIM5	Tripartite motif-containing protein 5	0.42	0.018
EFEMP1	EGF-containing fibulin-like extracellular matrix protein 1	0.42	0.009
TMEM248	Transmembrane protein 248	0.42	
TCEAL7	Transcription elongation factor A protein-like 7	0.42	0.030
PRRX1	Paired mesoderm homeobox protein 1	0.42	0.013
KRCC1	Lysine-rich coiled-coil protein 1	0.42	
SUSD5	Sushi domain-containing protein 5	0.42	0.011
HAPLN3	Hyaluronan and proteoglycan link protein 3	0.43	
EBF1	Transcription factor COE1	0.43	
CDR2	Cerebellar degeneration-related protein 2	0.43	0.008
HDAC4	Histone deacetylase 4	0.43	0.016
FGFR1	Isoform 21 of Fibroblast growth factor receptor 1	0.43	0.027
RTFDC1	Protein RTF2 homolog	0.43	0.009
BCAT1	Branched-chain-amino-acid aminotransferase, cytosolic	0.43	
PSAP	Isoform Sap-mu-6 of Prosaposin	0.43	
CDCA7	Isoform 2 of Cell division cycle-associated protein 7	0.43	

Gene Symbol	Description	Maximal fold downregulation	P-value for stringently downregulated proteins
AHRR	Isoform 2 of Aryl hydrocarbon receptor repressor	0.43	
IGFBP6	Insulin-like growth factor-binding protein 6	0.43	0.009
SDHAF2	Succinate dehydrogenase assembly factor 2, mitochondrial	0.43	
THBS2	Thrombospondin-2	0.44	0.011
TSC22D3	Isoform 2 of TSC22 domain family protein 3	0.44	0.035
PHLDA3	Pleckstrin homology-like domain family A member 3	0.44	0.041
TSEN34	tRNA-splicing endonuclease subunit Sen34	0.44	0.025
NLGN1	Neuroligin-1	0.44	
ZNFX1	NFX1-type zinc finger-containing protein 1	0.44	0.019
RNF114	E3 ubiquitin-protein ligase RNF114	0.44	0.040
CRIM1	Cysteine-rich motor neuron 1 protein	0.44	0.012
SPOCK1	Testican-1	0.44	0.028
JDP2	Isoform 2 of Jun dimerization protein 2	0.44	
ARL6IP4	Isoform 3 of ADP-ribosylation factor-like protein 6-interacting protein 4	0.44	
LONP2	Lon protease homolog 2, peroxisomal	0.45	0.011
MLF2	Myeloid leukemia factor 2	0.45	
CCDC28B	Coiled-coil domain-containing protein 28B	0.45	
ATP6AP2	Renin receptor	0.45	0.046
CSF1	Macrophage colony-stimulating factor 1	0.45	
HLA-C	HLA class I histocompatibility antigen, Cw-12 alpha chain	0.45	0.024
COL12A1	Collagen alpha-1(XII) chain	0.45	0.021
FLT1	Isoform 2 of Vascular endothelial growth factor receptor 1	0.45	0.015
TWIST1	Twist-related protein 1	0.45	
EPHB3	Ephrin type-B receptor 3	0.45	0.009
PIAS3	E3 SUMO-protein ligase PIAS3	0.45	
ARID5B	AT-rich interactive domain-containing protein 5B	0.45	0.042
TCF3	Isoform 3 of Transcription factor E2-alpha	0.45	0.026
C15orf39	Uncharacterized protein C15orf39	0.45	
SOGA1	Isoform 2 of Protein SOGA1	0.46	0.016
FER	Tyrosine-protein kinase Fer	0.46	0.016
CTSL	Cathepsin L1	0.46	0.026
TMEM168	Transmembrane protein 168	0.46	0.031
FOXP1	Forkhead box protein P1	0.46	0.009
CD248	Endosialin	0.46	0.008
LOX	Protein-lysine 6-oxidase	0.46	
ID1	DNA-binding protein inhibitor ID-1	0.46	
PPP2R3C	Serine/threonine-protein phosphatase 2A regulatory subunit B'' subunit gamma	0.46	
PEG10	Retrotransposon-derived protein PEG10	0.46	
LRRC42	Leucine-rich repeat-containing protein 42	0.46	0.046

Gene Symbol	Description	Maximal fold downregulation	P-value for stringently downregulated proteins
PDGFRA	Platelet-derived growth factor receptor alpha	0.46	0.013
RNF216	Isoform 2 of E3 ubiquitin-protein ligase RNF216	0.46	
TBKBP1	TANK-binding kinase 1-binding protein 1	0.46	
TSPYL4	Testis-specific Y-encoded-like protein 4	0.46	0.008
TCF7L1	Transcription factor 7-like 1	0.46	0.015
MKX	Homeobox protein Mohawk	0.47	
CCBE1	Collagen and calcium-binding EGF domain-containing protein 1	0.47	
IFIT3	Interferon-induced protein with tetratricopeptide repeats 3	0.47	0.039
SOX6	Transcription factor SOX-6	0.47	0.016
ZBED6	Zinc finger BED domain-containing protein 6	0.47	
SDC2	Syndecan-2	0.47	0.026
ZC3HAV1	Isoform 3 of Zinc finger CCCH-type antiviral protein 1	0.47	
TSPAN5	Tetraspanin-5	0.47	
CCDC85B	Coiled-coil domain-containing protein 85B	0.47	0.042
TMEM184B	Transmembrane protein 184B	0.47	
FOXF1	Forkhead box protein F1	0.47	
UBE2C	Ubiquitin-conjugating enzyme E2 C	0.47	
VANGL1	Vang-like protein 1	0.47	0.011
SDC1	Syndecan-1	0.47	0.038
CSNK1A1	Isoform 3 of Casein kinase I isoform alpha	0.47	
PCDHGB4	Protocadherin gamma-B4	0.47	
COL6A1	Collagen alpha-1(VI) chain	0.47	0.009
S100PBP	S100P-binding protein	0.48	
APCDD1	Protein APCDD1	0.48	
FAM168A	Protein FAM168A	0.48	
ANAPC11	Anaphase-promoting complex subunit 11	0.48	
FBXO18	Isoform 2 of F-box DNA helicase 1	0.48	0.009
C1S	Complement C1s subcomponent	0.48	
HIF1A	Isoform 3 of Hypoxia-inducible factor 1-alpha	0.44	
SLIT3	Isoform 4 of Slit homolog 3 protein	0.48	0.045
TCEAL9	Transcription elongation factor A protein-like 9	0.36	
APLP2	Amyloid-like protein 2	0.48	0.025
USP1	Ubiquitin carboxyl-terminal hydrolase 1	0.48	
NEO1	Neogenin	0.48	0.013
ADAMTS5	A disintegrin and metalloproteinase with thrombospondin motifs 5	0.48	
QSOX1	Sulfhydryl oxidase 1	0.48	
SQLE	Squalene monooxygenase	0.48	
PCDHGB5	Protocadherin gamma-B5	0.49	0.009
SLIT2	Slit homolog 2 protein	0.49	0.046
ZNF503	Zinc finger protein 503	0.48	

Gene Symbol	Description	Maximal fold downregulation	P-value for stringently downregulated proteins
LAMB2	Laminin subunit beta-2	0.49	0.026
PPRC1	Peroxisome proliferator-activated receptor gamma coactivator-related protein 1	0.48	
ALPK2	Alpha-protein kinase 2	0.49	0.011
MEX3B	RNA-binding protein MEX3B	0.49	
DENND6A	Protein DENND6A	0.49	
PRELP	Prolargin	0.49	
NBPF19	Neuroblastoma breakpoint family member 19	0.49	
TMEM132A	Isoform 2 of Transmembrane protein 132A	0.49	
RNF5	E3 ubiquitin-protein ligase RNF5	0.49	
CFLAR	CASP8 and FADD-like apoptosis regulator	0.46	
PRNP	Major prion protein	0.49	0.048
FLT1	Vascular endothelial growth factor receptor 1	0.49	0.012
LRRC17	Leucine-rich repeat-containing protein 17	0.45	0.028
OLFM2	Noelin-2	0.46	
FST	Follistatin	0.49	0.012
LAMA4	Laminin subunit alpha-4	0.49	0.038
PHLDA2	Pleckstrin homology-like domain family A member 2	0.49	0.033
MORF4L1	Mortality factor 4-like protein 1	0.49	0.030
ADM	ADM	0.49	
CDCA7L	Cell division cycle-associated 7-like protein	0.49	
IL6ST	Interleukin-6 receptor subunit beta	0.50	0.008
SPATA13	Isoform 6 of Spermatogenesis-associated protein 13	0.50	
PNMA1	Paraneoplastic antigen Ma1	0.49	
CXCL12	Isoform Delta of Stromal cell-derived factor 1	0.50	0.014
HBP1	HMG box-containing protein 1	0.37	
PDGFC	Platelet-derived growth factor C	0.46	
VWF	von Willebrand factor	0.40	
CDC7	Cell division cycle 7-related protein kinase	0.47	
OLFML2A	Olfactomedin-like protein 2A	0.48	
TMSB4X	Thymosin beta-4	0.49	
F13A1	Coagulation factor XIII A chain	0.46	
ESPL1	Separin	0.49	
ZNF282	Zinc finger protein 282	0.46	
TOB2	Protein Tob2	0.49	
CDK10	Cyclin-dependent kinase 10	0.50	
NFE2L2	Nuclear factor erythroid 2-related factor 2	0.44	
THBS3	Thrombospondin-3	0.42	

Appendix Table 3. Proteins upregulated during VACV infection

The 265 proteins which met the 'sensitive' criteria for upregulation are given. This requires them to be upregulated > 2 fold on average across the three replicates, at any time point compared to the 18h mock ('sensitive' criteria). P-values are given for proteins that additionally met the 'stringent' criteria of being quantified in all three replicates of the time-course, and $p < 0.05$ at the point of maximal upregulation (estimated using a BH-corrected two-tailed t-test).

Gene Symbol	Description	Maximal fold upregulation	P-value for stringently upregulated proteins
ZMYND19	Zinc finger MYND domain-containing protein 19	57.19	
IGKC	Ig kappa chain C region	38.99	
PER1	Period circadian protein homolog 1	33.13	
ZFP64	Zinc finger protein 64 homolog, isoforms 3 and 4	30.57	
NR4A1	Isoform 2 of Nuclear receptor subfamily 4 group A member 1	26.56	
NRIP1	Nuclear receptor-interacting protein 1	25.55	
ABLIM1	Actin-binding LIM protein 1	13.36	
TMEM26	Transmembrane protein 26	13.32	
ATXN7L3	Isoform 2 of Ataxin-7-like protein 3	10.97	
ZNF195	Zinc finger protein 195	9.38	
CD58	Lymphocyte function-associated antigen 3	7.65	
EXOC3L4	Exocyst complex component 3-like protein 4	6.98	
DNAH10	Dynein heavy chain 10, axonemal	5.94	
TBC1D32	Isoform 2 of Protein broad-minded	5.28	
EGR1	Early growth response protein 1	5.24	0.036
MDM4	Protein Mdm4	4.93	
PSTPIP2	Proline-serine-threonine phosphatase-interacting protein 2	4.59	
B9D1	B9 domain-containing protein 1	4.51	
IL1R1	Interleukin-1 receptor type 1	4.48	
ALG13	Putative bifunctional UDP-N-acetylglucosamine transferase and deubiquitinase ALG13	4.42	
NFKBIL1	NF-kappa-B inhibitor-like protein 1	4.34	
FOS	Proto-oncogene c-Fos	4.15	
RANGRF	Ran guanine nucleotide release factor	3.97	
STAB1	Stabilin-1	3.85	
ZDHHC18	Palmitoyltransferase ZDHHC18	3.84	
ELOVL4	Elongation of very long chain fatty acids protein 4	3.80	
APOE	Apolipoprotein E	3.72	0.015
SVBP	Small vasohibin-binding protein	3.44	
FSIP2	Fibrous sheath-interacting protein 2	3.25	
C2CD2L	Isoform 2 of C2 domain-containing protein 2-like	3.03	
MAFK	Transcription factor MafK	2.99	
DISC1	Disrupted in schizophrenia 1 protein	2.93	
DUSP1	Dual specificity protein phosphatase 1	2.93	

Gene Symbol	Description	Maximal fold upregulation	P-value for stringently upregulated proteins
WNT2B	Protein Wnt-2b	2.91	
PTGS2	Prostaglandin G/H synthase 2	2.90	
MSI1	RNA-binding protein Musashi homolog 1	2.80	
TNFRSF10D	Tumor necrosis factor receptor superfamily member 10D	2.79	0.060
CLK1	Isoform 3 of Dual specificity protein kinase CLK1	2.76	0.026
ID1	DNA-binding protein inhibitor ID-1	2.75	
TAGLN3	Transgelin-3	2.71	
F13A1	Coagulation factor XIII A chain	2.69	
TNC	Isoform 6 of Tenascin	2.69	
A2M	Alpha-2-macroglobulin	2.69	
LMBRD2	LMBR1 domain-containing protein 2	2.64	
PARP12	Poly [ADP-ribose] polymerase 12	2.55	
CSAD	Isoform 3 of Cysteine sulfinic acid decarboxylase	2.53	
APOA4	Apolipoprotein A-IV	2.51	
PLIN2	Perilipin-2	2.50	
LOXL1	Lysyl oxidase homolog 1	2.47	
ID3	DNA-binding protein inhibitor ID-3	2.46	
BNIP3	BCL2/adenovirus E1B 19 kDa protein-interacting protein 3	2.44	
CDC45	Isoform 3 of Cell division control protein 45 homolog	2.43	
PTTG1	Securin	2.43	
SDC3	Syndecan-3	2.34	
VRK3	Inactive serine/threonine-protein kinase VRK3	2.18	
CTGF	Connective tissue growth factor	2.16	
CBX2	Chromobox protein homolog 2	2.16	
EVC	Ellis-van Creveld syndrome protein	2.16	
ETS2	Protein C-ets-2	2.15	
RASSF1	Isoform C of Ras association domain-containing protein 1	2.12	
FGB	Fibrinogen beta chain	2.12	
GSTT1	Glutathione S-transferase theta-1	2.11	
PLG	Plasminogen	2.10	
HMCN1	Hemicentin-1	2.10	
PZP	Pregnancy zone protein	2.09	
ZSCAN20	Zinc finger and SCAN domain-containing protein 20	2.08	
STC1	Stanniocalcin-1	2.08	
COQ10B	Coenzyme Q-binding protein COQ10 homolog B, mitochondrial	2.07	
AKAP13	Isoform 3 of A-kinase anchor protein 13	2.06	
TPGS1	Tubulin polyglutamylase complex subunit 1	2.01	

Appendix Table 4. Proteomic and transcriptional assignment of VACV protein classes

The proteomic data was compared to previously assigned transcriptional classes (Yang *et al*, 2010, 2011, 2015), and proteomic classes (Croft *et al*, 2015). Croft *et al*. analysed VACV infection in two independent time-courses, with one from 0.5–9.5 hpi, sampled at 3 h intervals and a second time-course from 0.5–8.5 hpi sampled at 2 h intervals, with each generating four temporal classes of viral proteins. As in some cases the two classifications were not identical, the dataset presented here was compared to the 47 proteins that produced concordant temporal classes in the two time-courses.

Viral Gene	Temporal class from this proteomic study	Previously assigned transcriptional class	Fold Change with AraC (6 hpi)	Functional Category	Reference for function	Croft et al consensus
A35	1	E1.1	0.88	Host interaction	Yang et al. (2010)	
A37	1	E1.1	0.95	Unknown	Yang et al. (2010)	
A48	1	E1.1	1.35	DNA replication	Yang et al. (2010)	1
B8	1	E1.1	1.43	Host interaction	Yang et al. (2010)	
C11	1	E1.1	1.52	Host interaction	Yang et al. (2010)	1
E5	1	E1.1	1.08	Unknown	Yang et al. (2010)	
F11	1	E1.1	0.99	Host interaction	Yang et al. (2010)	1
F15	1	E1.1	0.71	Unknown	Yang et al. (2010)	
K1	1	E1.1	0.84	Host interaction	Yang et al. (2010)	1
N2	1	E1.1	1.12	Host interaction	Ferguson et al. (2013)	
VACWR018	1	E1.1	1.50	Unknown	Yang et al. (2010)	
A41	1	E1.2	1.58	Host interaction	Yang et al. (2010)	
A47	1	E1.2	0.91	Unknown	Yang et al. (2010)	
A52	1	E1.2	1.15	Host interaction	Yang et al. (2010)	
B11	1	E1.2	0.76	Unknown	Yang et al. (2010)	
B15	1	E1.2	0.90	Host interaction	Chen et al. (2008)	4
F3	1	E1.2	1.11	Host interaction	Yang et al. (2010)	
M2	1	E1.2	1.32	Host interaction	Yang et al. (2010)	
A46	2	E1.1	1.25	Host interaction	Yang et al. (2010)	1
A51	2	E1.1	1.09	Unknown	Yang et al. (2010)	1
A8	2	E1.1	0.98	Transcription	Yang et al. (2010)	
B12	2	E1.1	1.06	Unknown	Yang et al. (2010)	
B13	2	E1.1	1.07	Host interaction	Yang et al. (2010)	2
B19	2	E1.1	1.10	Host interaction	Yang et al. (2010)	
C1	2	E1.1	1.20	Unknown	Yang et al. (2010)	
C6	2	E1.1	1.10	Host interaction	Unterholzner et al. (2011)	2
D9	2	E1.1	0.70	Host interaction	Yang et al. (2010)	
K7	2	E1.1	0.76	Host interaction	Yang et al. (2010)	
O1	2	E1.1	1.01	Host interaction	Schweneker et al. (2012)	1
A20	2	E1.2	1.05	DNA replication	Yang et al. (2010)	
A23	2	E1.2	0.92	Transcription	Yang et al. (2010)	
A31	2	E1.2	0.98	Unknown	Yang et al. (2010)	2
A40	2	E1.2	1.23	Host interaction	Yang et al. (2010)	2
A44	2	E1.2	0.93	Host interaction	Yang et al. (2010)	
B18	2	E1.2	0.89	Host interaction	Colamonici et al. (1995)	
B6	2	E1.2	1.11	Unknown	Yang et al. (2010)	
C4	2	E1.2	1.11	Host interaction	Ember et al. (2012)	
C5	2	E1.2	1.14	DNA replication	Liu et al. (2018)	1
C7	2	E1.2	0.93	Host interaction	Yang et al. (2010)	
C9	2	E1.2	1.16	Host interaction	Liu & Moss (2018)	
D4	2	E1.2	1.17	DNA replication	Yang et al. (2010)	
F1	2	E1.2	0.53	Host interaction	Yang et al. (2010)	

Viral Gene	Temporal class from this proteomic study	Previously assigned transcriptional class	Fold Change with AraC (6 hpi)	Functional Category	Reference for function	Croft et al consensus
F12	2	E1.2	1.03	Virion association	Yang et al. (2010)	
F16	2	E1.2	1.14	Unknown	Yang et al. (2010)	
F4	2	E1.2	1.14	DNA replication	Yang et al. (2010)	1
F7	2	E1.2	0.88	Unknown	Yang et al. (2010)	
G5	2	E1.2	1.07	DNA replication	Yang et al. (2010)	
I4	2	E1.2	1.09	DNA replication	Yang et al. (2010)	
J2	2	E1.2	1.17	DNA replication	Yang et al. (2010)	
M1	2	E1.2	1.03	Host interaction	Ryerson et al. (2017)	1
VACWR011/ VACWR208	2	E1.2	1.16	Truncated	Yang et al. (2010)	
VACWR169	2	E1.2	1.14	Host interaction	Strnadova et al. (2015)	
A2	2	I	0.11	Transcription	Yang et al. (2010)	
A43	2	I	0.28	Host interaction	Yang et al. (2010)	
C13	2	I	0.19	Unknown	Yang et al. (2010)	
D10	2	I	0.31	Host interaction	Yang et al. (2010)	
A36	3	E1.1	0.83	Virion association	Yang et al. (2010)	
C23/B29	3	E1.1	1.05	Host interaction	Yang et al. (2010)	
E3	3	E1.1	1.19	Host interaction	Yang et al. (2010)	2
E4	3	E1.1	1.12	Transcription	Yang et al. (2010)	
H5	3	E1.1	0.96	Transcription	Yang et al. (2010)	
I3	3	E1.1	1.19	DNA replication	Yang et al. (2010)	
K3	3	E1.1	0.51	Host interaction	Yang et al. (2010)	
A18	3	E1.2	0.32	Transcription	Yang et al. (2010)	
A24	3	E1.2	0.66	Transcription	Yang et al. (2010)	2
A29	3	E1.2	0.95	Transcription	Yang et al. (2010)	
A49	3	E1.2	0.78	Host interaction	Mansur et al. (2013)	
A5	3	E1.2	0.83	Transcription	Yang et al. (2010)	
A50	3	E1.2	0.82	DNA replication	Yang et al. (2010)	
A55	3	E1.2	0.87	Host interaction	Yang et al. (2010)	
B1	3	E1.2	0.97	DNA replication	Yang et al. (2010)	
B17	3	E1.2	0.88	Unknown	Yang et al. (2010)	
B2	3	E1.2	1.07	Host interaction	Eaglesham et al. (2019)	
C12	3	E1.2	0.93	Host interaction	Yang et al. (2010)	
C2	3	E1.2	0.94	Host interaction	Yang et al. (2010)	
D1	3	E1.2	1.05	Transcription	Yang et al. (2010)	2
D12	3	E1.2	1.09	Transcription	Yang et al. (2010)	
D5	3	E1.2	0.86	DNA replication	Yang et al. (2010)	
D7	3	E1.2	0.98	Transcription	Yang et al. (2010)	
E1	3	E1.2	0.82	Transcription	Yang et al. (2010)	
E9	3	E1.2	1.12	DNA replication	Yang et al. (2010)	2
F2	3	E1.2	1.15	DNA replication	Yang et al. (2010)	
F6	3	E1.2	0.52	Unknown	Yang et al. (2010)	
G2	3	E1.2	0.55	Transcription	Yang et al. (2010)	
G5.5	3	E1.2	0.96	Transcription	Yang et al. (2010)	
J3	3	E1.2	1.03	Transcription	Yang et al. (2010)	
J4	3	E1.2	0.93	Transcription	Yang et al. (2010)	2
J6	3	E1.2	0.74	Transcription	Yang et al. (2010)	2
N1	3	E1.2	0.52	Host interaction	Yang et al. (2010)	
VACWR013	3	E1.2	1.54	Host interaction	Yang et al. (2010)	
A1	3	I	0.17	Transcription	Yang et al. (2010)	
A12	3	I	0.35	Virion association	Yang et al. (2010)	3
A30.5	3	I	0.58	Virion association	Maruri-Avidal et al. (2013)	
A6	3	I	0.19	Virion association	Yang et al. (2010)	3
E7	3	I	0.25	Unknown	Yang et al. (2010)	3
G8	3	I	0.19	Transcription	Yang et al. (2010)	

Viral Gene	Temporal class from this proteomic study	Previously assigned transcriptional class	Fold Change with AraC (6 hpi)	Functional Category	Reference for function	Croft et al consensus
H7	3	I	0.14	Virion association	Yang et al. (2010)	
I6	3	I	0.37	Virion association	Yang et al. (2010)	
K2	3	I	0.42	Host interaction	Yang et al. (2010)	
K4	3	I	0.29	DNA replication	Yang et al. (2010)	
A26	3	L	0.80	Virion association	Yang et al. (2010)	
C3	3	L	0.31	Host interaction	Yang et al. (2010)	4
C8	3	L	0.43	Unknown	Yang et al. (2010)	4
A33	4	E1.1	0.82	Virion association	Yang et al. (2010)	
A4	4	E1.2	0.82	Virion association	Yang et al. (2010)	
A56	4	E1.2	0.62	Virion association	Yang et al. (2010)	4
B5	4	E1.2	0.63	Virion association	Yang et al. (2010)	
C10	4	E1.2	0.96	Host interaction	Yang et al. (2010)	2
C22/B28	4	E1.2	0.55	Truncated	Yang et al. (2010)	
E2	4	E1.2	0.80	Virion association	Yang et al. (2010)	
F5	4	E1.2	0.78	Unknown	Yang et al. (2010)	4
F8	4	E1.2	0.66	Unknown	Yang et al. (2010)	
A15	4	I	0.65	Virion association	Yang et al. (2010)	
A16	4	I	0.67	Virion association	Yang et al. (2010)	
A19	4	I	0.36	Virion association	Satheshkumar et al. (2013)	3
A22	4	I	0.80	DNA replication	Yang et al. (2010)	
A25	4	I	0.68	Virion association (truncated)	Yang et al. (2010)	
A27	4	I	0.76	Virion association	Yang et al. (2010)	
A3	4	I	0.72	Virion association	Yang et al. (2010)	3
A30	4	I	0.54	Virion association	Yang et al. (2010)	
A32	4	I	0.33	Virion association	Yang et al. (2010)	
A34	4	I	0.41	Virion association	Yang et al. (2010)	
A38	4	I	0.16	Virion association	Yang et al. (2010)	
A42	4	I	0.54	Virion association	Yang et al. (2010)	
B16	4	I	0.35	Host interaction	Yang et al. (2010)	
D11	4	I	0.46	Transcription	Yang et al. (2010)	
D13	4	I	0.33	Virion association	Yang et al. (2010)	
D6	4	I	0.44	Transcription	Yang et al. (2010)	
D8	4	I	0.53	Virion association	Yang et al. (2010)	4
E11	4	I	0.35	Virion association	Yang et al. (2010)	
E6	4	I	0.38	Virion association	Yang et al. (2010)	3
E8	4	I	0.40	Virion association	Yang et al. (2010)	
F13	4	I	0.47	Virion association	Yang et al. (2010)	4
G4	4	I	0.29	Virion association	Yang et al. (2010)	4
H1	4	I	0.48	Transcription	Yang et al. (2010)	3
H3	4	I	0.68	Virion association	Yang et al. (2010)	4
I1	4	I	0.19	Virion association	Yang et al. (2010)	
I8	4	I	0.62	Transcription	Yang et al. (2010)	
J1	4	I	0.42	Virion association	Yang et al. (2010)	3
L4	4	I	0.67	Transcription	Yang et al. (2010)	3
O2	4	I	0.56	Unknown	Yang et al. (2010)	
A10	4	L	0.74	Virion association	Yang et al. (2010)	3
A11	4	L	0.53	Virion association	Yang et al. (2010)	4
A13	4	L	0.51	Virion association	Yang et al. (2010)	
A17	4	L	0.55	Virion association	Yang et al. (2010)	3
A2.5	4	L	0.65	Virion association	Yang et al. (2010)	
A21	4	L	0.74	Virion association	Yang et al. (2010)	
A28	4	L	0.84	Virion association	Yang et al. (2010)	
A39	4	L	0.41	Truncated	Yang et al. (2010)	
A45	4	L	0.63	Virion association	Yang et al. (2010)	

Viral Gene	Temporal class from this proteomic study	Previously assigned transcriptional class	Fold Change with AraC (6 hpi)	Functional Category	Reference for function	Croft et al consensus
A7	4	L	0.47	Transcription	Yang et al. (2010)	
A9	4	L	0.58	Virion association	Yang et al. (2010)	
B7	4	L	0.50	Host interaction	Yang et al. (2010)	4
D2	4	L	0.56	Virion association	Yang et al. (2010)	
D3	4	L	0.69	Virion association	Yang et al. (2010)	
E10	4	L	0.56	Virion association	Yang et al. (2010)	4
F10	4	L	0.65	Virion association	Yang et al. (2010)	
F17	4	L	0.61	Virion association	Yang et al. (2010)	3
F9	4	L	0.72	Virion association	Yang et al. (2010)	
G1	4	L	0.52	Virion association	Yang et al. (2010)	
G3	4	L	0.85	Virion association	Yang et al. (2010)	
G6	4	L	0.63	Host interaction	Yang et al. (2010)	
G7	4	L	0.55	Virion association	Yang et al. (2010)	
G9	4	L	0.53	Virion association	Yang et al. (2010)	
H2	4	L	0.59	Virion association	Yang et al. (2010)	
H4	4	L	0.49	Transcription	Yang et al. (2010)	3
H6	4	L	0.64	Transcription	Yang et al. (2010)	
I7	4	L	0.54	Virion association	Yang et al. (2010)	
J5	4	L	0.62	Virion association	Yang et al. (2010)	
L1	4	L	0.81	Virion association	Yang et al. (2010)	
L3	4	L	0.65	Transcription	Yang et al. (2010)	
L5	4	L	0.81	Virion association	Yang et al. (2010)	