From the Center for Infectious Medicine
Department of Medicine
Karolinska Institutet, Stockholm, Sweden

DEGRADOMICS - A STUDY OF THE CELLULAR PROTEOLYTIC LANDSCAPE IN ENTEROVIRUS INFECTIONS

Sebastian Kapell

Karolinska Institutet

Stockholm 2019

Cover illustration by Sabina Kapell.
Published by Karolinska Institutet.
Printed by E-Print AB 2019.

ISBN 978-91-7831-477-5

Degradomics - a study of the cellular proteolytic landscape in enterovirus infections

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Sebastian Kapell

Principal Supervisor:
Malin Flodström-Tullberg, Professor
Karolinska Institutet
Department of Medicine, Huddinge
Center for Infectious Medicine (CIM)

Co-supervisor(s):
Jonas Klingström, Associate Professor
Karolinska Institutet
Department of Medicine, Huddinge
Center for Infectious Medicine (CIM)

Lena Eliasson, Professor
Lund University
Department of Clinical Sciences
Islet Cell Exocytosis

Opponent:
Martin Denis Ryan, Professor
University of St Andrews
School of Biology

Examination Board:
Janne Lehtiö, Professor
Karolinska Institutet
Department of Oncology-Pathology

Joey Lau Börjesson
Uppsala University
Department of Medical Cell Biology

Nelson Gekara, Assistant Professor
Umeå University
Department of Molecular Biology
Molecular Infection Medicine Sweden (MIMS)

“Night science, hesitates, stumbles, recoils, sweats, wakes with a start. Doubting everything, it is forever trying to find itself, question itself, pull itself back together. Night science is a sort of workshop of the possible where what will become the building material of science is worked out. Where phenomena are still mere solitary events with no link between them. Where thought makes its way along meandering paths and twisting lanes, most often leading nowhere. What guides the mind, then, is not logic but instinct, intuition. The need to understand.”

François Jacob

ABSTRACT

The common human pathogen, enteroviruses (EVs) are a genus of single stranded RNA viruses that include Coxsackie B virus (CVBs). The viral RNA encodes a polyprotein containing two viral proteases $2A^{pro}$ and $3C^{pro}$, that process the polyprotein into structural and non-structural proteins during virus replication. Previous research has shown that aside from their role in polyprotein processing, the proteases also cleave host proteins in order to modulate different cell functions such as translation, transcription and innate immunity.

The innate immune response, and in particular the type 1 interferons (IFN), have an important role in controlling virus spread and protecting neighbouring cells from infection. Similarly, type III IFNs modulate the permissiveness of cells to CVBs. In **paper I**, we found that CVB3 has evolved a mechanism to evade the type III IFN response in a comparable manner to that previously shown for the type I IFNs. This perturbation is likely mediated via the proteolytic cleavage of the signal transduction proteins IPS-1 and TRIF by $2A^{pro}$.

When investigating virus-associated diseases, high quality reagents are essential especially when the aim is to detect virus in serum or tissue samples. We recognised a need for reagents capable of detecting CVBs and in **paper II** we described the development of CVB specific antibodies against $2A^{pro}/3C^{pro}$ and show their utility in western blotting, confocal microscopy, immunohistochemistry and flow cytometry.

Several observations have suggested a role for CVB infection in the development of type 1 diabetes (T1D). Interestingly, β-cells infected with CVBs in vitro have defective insulin secretion. The aim of **paper III** was to investigate whether β-cell dysfunction observed during CVB infection could be attributed to the activity of the viral proteases. We found that the viral proteins $2A^{pro}$, $3C^{pro}$ and 3A individually exert negative effects on glucose stimulated insulin release (GSIS) and voltage stimulated exocytosis in β-cells.

Based on the results in **paper I** and **III**, it is evident that $2A^{pro}$ and $3C^{pro}$ have multifaceted roles in the viral replication cycle and in the modulation of the host cell. In **paper IV** we wanted to define the viral proteases specific targets in Coxsackeivirus B3 (CVB3) permissive cell lines of varied origin (HeLa, CaCo-2 and EndoC-βH1). By enriching for the N-terminal peptides using subtiligase labelling combined with LS-MS/MS, we identified and validated 81 host cell protease substrates. Among the substrates we identified, the protein TCF7L2 was a target of $2A^{pro}$ mediated cleavage. TCF7L2 is an important transcription factor involved in maintaining β-cell functionality and glucose stimulated insulin secretion. This finding provides a potential mechanistic explanation for the observation that β-cells infected with CVBs *in vitro* are defective in insulin secretion.

The studies presented in this thesis increase our understanding the molecular virology of enteroviruses. Moreover, they open a new chapter of research in examining how disease pathology might be caused by the activity of virally encoded proteases.

LIST OF SCIENTIFIC PAPERS

I. Katharina Lind, Emma Svedin, Erna Domsgen, **Sebastian Kapell**, Olli H. Laitinen, Markus Moll and Malin Flodström-Tullberg. Coxsackievirus counters the host innate immune response by blocking type III interferon expression. The Journal of General Virology. 2016 vol. 97(6) pp. 1-12.

II. Olli Laitinen, Emma Svedin, **Sebastian Kapell**, Minna Hankaniemi, Pär Larsson, Erna Domsgen, Virginia M. Stone, Juha A.E. Määttä, Heikki Hyöty, Vesa P. Hytönen, Malin Flodström-Tullberg. New Coxsackievirus 2Apro and 3Cpro protease antibodies for virus detection and discovery of pathogenic mechanisms. The Journal of Virological Methods. 2018. 255. pp. 29-37.

III. Emma Svedin, Erna Domsgen, **Sebastian Kapell**, Omar Hmeadi, Anna Edlund, Sebastian Barg, Lena Eliasson, Malin Flodström-Tullberg. Coxsackievirus affects multiple steps in the insulin secretion pathway leading to impaired insulin release by infected beta cells. *Manuscript.*

IV. **Sebastian Kapell**, Mohsan Saeed, Milica Tesic Mark, Henrik Molina, Soile Tuomela, Juha A.E. Määttä, Vesa P. Hytönen, Charles Rice, Malin Flodström-Tullberg. Coxsackie B virus protinase substrate repertoire defined in human proteome by N-terminal proteomics uncovers cleavage of TCF7L2 by $2A^{pro}$. *Manuscript.*

CONTENTS

LIST OF ABBREVIATIONS

Abu	α-aminobutyric acid
Aβ	Amyloid β-peptide
AD	Alzheimer's disease
ALS	Amyotrophic lateral sclerosis
APP	Amyloid precursor protein
Bak1	BCL2 antagonist/killer 1
Bax	BCL2 associated x, apoptosis regulator
Bid	BH3 interacting domain death agonist
CAR	Coxsackievirus adenovirus receptor
CTL	Cytotoxic T lymphocytes
COFRADIC	Combined fractional diagonal chromatography
CVB	Coxsackievirus group B
DAF	Decay-accelerating factor
DOSE	Ontology Semantic and Enrichment
Dpi	days post infection
eIF4G	Eucaryotic intiation factor 4G
ECMV	Encephalomyocardits virus
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmic reticulum
EV	Enterovirus
EV-B	Enterovirus B
EV71	Enterovirus 71
FFPE	Formalin-fixed and paraffin embedded
FMDV	Foot-mouth disease virus
GZMA	Granzyme A

GZMB	Granzyme B
HCV	Hepatitis C virus
HFMD	Hand–foot and mouth disease
HIV	Human immunodeficiency virus
HOS	Higher order structure
HPV	Human papillomavirus
Hpi	Hours post infection
ICAM-1	Intracellular adhesion molecule 1
IIAE3	Infection-induced acute encephalopathy 3
IFN	Interferon
IHC	Immunohistochemistry
IRES	Internal ribosome entry site
IRF7	Interferon regulatory factor 7
ISG	Interferon stimulated gene
iTRAQ	Isobaric tags for relative and absolute quantitation
MAIT	Mucosal associated invariant T cell
MAVS	Mitochondrial antiviral signalling protein
MDA5	Melanoma differentiation-associated protein 5
MMP	Matrix metalloprotinase
MHC	Major histocompatibility complex
miRNA	MicroRNA
MOI	Multiplicity of infection
MyD88	Myeloid differentiation factor 88
NK	Natural killer
NF-κB	Nuclear factor kappa light chain enhancer of activated B cells
NOD	Non-obese diabetic

NPC	Nuclear pore complex
ORF	Open reading frame
PAMP	Pathogen associated molecular patterns
PCBP2	Poly(RC) Binding protein 2
PDB	Protein data bank
Poly I:C	Polyinosinic:polycytidylic acid
PRR	Pattern recognition receptor
PTB	Polypyrimidine tract binding protein 1
PTMs	Post translational modifications
RIG-I	Retinoic acid inducible gene I
Rip	Regulated intramembrane proteolysis
PVDF	Polyvinylidene difluoride
PVR	Polio virus receptor
RP-HPLC	Reverse phase liquid chromatography
Rnase L	Ribonuclease L
SCA2	Spinocerebellar ataxia type 2
SNP	Single nucleotide polymorphism
SRP20	Splicing Factor, Arginine/Serine-Rich 20-KD
T1D	Type 1 diabetes
TAILS	Terminal isotopic labelling
TCF	T cell receptor
TCF7L2	Transcription Factor 7 Like 2
TEV	Tobacco etch virus
TLE2	Transducin Like Enhancer Of Split 3
TLR	Toll like receptor
TNBS	2,4,6-trinitrobenzenesulfonic acid

TRIF	TIR-domain containing adapter-inducing IFNβ
uORF	Upstream open reading frame
QFWB	Quantifiable fluorescence based western blot
VPg	Viral protein genome linked
WT	Wild type

1 INTRODUCTION

The great naturalist and father of taxonomy, Carl Linnaeus once wrote in *Philosophia Botanica (1751), "natura non facit saltum*" which directly translates from Latin into "nature does not make jumps". This statement held onto the belief of the time that species were discrete and distinct from each other, as during his creation, God had already set the boundaries of the natural world for all that grow and crawl. However, it was through Linnaeus own study and systematic cataloguing of nature that he found continuity and fluidity of species across the different taxa. Perhaps this seeded the idea that new species could develop, since he later on came to cross out the passage *"natura non facit saltum*" in his own copy of *Philosophia Botanica*. This turned out to be correct a century later when Charles Darwin published *On the Origin of Species* (1859) and the theory of evolution was firmly established, stating that all species arise and develop from variation and the selective pressure by natural selection. Linnaeus act of crossing out that single passage in his book, underpins the most essential aspects of a scientific mind, elastic thinking and scepticism against all truths. Now at the start of the 21st century if Linnaeus himself could have access to our understanding of modern biology he would be astounded by the complexity but also the highly structured architecture of life. Never is this truer then when studying the biology of viruses, the most ubiquitous and abundant biological entity on the earth.

From their discovery in the early 19th century, starting with the identification of a few plant and animal pathogens, our understanding of viruses has been radically transformed. With the advent of DNA/RNA sequencing technology we now know it is very likely that the greatest amount of genetic diversity in nature resides within viral genomes (*1*). Furthermore, viruses are polyphyletic (having many evolutionary origins) and infect diverse and distant hosts across the tree of life (*2*). The continued arms race between higher order organisms and entities such as viruses over the millennia has been pivotal in shaping a diverse and effective immune system (*3*, *4*). As an obligate intracellular parasite, viruses interact within the cellular machinery subverting and perturbing some functions whilst stimulating others, thereby tightly integrating themselves into the inner working of the cell. Therefore, it is through the careful study of the relationships between pathogen and host that we can decipher new and exciting biology. Some of the most transformative discoveries in immunology and medicine can be directly traced back to these fascinating relationships.

The aim of this thesis is to add to our understanding of enterovirus biology especially in relation to the virus encoded proteases, defining the role of protease activity in virus replication and human disease. Hopefully, part of this thesis work also contributes to the new and exciting field of degradomics, the study of proteases and their substrate repertoire and ultimately how substrate cleavage shapes cellular function. Four articles form the basis of this thesis, **paper I** describes how CVBs has evolved a mechanism to block the type III interferon response upon infection. In **paper II** the development of two antibodies are described that are reactive against the CVB proteases 2A^{pro} and 3C^{pro}. The utility of these antibodies for detecting enterovirus B (EV-B) species in methodologies such as

immunohistochemistry (IHC), western blot (WB), flow cytometry and confocal microscopy is demonstrated. In **paper III** the effect of CVB infection and overexpression of viral proteases on β-cell function is investigated. The final study, **paper IV,** is an in-depth characterisation of the $2A^{pro}$ and $3C^{pro}$ substrate repertoire in multiple cell lines.

1.1 PROTEASES AND THEIR ROLE IN CELLULAR FUNCTIONS

1.1.1 CLASSIFICATION AND BIOCHEMISTRY

The most fundamental process in regulating the function of a cell is that which is carried out by the enzymes that break peptide bonds known as proteases (*5*). Proteolysis is the most universal of post translational modifications (PTM) of cellular proteins. There are exceptionally few proteins resistant to proteolysis and in essence every polypeptide will at some point be broken down to its most basic components. Amino acids can then be re-cycled and re-purposed for translation.

Proteolysis integrates seamlessly into the most crucial processes within a cell such as removing signalling peptides and PTM´s from synthesized proteins, allowing for their appropriate compartmentalization and function. Proteolytic processing may even target a protein into the cell´s secretory pathways or release proteins anchored in the plasma membrane, a process known as regulated intramembrane proteolysis (Rip; (*6*)), which can thereby promote cell to cell communication. Many hormones and enzymes are intimately regulated by proteolysis as they require the removal of pro-domains to allow them to assume their active conformation. It is therefore not surprising that any alteration in the expression or activity of a protease may contribute to pathology in many diseases. This makes the pursuit of deciphering protease biochemistry and biology a prerequisite for understanding general cell physiology as well as tissue pathology in various diseases.

There are over 550 proteolytic enzymes annotated for *Homo Sapiens*, almost all of which are hydrolases, meaning they are biochemical catalysts that use a water molecule to break a chemical bond. A typical structure for such an enzyme usually consists of an indentation where the active site is located (*7*). An active site consists of residues either in a triad or dyad conformation that participate in a catalytic reaction that breaks a peptide bond. Each binding pocket in the protease is numbered S1, S2, S3, S4 (and so on) with S1 being located just next to the scissile bond (a covalent chemical bond that can broken). The same numbering scheme is used for the substrate binding sites; P1, P2, P3, P4 (*8*). Interactions of substrates with positions outside of the active site important for proteolytic cleavage are termed exosites (*9*).

As it is with most classification schemes there are often multiple ways of categorising any biological entity. For proteases one of the most fundamental divisions that can be made is by specificity. For instance, if the proteolytic enzyme targets the N- or C- terminus it would be defined as an “exopeptidase”, examples of which include (aminopeptidases and carboxypeptidases). However, if a peptide bond is hydrolysed inside a peptide chain then the protease would instead be labelled as an “endopeptidase” and such proteases include pepsin and trypsin. Furthermore, classification can be performed based on the catalytic type, meaning the type of residues that occupy the catalytic site; for instance, serine, threonine or cysteine types. If occupied by aspartate or glutamate, or if there is a metal ion bound to the amino acid side chain, the protease would be named either as aspartyl, glutamyl or matrix metalloprotinase (MMP) (*10*).

Compared to other groupings of various enzymes and their families, neither of these classifications take note of the evolutionary origin of the different proteases. The most generally accepted systematic organisation of peptidases which has been introduced is based on sequence homology, and the system that was created based on this organised principle is termed MEROPS (*11*, *12*). The system relies on dividing the known peptidases into clans and families.

1.1.2 REGULATION OF PROTEASE ACTIVITY

Since proteases are involved in many processes within a cell and with proteolysis essentially being an irreversible reaction it is important that it is tightly regulated. It is interesting to note however that there are some enzymes known with ligation activity that is either naturally occurring (*13*) or that that has been engineered (*14*, *15*), although they seem to occur much less frequently in nature compared to proteases. Regulation of protease activity involves classical mechanisms such as producing proteases as zymogens, which require proteolytic processing before they can become active. Proteases can also be spatially separated from their substrates through localisation to different cellular compartments. Moreover, proteins may act as protease inhibitors by binding to the active site of the protease, or the enzyme may be in an inactive state until it is exposed to its required co-factor (various ions). For some proteases, additional regions outside the active site have been identified where ligand binding can function as an allosteric inhibitor or activator. These allosteric binding sites can provide more selective modulation and the protease may even be transiently activated (*16*).

Several classical proteolytic cascades have been described that demonstrate the activation of proteases. Most notably, these include the proteases connected to cell death, the cysteine proteases (caspases) (*17*) and pathways involved in blood coagulation which require serine protease activity (*18*). Another well described proteolytic cascade related to the innate immune system is the complement system or the complement cascade, in which, upon target recognition, a series of serine proteases are activated (*19*).

Due to advances in the annotation of substrates that are cleaved by proteases that belong to various families, a new view of protease regulation has started to form termed the “protease web” (*20*). Proteases process other proteases, but also cleave inhibitors that act on a diverse number of other proteases. This creates complex regulatory circuits with feedback and amplification of signals via proteolytic cascades. Therefore, proteolysis within a cell is highly interconnected across groups of proteases that even belong to different families, making the regulation of any individual protease a highly complex matter.

1.1.3 CELL DEATH AND PROTEASE ACTIVATION

Initially proteases were thought of as a method for the cell to degrade and recycle proteins but as the research field expanded it became clear this was only one role of the proteases in biological systems. Now proteases are considered to be equally essential component of many cell signalling pathways. This is illustrated very well when considering the process of

cell death. It has long been known that caspases have an important role in both the initiation and execution of apoptosis (*21*). As the process of cell death was explored it became clear that even proteases outside of the caspase family were involved, such as granzymes (*22*), cathepsins (*23*) and calpains (*24*).

Of all the eleven caspases encoded in the human genome, roles for seven of them in the process of apoptosis have been described so far (2-3, 7-10) and these caspases can be further sub-divided into either initiator caspases (2, 8, 9, 10), or executioner caspases (3, 6 and 7) (*25*). The initiator caspases are activated either by the intrinsic (internally triggered and mediated via the mitochondrion) or extrinsic pathways (external triggered by ligand binding to surface receptor) (*26*). External signalling by binding of tumour necrosis factor (TNF) to the death receptor leads to downstream activation of caspases 8 and 10. In turn these proteases mediate the downstream cleavage of effector caspases 3, 6, 7 and caspase specific substrate cleavage resulting in apoptosis of the cell. The intrinsic pathway can be triggered by different types of cell stress, causing the oligemerisation of BCL2 Antagonist/Killer 1 (Bak1) and BCL2 Associated X, Apoptosis Regulator (Bax) forming pores in the mitochondrial membrane. Released cytochrome c from the mitochondrion in turn binds to the apoptotic protease-activating factor 1 (Apaf-1) forming the apoptosome thereby activating caspase 3 (*27*), thereby triggering apoptosis.

Cell death can also be induced in a target cell by the release of cytotoxic granules containing serine proteases named granzymes, from cytotoxic T lymphocytes (CTL) (*28*), natural killer (NK) (*29*) cells or Mucosal associated invariant T (MAIT) cells (*30*, *31*) thru the process of CTL-mediated cytotoxicity (*32*, *33*). So far, five different granzymes have been identified in humans (A, B, H, K, M) (*34*, *35*) for which only a few the mechanism of their cytotoxic activity have been understood. Granzyme A (GZMA), can induce caspase-independent cell death partly by cleaving mitochondria associated protein (NDUFS3) and disrupting mitochondrial metabolism (*36*, *37*). Granzyme B (GZMB) is one of the best characterised granzymes and it was first thought to trigger cell death by directly activating caspase 3 by cleavage (*38*). However, follow up studies concluded that it was mainly granzyme mediated cleavage of BH3 Interacting Domain Death Agonist (Bid) that triggered apoptosis by the intrinsic/mitochondrial pathway (*39*, *40*).

The lysosomal cysteine proteases, cathepsins, are known for their roles in degrading cargo in the endocytic pathway and for their involvement in normal protein turnover. Due to cellular stress or other triggers such lysothropic agents or viral proteins, cathepsins can translocate into the cytosol (*41*). Interestingly, some cathepsins even retain their proteolytic activity at neutral pH (cytosolic conditions) with cathepsins B, L, K and S remaining stable outside the acidic lysosomal compartment (*42*, *43*). It has also been demonstrated that the proteases B, K, L, and S, along with cathepsin H could possibly cleave Bid (*44*) providing a rationale for how the cathepsins contribute to the induction of apoptosis.

Similarly, the non-lysosomal cysteine proteases that belong to the calpain family can cleave the pro-apoptotic protein Bid as well as survival proteins such as Bcl-2 and Bcl-X_L (*24*). Direct cleavage of caspase 7 by calpain has been observed, possibly inactivating the

caspase (*45*) providing another example of intersection between calpain activity and apoptosis.

These different families and types of proteases are all intricately integrated in orchestrating specific outcomes, mainly the initiation and execution of the cell death pathways. It is easy to appreciate the depth of regulation for which proteolytic processes directly give rise to by alterations seen after extreme or subtle changes in the protease web.

1.1.4 PROTEASES IN HEALTH AND DISEASE

The type of diseases caused by the deregulation of protease activity are as varied as the processes that proteolysis regulates. This observation is most striking when considering the hereditary diseases caused by mutations in protease genes (*46*). Alterations in expression patterns or changes in protease structure are the main contributors to the many pathological processes including; cardiovascular disease, neurodegenerative disorders, osteoporosis and cancer. To provide a comprehensive survey of these pathologies and the mechanistic explanations for how proteolysis is involved is well beyond the scope of this thesis. However, some representative examples are provided in the following paragraphs, for an excellent review on the topic see the following reference (*47*).

A broad array of proteases ranging from cysteine and serine proteases to metalloproteases have been implicated in cardiovascular disease (*48*). They have been shown to contribute to various pathologies through a diverse range of processes such as matrix turnover, cellular migration and formation of new blood vessels (angiogenesis).

The causes of the neurodegenerative disorder Alzheimer's disease (AD) is generally unknown. Although that formation of cerebral plaques composed of amyloid β-peptide (Aβ) is thought of as a pathological hallmark of the disease (*49*). Since these plaques are inherently derived from the processing of the amyloid precursor protein (APP) there is a direct dependency on the proper function of three proteases; α-, β-, and γ-secretases (*50*, *51*). Therefore, the proteases not only present a direct link to the pathology but also provide a target for implementing therapeutic strategies to limit Aβ plaque formation (*49*). So far, such therapeutic strategies have sadly proven ineffective, Verubecestat a β-secretase inhibitor did not reduce the cognitive or functional decline in mild-to-moderate patients suffering from AD (*52*).

Osteoporosis is characterized by the loss of bone density and mass due to degeneration by osteoclasts as well as inadequate regeneration of the bone matrix by osteoblasts (*53*). Approximately 90% of the bone mass consists of type I collagen fibres, making bone vulnerable to the effect of potent collagenases such as the cysteine proteases including cathepsins K (*54*). The proteases expressed by osteoclasts are thought to play a key role in the degradation of bone matrix. This has led to the development of compounds targeting cathepsin K for use in osteoporosis treatment and such selective inhibitors have entered clinical trials (*55*). As an example, Odanacatib, a caphepsin K inhibitor was recently evaluated in a phase III clinical trial in postmenopausal women. The study concluded that there were a clinically relevant reduction in bone fractures (*56*). Although the compound

was later discontinued as adverse drug effects had been noted. Regardless the result of the study increased the interest for cathepsin K as a potential therapeutic target for osteoporosis (*57*).

Research on the very broad area of cancer has, during the last few decades, demonstrated roles for proteases in the early events of tumour growth and the late stages in tumour progression, invasion and metastasis (*58*). The pericellular proteases (such as MMPs) have, in particular, been recognized for their importance, since the induced expression of MMP2 and MMP9 contribute to pericellular proteolysis and remodelling of the tumour microenvironment (*59*).

When studying examples of protease mediated or associated diseases, the realisation comes quickly that combined with previously mentioned protease web, cell pathophysiology is never isolated but instead interconnected across a multidirectional proteolytic network spanning several families of proteases. This underlies the point that only by viewing these events as a system of activating and inhibiting reactions it becomes possible to describe the outcomes of proteolytic deregulation in any given cell. Moreover, this outcome will also be dictated by both the expression as well as the activation state of each individual protease.

1.1.5 METHODOLOGIES FOR PROTEASE SUBSTRATE DISCOVERY

The early efforts in identifying protein substrates relied predominantly on defining the sequence specificities at cleavage sites for individual proteases. In order to extrapolate these specificities into novel peptides and experimentally verify a substrate *in vitro* the methodologies applied for this purpose were varied, ranging from expressing peptide substrates by phage display (*60*) to using synthetic peptide libraries (*61*). There is however an inherent limitation with these approaches (aside from being laborious), relating to the fact that only one aspect that defines a true cleavage site is tested; namely the primary amino acid sequence. The tertiary structure that may either block or facilitate protease-substrate interactions is not addressed. It is possible that proteolytic cleavage occurring might be entirely dependent on tertiary structure of the protein.

These aforementioned limitations prompted the move into gel-based methodologies, such as 2D gel-electrophoresis (*62*). Separating complex protein mixtures, allows for comparative analysis between control and treated samples (those that have undergone proteolysis). After proteome separation on a gel, differentially stained spots can be excised and analysed by mass spectrometry. Limitations with this approach and variations of it, lie in the resolution and sensitivity required to identify substrates, since adequate material is needed for the staining and isolation steps. Deriving the cleavage site and thereby the protease specificity is also challenging by such methodologies.

Another method for selecting and enriching for proteolytically processed peptides takes advantage of the N-termini generated by proteolysis (*63*). These termini are unique in the sense that they are rich in reactive nitrogen atoms at the N-terminal α-amine. It is through this property that these free N-termini can undergo reactions and subsequent conjugation with various molecules. Several methodologies utilising such conjugation strategies have

been developed, with the aim to enrich for peptides that can be subsequently analysed by mass spectrometry.

1.1.6 NEGATIVE SELECTION STRATEGIES

Negative selection of peptides by combined fractional diagonal chromatography (COFRADIC) (*64*) uses the compound 2,4,6-trinitrobenzenesulfonic acid (TNBS) to label free N-termini, increasing their hydrophobicity and allowing for their separation by reverse phase liquid chromatography (RP-HPLC). Alternative approaches applying negative selection have also been developed, an example of this type of methodology being terminal isotopic labelling (TAILS) (*65*, *66*). Free N-termini are blocked by dimethylation or isobaric tags for relative and absolute quantitation (iTRAQ) labelling, and selection comes after the peptides have been blocked and digested with trypsin, in order to be reacted with an aldehyde-derivatised polymer. Unbound peptides can then be filtered out and analysed by LC-MS/MS and the identification of peptides can thus occur. A caveat of this procedure is that the incomplete capture of peptides (thereby incomplete negative selection) could potentially cause false positives.

1.1.7 POSITIVE SELECTION STRATEGIES

Another methodology that has been developed that instead relies on positive selection of N-terminal target peptides is termed subtiligase labelling (*67*, *68*) (Figure 1). At the heart of this method lies a bacterial protease (subtilisin) that has been genetically engineered by the introduction of two point mutations (*15*, *69*). These mutations have changed the catalytic serine to a cysteine and modified the structure of the catalytic site by exchanging a proline in position 225 to alanine. These two modifications abrogate the proteolytic activity of the protease and instead provide the enzyme with ligation activity (subtiligase). By utilising this enzyme it is possible to specifically incorporate a biotinylated-ester, containing a non-naturally occurring amino acid (α-aminobutyric acid or Abu)) at the site of the free amine termini generated by the proteolytic cleavage of the protein (*70*). Positive selection is then achieved by capturing the labelled peptides on streptavidin beads with a biotinylated-tag designed with a tobacco etch virus (TEV) protease cleavage site. The tag can, via TEV protease mediated cleavage, liberate the peptide from the bead. After this cleavage, a dipeptide that has the Abu modification, remains attached to the N-terminus which during LC-MS/MS analysis can be used for verifying true positives. This negates one of the problems with negative selection as the risk for false positives is greatly reduced.

Since the subtiligase labelling is rate limited and the efficacy of the enzyme low, this increases the amount of sample protein required for the labelling reaction. This can be a bottleneck for extensive and deep coverage of peptides identified if the sample preparation is limited. Furthermore, it is known that the subtiligase enzyme has lower activity when labelling acidic amino acid (asparagine and glutamine) or branched-chained residues like proline, valine and isoleucine (*15*). Instead the enzyme has enhanced activity against small amino acids like glycine, alanine and serine. Recently efforts have been made to counter biased labelling specificity of the enzyme by screening point mutants of subtiligase, thereby

finding mutants with less biased labelling, these efforts have been successful and improved the existing technology (*71*).

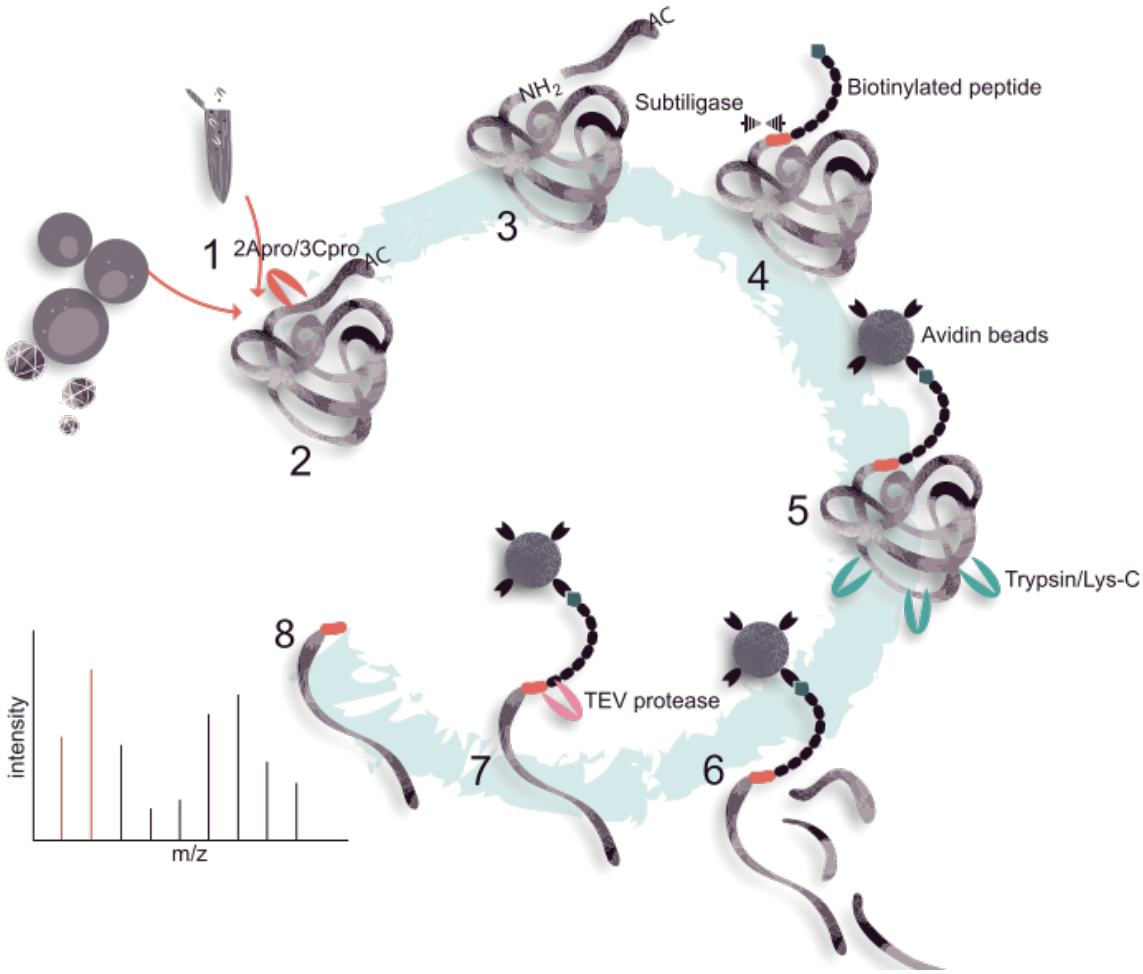

Figure 1. Subtiligase labelling of protein lysates enrich for the free N-terminome of proteolytically cleaved proteins*. (1) Protein lysates prepared for labelling reaction. (2) Intramolecular cleavage of protein by active proteases. (3) Cleavage generates free N-termini within the protein. (4) Ligation of N-termini with biotinylated peptide ester tag (containing TEV cleavage site), reaction catalysed by subtiligase. (5) Binding of tagged peptide to agarose streptavidin beads and fragmentation of protein by trypsin/Lys-C treatment. (6) Removal of unbound peptides generated by intramolecular cleavage by trypsin/ Lys-C. (7) TEV protease cleavage removs peptide ester tag from streptavidin bead complex leaving the Abu modification in extreme N-terminus of the captured peptide. (8) Analysis by LC-MS/MS.*

Undoubtedly the pace at which protease specific substrates and cleavage sites are identified for many proteases, has significantly advanced during the last decade. These advancements have happened synergistically together with new innovative enrichment methodologies as previously described, but also due technical developments in proteomics and mass spectrometry. With the generation of large dataset, initiatives like that of the MEROPS (*12*), DegraBase (*72*) and TopFind (*73*) databases, can collect highly annotated

degradomes generated under various experimental conditions. Integration of data together with rational model building based on our understanding of protease biology hold the promise one day describing the complete proteolytic landscape within a cell. It is only by studying the whole proteolytic landscape true insight into disease states caused by deregulated proteolysis will be gained. Finally, such insights will provide therapeutic opportunities for negating the harmful changes occurring within the protease web during different pathologies.

1.2 THE MOLECULAR VIROLOGY OF ENTEROVIRUSES

1.2.1 TROPISM

The members of the *Picornaviridae* family in the genus of enteroviruses are small single stranded positive sense RNA viruses. These enteroviruses gain entry into cells via surface receptors, and many of these transmembrane proteins have been described poliovirus receptor (PVR), Coxsackievirus and adenovirus receptor (CAR), complement decay accelerating factor (DAF) and intracellular adhesion molecule 1 (ICAM-1) (*74*). For a virus, the surface receptor is the first requirement in establishing a successful infection. Therefore, unsurprisingly, virus entry via receptor binding is an important determining factor in dictating host and tissue tropism. In the case of poliovirus, PVR that is expressed by lymphoid tissue of the gastrointestinal tract enables the virus to gain access to the cells thereby establishing the primary site of infection (*75*, *76*). If later, the poliovirus infection becomes systemic then the neuronal cells (expressing PVR) can also be infected. However, it is too simplistic to regard the permissiveness of a cell to virus infection as solely due to the expression of a surface receptor, as other intracellular host proteins are needed for the successful completion for viral replication (*77*, *78*). Factors that regulate the ability of a cell to sustain viral propagation may be modulated further by external immunoregulatory molecules like the interferons (*79*). Surveying the literature, it is clear that enteroviruses in both *in vitro* cell culture as well as *in vivo* animal models, can infect and replicate in a broad range of cell types and hosts (*80*).

1.2.2 REPLICATION AND PROTEOLYTIC PROCESSING OF THE VIRAL POLYPROTEIN

Upon entry into a cell, the virus capsid undergoes uncoating which may be conditional on acidification of the endocytic compartment, depending on enterovirus species (*81*, *82*). Conformational changes in the capsid facilitate the interaction of the structural proteins VP1 and VP4 with the endocytic membrane (*83–85*). This allows for the release of the positive sense viral RNA (approximately 7.4 kb in size) into the cytosol. The viral RNA is then used directly to produce the viral polyprotein by cap-independent translation mediated by RNA secondary structures termed internal ribosome entry site (IRES) (*86*). For the last 50 years it was thought that enteroviruses encoded one open reading frame (ORF) that makes the polyprotein, but recently this long-established dogma was disproven after the identification of an additional ORF encoding a 56-76 amino acid long (predicted) transmembrane protein, present in some enterovirus species (*87*). This factor, designated as upstream open reading frame (uORF) protein, is thought to be important in facilitating efficient virus release from an infected cell.

The enterovirus polyprotein is divided into different regions (Figure 2) that consist of the structural capsid proteins in the P1 region (VP4-VP1), and the non-structural proteins in the P2 region (2A-2C), together with the P3 region (3A-3D) (*88*). Processing of the viral polyprotein by the viral encoded proteases $2A^{pro}$ and $3C^{pro}$ generates the individual functional proteins. An interesting exception from this mode of polyprotein processing can

be seen with foot-and-mouth disease virus (FMDV) (*89*, *90*). In the polyprotein of FMDV, P1 and P2 separation is not mediated by $2A^{pro}$ proteolytic cleavage, but instead it is caused by a series of amino acid residues located within the $2A^{pro}$ sequence that prevent the linking of the P1 and P2 segments during translation. This process is termed "ribosomal skipping" and has been observed in multiple positive and negative RNA stranded viruses (*91*).

Collectively, the structural and non-structural viral proteins translate the viral RNA, modulate the host cell to facilitate the replication (*92*), subvert the cell innate immune response (*93*), and finally initiate encapsidation and formation of new infectious virions (*94*). Release of newly formed virions was traditionally thought to mainly occur by cell lysis, however this classical view is being challenged, as observations have been made suggesting that viruses can be shed in extracellular vesicles (EVs) (*95*, *96*), offering an alternative method of viral dissemination. Further support of the latter mode of transmission comes from reports that enteroviruses can establish persistent infections in both tissue and cell models with little cytopathic effects (*97–99*).

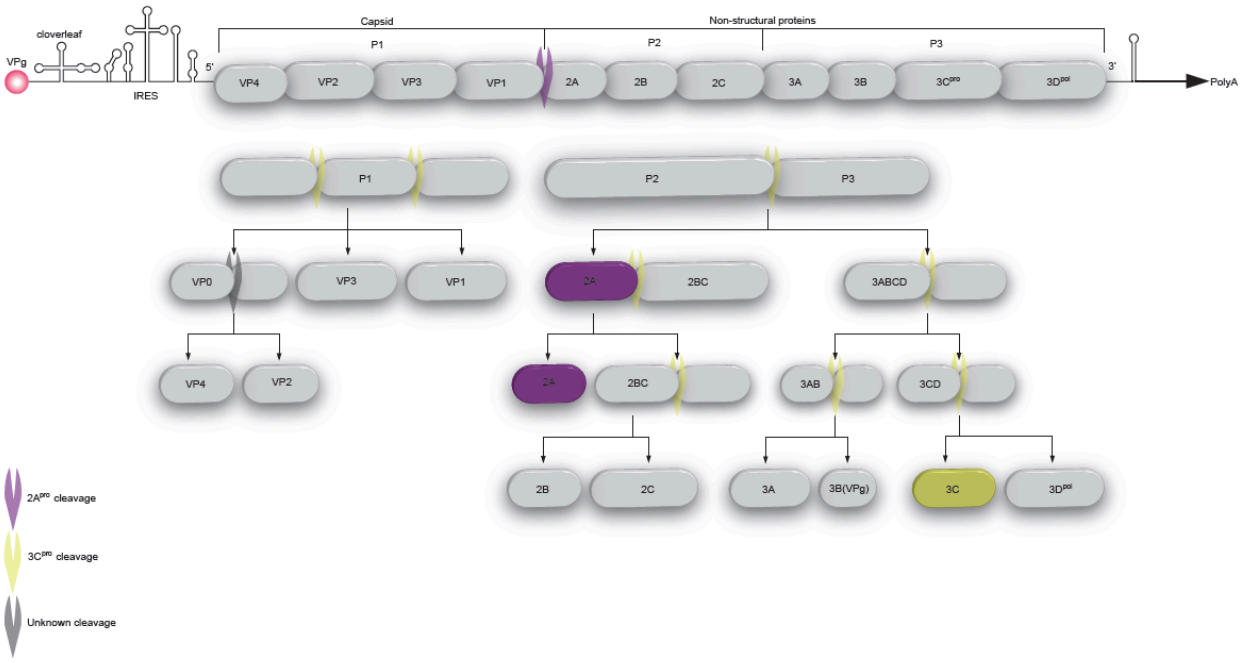

Figure 2. Schematic depicting the organisation of enterovirus RNA and polyprotein, highlighting sites of proteolytic processing attributed to $2A^{pro}$ and $3C^{pro}$. *RNA secondary structures (IRES) indicated outside the polyprotein segments, with the non-structural protein, viral protein genome linked (VPg) coloured red. Cuts in the polyprotein is highlighted by the colours purple ($2A^{pro}$) yellow ($3C^{pro}$) with black colour indicating autocatalytic cleavage.*

1.2.3 THE VIRAL ENCODED PROTEASES $2A^{PRO}$ AND $3C^{PRO}$

For most enteroviruses, the processing of the polyprotein is divided between the two viral encoded proteases $2A^{pro}$ and $3C^{pro}$. The scissile bond between VP1 and 2A is cleaved by $2A^{pro}$, whereas all other sites separating the different viral proteins are cleaved by $3C^{pro}$ (Figure 2) (*93*). Processing of VP0 into VP4 and VP2 is an exception as neither of the two

proteases have been shown to cleave VP0, instead structural modelling has inferred that this event occurs autocatalytically (*100*). The discovery and mechanism of this process is discussed in detail in chapter 1.2.4.

Structurally and biochemically, the enterovirus proteases $2A^{pro}$ and $3C^{pro}$ both belong to the chymotrypsin-related endopeptidases. Each protease is comprised of two separate domains (Figure 3). The structural elements themselves are made up of antiparallel β-sheet barrels, together with a β-sheet pile for $2A^{pro}$ and a conserved ion binding domain in both proteases (zinc ion for $2A^{pro}$ and chloride ion for $3C^{pro}$). The structure of both $2A^{pro}$ (*101*) and $3C^{pro}$ (*102*) proteases have been solved multiple times for different species of enteroviruses and can be found in the protein data bank (PDB) (*103*). Conservation of $2A^{pro}$ and $3C^{pro}$ sequences between different species of the same genus ranges between 50-75% (*93*). Moreover, the region in the vicinity of the catalytic site is partially conserved for both $2A^{pro}$ and $3C^{pro}$ within the enterovirus genus. This suggests that sequence specificity, and in addition substrate specificity and repertoire, may be similar irrespective of the enterovirus species. Indeed, this has been documented within the scientific literature as multiple substrates are shared between enteroviruses. Examples of such conserved substrates are discussed in chapter 1.3.2 (*93*).

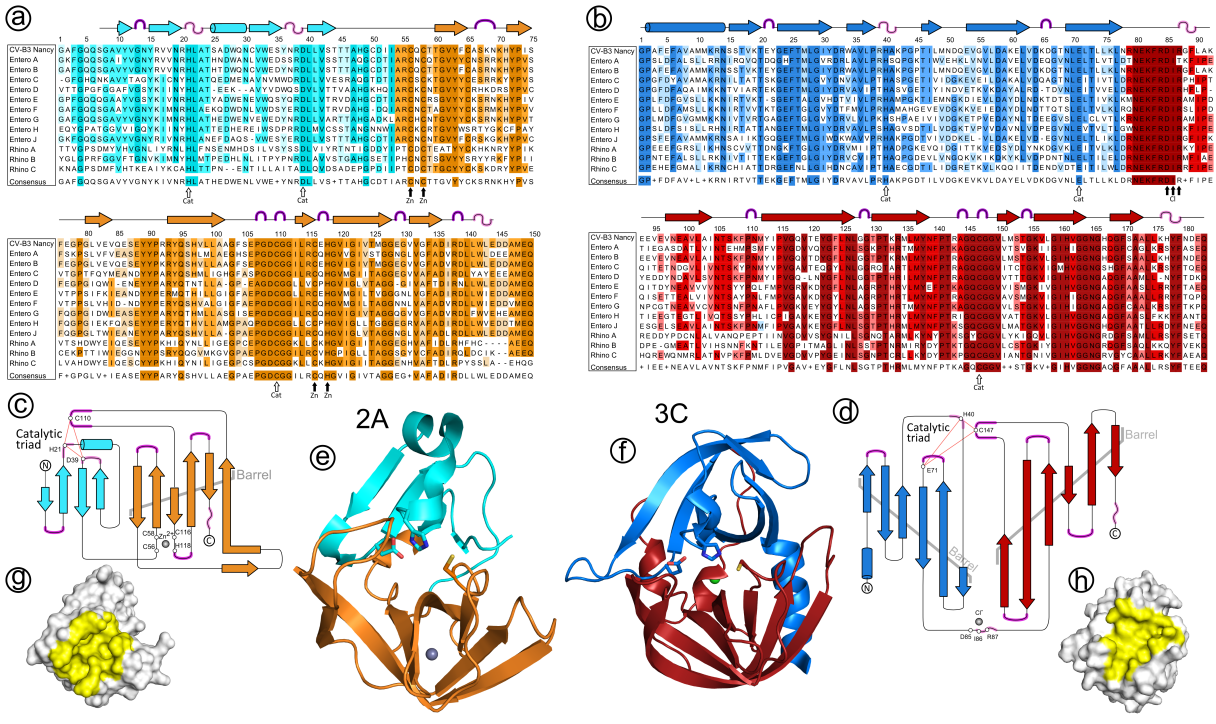

Figure 3. Sequence alignment and graphical representations of $2A^{pro}$ and $3C^{pro}$ structure. *(a) and (b) Primary sequence alignment of a representative sequence from each virus species in the enterovirus genus. The structural features of the catalytic triad and ion binding amino acids are highlighted with arrows. (c) and (d) topological schematic over $2A^{pro}$ and $3C^{pro}$. (e) Ribbon model of Enterovirus 71 (EV71) $2A^{pro}$ (PDBID: 4FVB). (f) Ribbon model of CVB3 $3C^{pro}$ (PDBID:2VB0). (g) and (h) Surface representations of each protease with the active site of marked in yellow. Figure adapted from* (*93*) *with permission.*

Efforts have been made to develop tools able to predict novel substrates based on the sequence cleavage specificity of the $2A^{pro}$ and $3C^{pro}$ proteases, the most notably tool being the NetPicoRNA server (*104*). Relying on the available polyprotein sequences of annotated enteroviruses, the authors applied a neuronal network algorithm to define amino acid logos as well as the surface accessibility of the cleavage site within the polyprotein. This work highlighted the importance of the conserved amino acid identities at positions P_4, P_2, P_1, $P_{1'}$, $P_{2'}$, especially for $2A^{pro}$ where the P_1 amino acid was exclusively a glycine whereas the $3C^{pro}$ logo was characterised by a glutamate in position P_1 and glycine at $P_{1'}$. A tool like the NetPicoRNA server has allowed researchers to investigate the susceptibility of cellular proteins of interests to cleavage by the enterovirus proteases $2A^{pro}$ and $3C^{pro}$. Since the tool was published in 1996, the server has now been discontinued. But with the advancements in computational analysis in the area of machine learning and with the availability of more extensive datasets of protease substrates, it is very likely that more accurate and refined prediction algorithms will be implemented in the near future. In fact, such tools are already available (iPROT-Sub, PROSPERous, PeptideCutter) for a range of endogenous protease substrates (*105–107*).

Determination of the protease specificity based on cut sites in the polyprotein provides logos for the amino acids which are accepted (or preferred) in the cut site. Ultimately however this approach could be somewhat misleading in predicting new cellular substrates, since the conservation of amino acids in the cut sites can be due to evolutionary constraints rather than the connection with protease cleavage. The type of logos that rely exclusively on viral polyprotein sequences may therefore be limited, and this holds further truth as the higher order structure (HOS) of the substrate is not incorporated into this type of predication model. Hopefully, as more protease substrates are structurally determined it will become possible to identify patterns in tertiary structures and as such improve the existing *in silico* prediction tools.

1.2.4 THE AUTOCATALYTIC CLEAVAGE OF VP0

The definition of autocatalysis is a reaction in which the product is the catalyst for the same reaction. In a classical study (*108*) published in 1968 the mechanism by which poliovirus is processed started to be elucidated. It was evident from pulse chase experiments performed within a short time frame (30min) viral polyprotein is processed by some mechanism. In the same study the capsid protein VP0 was identified, but it took two more years before the same authors could show that VP0 gave rise to the smaller capsid proteins VP2 and VP4 (*109*). For many years not much work was done to investigate the process from which VP2 and VP4 is generated from VP0. It was clear however from cell-free experiments in extracts from rabbit reticulocytes that processing did not take place solely by the expression of full-length polyprotein (*110*).

Eventually the hypothesis was proposed that cleavage of VP0 would only occur during the final assembly of virions containing RNA (*111*). Late in the year of 1981 another study could by partial amino-terminal sequencing determine the amino terminus of each capsid protein (*112*). It was concluded that cleavage of VP0 was distinct from other cleavages in

the polyprotein, again pointing to its direct connection with viral morphogenesis. The same authors succeeded in producing infectious poliovirus particles by cell-free translation of poliovirus RNA in HeLa cell extract. In these experiments accumulation of VP0 and faint amounts of VP2 could be detected by SDS-PAGE (*112*). This was an advance in the field, since other research groups had only succeeded in assembling immature capsid intermediates (containing uncleaved VP0) (*113*, *114*).

As the provirus structure VP0 is uncleaved, these capsids were not as stable as mature virions at elevated temperature, detergents and high pH (*113*). These biochemical properties made them difficult to isolate and study. This issue was circumvented in 1994 with the development of a specialized protocol (*100*), allowing the isolation and crystallization of the empty capsid intermediates at a 2.9 Å resolution. A structure was modelled in which VP0 was intact and based on this structure the hypothesis that the cleavage was autocatalytic in nature proposed.

The hypothesis was based on four key observations. First, close to the VP0 scissile bond there is a histidine side chain that could potentially catalyse the hydrolysis of the bond. Second, solvents were detected between the bond and the histidine (potentially water). Third, there are two carbonyl oxygen atoms oriented in the same direction that upon ligand binding could increase the positive charge in the scissile bond and accelerate hydrolysis. Fourth, within the depression beneath where the scissile bond resides there are hydrophobic residues, strengthening the possibility that ligands could bind to the site, such as RNA. This would be in line with the previous postulations that cleavage is RNA dependent. Based on the observations made from this structural modelling the hypothesis that VP0 cleavage is autocatalytic in nature was firmly established.

1.3 ENTEROVIRUSES AND THE HOST CELL

1.3.1 VIRAL ASSOCIATION WITH TYPE 1 DIABETES DEVELOPMENT

Of all the infectious diseases known to affect humans, some of the most common can be attributed to viruses belonging to the *Picornaviridae* family, especially those within the enterovirus genus. The common cold, poliomyelitis, myocarditis, meningitis, and hand–foot and mouth disease (HFMD) are to name but a few diseases caused by enteroviruses (*115*). In addition chronic and inflammatory diseases related to autoimmunity, type 1 diabetes (T1D), allergies and exacerbations in asthma have also been associated with enterovirus infections (*116*).

T1D is a chronic disease caused by the loss of insulin secreting β-cells in the pancreas resulting in an inability to correctly regulate blood glucose levels. The association between enteroviruses and especially Coxsackievirus B (CVB) and T1D is longstanding as the first report dates back to the late 1960s. In the study it was shown that, the incidence of CVB4 specific antibodies significantly correlated with the seasonal disease presence and onset of juvenile diabetes (*117*). In this study, one should note that of the six serotypes within the CVB group, five were tested and the serology of CVB1, CVB2, CVB3 and CVB5 did not show significant correlations. Further epidemiological and clinical studies trying to substantiate this association have been conducted with varied outcomes (*118*).

In 2004 a systematic review was conducted on the CVB serology in connection to T1D development examining 26 case-control studies published between the years 1966-2002 (*119*). The authors concluded that the evidence was too inconsistent to convincingly prove an association. About 7 years later another meta-analysis surveying observational molecular studies (24 studies and 2 abstracts from 1965 to 2010) measured two outcomes, enterovirus association with either T1D or autoimmunity (*120*). Despite variation in study design and statistical heterogeneity across/between the studies, significant associations between enterovirus infection and T1D, as well as autoimmunity, could be substantiated.

In a more recent example, the Diabetes Virus Detection (DiViD) study (*121*) in which a small pancreatic biopsy was collected from adults with recent onset diabetes, enterovirus RNA was detected by RT-PCR in four out of the six donors (*122*). Viral capsid protein (VP1) positivity in pancreatic islets was also detected in all six donors by immunohistochemical (IHC) staining. However concerns have been raised about the reliability of these findings as positivity for viral RNA (by RT-PCR) was only found in one out of six donors in the biopsy sample (*123*). The former report documenting RNA positivity in four out of the six cases came from the analysis of media in which the isolated islets had been cultured in prior to use in functional assays (105), whereas the results in the latter report were obtained directly from the biopsies (106). Due to the inconsistency in the results despite same detection method for viral RNA, laboratory contamination cannot be ruled out.

Many of the observational molecular studies conducted so far have relied upon the use of a particular VP1 antibody (clone 5D8/1, Dako). It has been reported that this specific antibody cross-reacts with endogenous mitochondrial antigens (*124*). Non-specific staining for the 5-D8/1 clone have also been reported to occur in IHC when using heart tissue derived from both mouse and human (*125*). This raised concerns about the validity of studies utilising this reagent, although it has been argued that under optimised experimental conditions it is likely that detection is reflective of presence of enterovirus antigens in infected tissue (*126*). VP1 positive islets have been found in tissue derived from patients with both type 1 and type 2 diabetes, although with a higher frequency in T1D (*127*). This antibody was also used in the DiViD study, and interestingly all six donors and two out of the nine non-diabetic controls were positive for VP1 in islets (*122*). This make cross-reactivity of the VP1 antibody difficult to exclude and as such results from studies employing this reagent should be interpreted with caution.

It is challenging to summarise over 50 years of research examining the link between type T1D development and enterovirus infections as the results have been as varied as the studies conducted. To date such a causal link has yet to be convincingly proven although a substantial amount of circumstantial observations have been reported. If such an association holds true however, several mechanisms through which enteroviruses could trigger T1D have been proposed (*128*). One process could be via the direct cytolytic infection of β-cells causing cell lysis and thereby disease development. Perhaps a more likely scenario that ties in with the conventional understanding of T1D pathology as an autoimmune condition (*129*, *130*) would be an autoimmune response triggered by an enterovirus infection. Prolonged persistent infections of enteroviruses in various tissues have been observed, in for instance the heart, gut and pancreas (*131–133*). If β-cells would sustain a persistent infection it could potentially break self-tolerance by presenting viral antigens via major histocompatibility complex (MHC) class I, thereby exposing the cell to the immune system. There is also the possibility that a viral antigen shares an immunogenic epitope with a host protein. This could lead to bystander activation causing activation of autoreactive T and B cells and subsequent β-cell destruction (*134*).

1.3.2 ENTEROVIRUS IMMUNE EVASION STRATEGIES

The first main obstacle that an invading pathogen faces upon entering the host cell is innate immunity and the need to overcome the interferon (IFN)-mediated antiviral response. This response consists of the recognition of viral nucleic acids, known as pathogen-associated molecular patterns (PAMP). These PAMPs, in the case of picornaviruses, are detected by intracellular proteins known as pattern recognition receptors (PRRs) that consist of the toll like receptors (TLRs), a large family of receptors (e.g. TLR3, TLR7, TLR8 and TLR9) to name a few, all of which have their own preference for specific PAMPs. Cytosolic retinoic-acid inducible gene I (RIG-I) or melanoma differentiation-associated 5 (MDA5) are other known PRRs (*135*). The recognition of viral nucleic acids by PRRs then leads to the production and secretion of Type I interferons (IFNs), and thus by either autocrine or paracrine signalling through the type I IFN receptor (IFNAR), an antiviral state is induced in the target cells, promoting restriction of viral replication or virus clearance (*136*, *137*).

The ways in which enteroviruses counter the host antiviral response are diverse (Figure 4). However, a strong theme in the different strategies that enteroviruses have evolved to perturb innate immunity responses revolves around the employment of the virus encoded proteases $2A^{pro}$ and $3C^{pro}$. During infection with different enteroviruses, such as poliovirus, CVB3 or EV71, the viral sensors RIG-I and MDA5 are directly cleaved by $3C^{pro}$ (*138*, *139*) and MDA5 by $2A^{pro}$ respectively. Other proteins involved in inducing and mediating the IFN response are likewise targeted. For instance, RIG-I/MDA5 downstream signalling proteins such as mitochondrial antiviral signalling protein (MAVS) is cleaved by $2A^{pro}$ (*139–141*). Moreover, the recognition of viral nucleic acid in the endocytic compartment by TLR3, is blocked by $3C^{Pro}$ cleavage of the signalling adaptor protein, TIR Domain Containing Adaptor Inducing Interferon-Beta (TRIF). Even at the level of transcription factors, several proteins involved in the activation of interferon transcription are targeted by the proteases, like IRF7 (*142*, *143*) and IRF9 (*144*) which are both susceptible to cleavage by $3C^{pro}$ in EV71 infections. These observations demonstrate the multiple intersections between the viral proteases and innate immune signalling pathways, with protease mediated modulation seen in the early stages of recognition, during the propagation of signals and finally when activation of interferon transcription occurs.

Aside from interfering with innate signaling pathways, enteroviruses can also modify host gene expression to perturb interferon production at the stage of RNA processing and trafficking. Enterovirus encoded proteases are involved in disrupting RNA polymerase I, II, and III transcription (*145–147*). Several host proteins involved in transcription have been identified as viral protease substrates shedding light on how the disruption of transcription is accomplished (*92*). Moreover, during enterovirus infection nucleocytoplasmic transport is altered with the subsequent accumulation of some certain host factors in the cytoplasm. These host factors either stimulate the IRES-mediated translation of viral polyprotein or mediate the switch between viral translation and RNA replication, and include La autoantigen (La) (*148–151*), Poly(RC) Binding protein 2 (PCBP2) (*152–154*), and polypyrimidine tract binding protein 1 (PTB) (*155*, *156*), Splicing Factor, Arginine/Serine-Rich 20-kD (SRP20) (*157*). The precise process by which nucleocytoplasmic transport is modified during infection is complex and not fully understood, but it is thought that in part, it is due to the proteolytic cleavage by $2A^{pro}$ of several proteins, such as Nup62, Nup98 and Nup153, which are part of the nuclear pore complex (NPC) and this has been described for both poliovirus and rhinovirus (*158–160*).

A central theme in the enterovirus replication cycle is the shutting down of host translation, thereby effectively blocking the expression of antiviral genes. This modulation needs to be balanced and precise as the same cellular machinery is also adopted by the virus to produce it's viral mRNA and proteins. In exerting control of host translation, enteroviruses have evolved to cleave the translational initiation factors such as, eukaryotic translation initiation factor 4 G (EIF4G) and poly (A)-binding protein (PABP) (*161*). EIF4G is a central scaffolding protein essential for the recruitment of the cap-binding protein eIF4F to the ribosome and the mediation of cap-dependent translation (*162*), hence, cleavage of EIF4G represses cap-dependent translation. To achieve complete blockage of cap-dependent translation, PABP likewise cleaved by the viral proteases. PABP binds the 3′ poly (A) tails of host RNA transcripts, forming the loop structure of actively translating mRNAs. PABP

cleavage has been attributed to both $2A^{pro}$ and $3C^{pro}$ in poliovirus and CVB3 infection (*142*, *163*). Viral protease mediated cleavage events are on the other hand not the only mode through which enteroviruses can modulate host translation. Enterovirus induced expression of host micro RNA miR-141 which leads to down-regulation of the translation initiation factor eIF4E which in turn negatively impacts cap-dependent translation while leaving cap-independent translation unaffected (*164*).

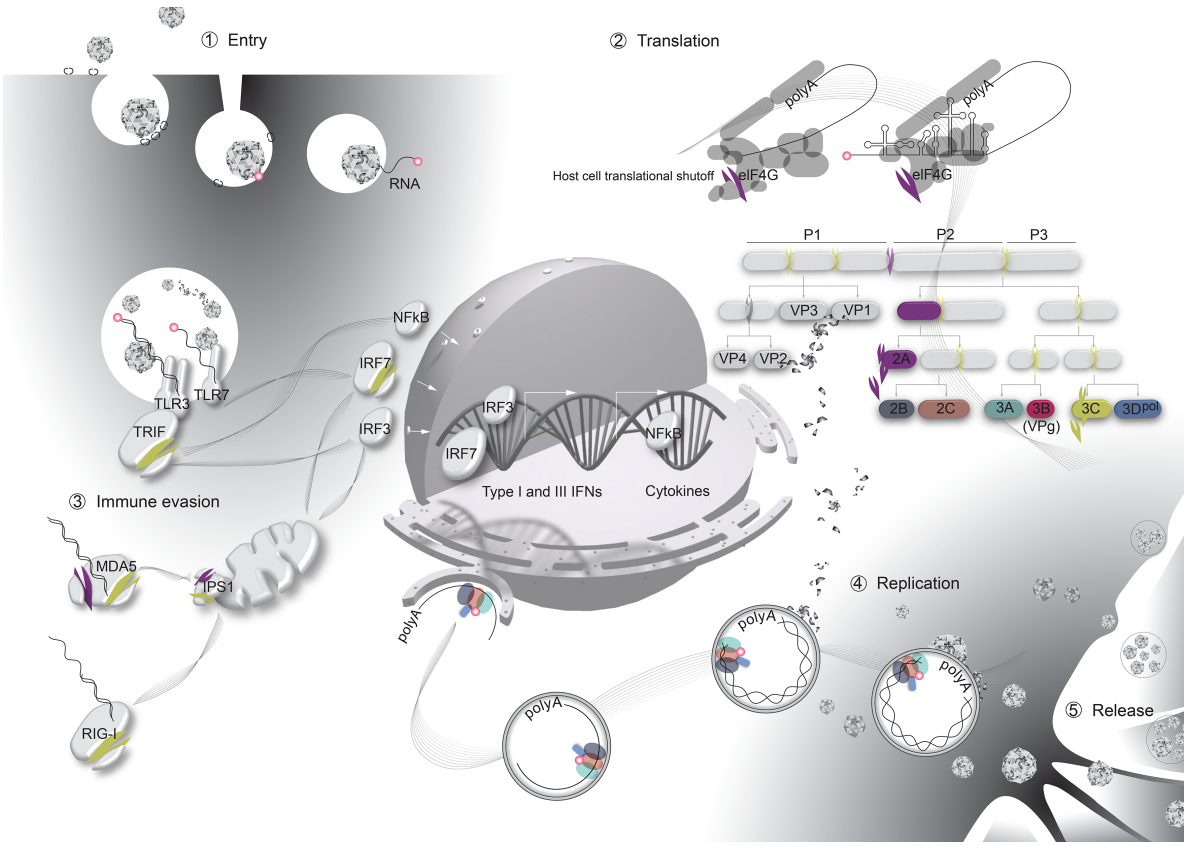

Figure 4. Viral replication cycle and innate immune evasion. *(1) Entry of the virus into the cell mediated via transmembral surface receptors leading to the release of viral RNA into cytoplasm. (2) Cap-independent translation and subsequent processing of viral polyprotein into functional structural and non-structural proteins. (3) Modulation of innate signalling pathways, leading to the perturbation of the interferon response. (4) Replication of viral RNA and packaging into newly formed virions. (5) Release of virus progeny either by cell lysis or by extracellular vesicles. Figure adapted from* (*93*) *with permission.*

Several strategies used by enteroviruses, to achieve immune evasion have been discussed here, namely the interruption of innate signalling pathways, and the direct targeting transcription or translation. The examples given were chosen on the basis of the involvement of the viral proteases in each evasion strategy. There are however more unconventional ways in which enteroviruses have evolved to evade detection or counter an immune response. It has been shown that enteroviruses such as CVB3 and EV71 can be contained in and therefore disseminated from an infected cell through the shedding in extracellular vesicles (*165*, *166*) thereby possibly avoiding their detection by neutralising antibodies. Interestingly EV71 produces vesicles enriched in miRNA-146A, a micro RNA

with the ability to suppress the expression of interferon I, thereby promoting viral replication and spread (*165*). These are just some of the ways in which enteroviruses have evolved to circumvent an effective immune response during infection. Although there have been many advances in our understanding of these processes over the decades, much remains to be discovered, especially in regards of finding new viral protease substrates which upon targeting could perturb the innate immune response.

2 AIMS OF THE THESIS

The body of work that makes up this thesis has as its objective to provide insights into enterovirus biology and to elucidate the relationship between infections and human disease. The experimental work focuses on defining the viral protease substrate repertoire and provides mechanistic explanation of how enteroviruses may cause β-cell dysfunction, in the possible context of T1D development.

Specific aims:

Paper I: To investigate if CVBs are capable of perturbing type III interferon production upon infection by subverting innate immune signalling pathways via the activity of virus encoded proteases $2A^{pro}$ and $3C^{pro}$.

Paper II: To develop highly specific antibodies against the viral encoded proteases $2A^{pro}$ and $3C^{pro}$, that could be used in a wide variety of methodologies for enterovirus detection.

Paper III: To investigate the impact that CVB infection and the ectopic expression of viral proteases $2A^{pro}$ and $3C^{pro}$ have on β-cells and their capacity to secrete insulin.

Paper IV: To perform a thorough survey of the CVB specific $2A^{pro}$ and $3C^{pro}$ substrate repertoire in multiple cell lines. Furthermore, to explore the functional impact that changes in the proteolytic landscape might have on the infected cell and how this might relate to human disease.

3 METHODOLOGICAL APPROACH

3.1 ETHICAL CONSIDERATIONS

In **paper II** experiments were conducted using materials derived from the following murine model strains, C57BL6 (JAX stock #000664) and non-obese diabetic mice (NOD) (*167*). These mice were kept according to institutional and national regulatory guidelines and with appropriate ethical permission. In **paper III,** infection experiments were carried out using human pancreatic islets isolated from organ donors by the Nordic Network for Clinical Islet Transplantation, Uppsala University. All experiments using human derived material were conducted according to the ethical directives described in the Declaration of Helsinki and approved by the regional ethics committee. The numbers for the relevant ethical applications submitted are as follows: Dnr. S46-14 (**paper II**), Dnr. 489/03 (**paper III**).

3.2 VIRUSES

The well characterised and widely used CVB3 strain Nancy was used in nearly all studies (**paper I, II III**) with the addition of CVB4 E2 in **paper III**. In **paper IV** in the infection experiments the virus strain used was CVB3 Woodruff encoding a non-fused enhanced green fluorescent protein (EGFP) as a marker (*168*). In **paper IV** Vesicular Stomatitis Virus (VSV-GFP) was also used (*169*). In **paper II** a cell array containing GMK, HeLa and RD cells infected with various virus species was utilised (*170*). For all of the virus stocks used in the different studies, the viruses had been propagated and titered in HeLa cells.

3.3 CELL LINES

Depending on the experimental context and hypothesis tested, various cell lines were used. In **paper I and paper II**, HeLa cells were used as a cell model as they are highly permissive to CVBs and can mount an intact type III interferon response upon viral infection (*171*). For the studies examining insulin secretion after CVB infection (**paper III**) the rat β-cell line INS-1 832/13 was chosen as a model due to its previously established suitability in examining β-cell insulin secretion (*172*). In the final study (**paper IV**) that aimed to define the protease substrate repertoire three different cell lines were included: HeLa, CaCo-2 (human epithelial colorectal adenocarcinoma cells) and EndoC-βH1 (genetically engineered human beta cell line) (*173*).

3.4 IMMUNOCYTO- AND IMMUNOHISTOCHEMISTRY

In probing CVB infected pancreatic tissue **in paper II** immunocyto- and immunohistochemistry were used. Tissue that had been formalin fixed and paraffin embedded (FFPE), sectioned, heat-fixed and treated for antigen retrieval. The antibodies developed against the proteases had been titered for optimal performance and detected by a biotinylated anti-rat IgG antibody. Visualization could then be achieved by using a horseradish peroxidase-coupled streptavidin reagent.

3.5 RNA ISOLATION AND REAL-TIME PCR

Expression of interferon induction, cellular surface receptors in **paper I** or the presence of viral RNA, insulin in **paper III** were estimated by isolating RNA from samples and using RT-PCR as described in (*141*, *174*).

3.6 FLOW CYTOMETRY

To validate robust infection of target cells with CVB in **paper I, II, IV**, cells were fixed in paraformaldehyde and analysed by flow cytometry. VP1 or EGFP tagged virus were detected using BD FACS Calibur or BD Accuri, analysis of data was performed with the FlowJo software, version 10.

3.7 WESTERN BLOT

To detect proteins as well as cleavage products, western blot was performed in **paper I, II, III, IV** according to standard procedure (*175*) separating protein lysates by SDS-PAGE and transferring proteins to polyvinylidene difluoride (PVDF) membrane using the I-BLOT system (*176*). P**aper IV** utilised protein transfer to nitrocellulose membrane and for protein detection and for quantifiable fluorescence based western blot (QFWB) the modern LI-COR Odyssey (*177*) was used.

3.8 MONITORING OF HORMONE SECRETION

The degree of insulin release was evaluated by stimulating β-cells with glucose and then measure insulin content in media by enzyme-linked immunosorbent assay (ELISA). In **paper III** glucose stimulated insulin secretion was carried out using islets from organ donors or INS-1 832/13. Due to the technical difficulties in performing this study multiple complementary methodological approaches were applied as described below.

First, to evaluate the effect of viral proteases on insulin secretion, INS-1 832/13 cells were transfected with constructs containing $2A^{pro}$ or $3C^{pro}$. These experiments were problematic due to low transfection efficiency (10-15%). In order to address this issue cells were co-transfected with a construct containing human growth hormone (hGH). This approach has previously been used as a proxy for evaluating secretion as hGH, when ectopically expressed, is co-secreted with insulin (*178*). This made it possible to study insulin secretion despite low transfection efficiency.

Secondly, to further evaluate the effect of viral protease expression on the secretion of individual INS-1 832/13 β-cells, patch clamp studies were conducted. Patch clamp is an electrophysiological method used to monitor ion currents by measuring changes in the overall capacitance of a living cell and subsequently depolarization and insulin secretion (*179*). β-cells were first transfected with non-structural viral proteins. The day after transfection individual cells that were GFP positive were pierced with a fine glass pipette. As transfection of β-cells with $2A^{pro}$ were both cytotoxic and blocked efficient protein

translation, the patch clamp studies were carried out by adding active recombinant $2A^{pro}$ to the pipette. Calcium currents were evaluated by depolarising cells from -70 mV or -50 mV to +50 mV for 50 ms. To study exocytosis the cells were subjected to a sequence of depolarisations ranging from -70 mV to 0 mV. As insulin granules dock with the plasma membrane and release their cargo to the extracellular space, the cell surface area changes and thereby the cell membrane capacitance, which indirectly indicates the secretion efficiency.

Thirdly, validation of previous results generated from patch clamp experiments was performed by an independent method applying total internal reflection fluorescence (TIRF) microscopy (*180*). This allows for the monitoring and quantification of cell surface events such as fusion of insulin granules with the plasma membrane.

3.9 SUBTILIGASE LABELLING

In **paper IV** for the purpose of protease substrate discovery, we applied a method known as subtiligase labelling (*67*). As newly translated proteins become acetylated at the α-amino-termini, any unblocked peptides will be mainly produced after proteolytic cleavage. This is a feature exploited in this labelling strategy. Through the activity of the genetically engineered subtiligase enzyme, unblocked N-termini are labelled with biotinylated synthetic peptide (harbouring an amino acid sequence ENLYFQS) and a non-natural amino acid mass tag (α-aminobutyric acid, Abu). Peptides are then bound on streptavidin beads and later released by TEV protease mediate cleavage within the TEV recognition site (ENLYFQ↓S). This approach is designed to enrich the N-terminal side of peptides in complex samples by a catch-and-release strategy for further peptide identification by LC-MS/MS. A comparison with other substrate discovery methodologies and a more detailed description and discussion regarding of subtiligase labelling can be found in section 1.1.7 of this thesis.

3.10 LENTIVIRUS TRANSDUCTION

In order to stably express a protein that was resistant to $2A^{pro}$ and $3C^{pro}$ cleavage (and thereby prevent protease mediated cleavage) in HeLa cells, lentivirus transduction was used in **paper IV** (*181*, *182*).

3.11 BIOINFORMATIC ANALYSIS

For figure generation and statistical testing in **papers I and III,** the commercially available GraphPad software version 8 was used. In **paper IV** custom code was written using the scripting language python (3.6), aside from the python standard library several other libraries were implemented. These include (but are not limited to), Pandas (*183*), Matplotlib (*184*), Numpy (*185*) and Networkx (*186*). In the R (3.5) scripting language the following packages were utilised for gene ontology enrichment analysis and pathway analysis; Disease Ontology Semantic and Enrichment analysis (DOSE; (*187*) and ReactomePA (*188*). Publically available databases such as DisGenet (*189*), DegraBase (*72*), IntAct (*190*),

and Reactome (*191*) were also integrated into the analysis pipeline. For generating amino acid sequence logos the R package ggseqlogo was implemented (*192*). Statistical analysis was performed using the python library SciPy (*193*) or by built in statistical test in the software Prism.

4 RESULTS AND DISCUSSION

In the four papers in my thesis, I have investigated different aspects of enterovirus biology. With a special focus on studying the proteolysis mediated by viral proteases and which role substrate cleavage events have in the virus replication cycle. In **paper I** it is shown how CVB3 has evolved mechanisms to block type III interferon expression during infection and how $2A^{pro}$ are directly involved in perturbing the innate immune response. **Paper II** reports the development of $2A^{pro}$ and $3C^{pro}$ specific antibodies and their utility in detecting EV-B species in a range of methodologies. **Paper III** is an investigation of the cellular dysfunction observed in insulin secreting β-cells during CVB infection, showing that the viral proteases may contribute to this phenotype. Finally, **paper IV** is a comprehensive survey of the $2A^{pro}$ and $3C^{pro}$ specific substrate repertoire in multiple cells lines. Collectively, these papers define the host protein cleavage events mediated by the viral proteases and their functional/phenotypic relevance is explored in the context of enterovirus associated disease.

4.1 COXSACKIEVIRUS COUNTERS THE HOST IMMUNE RESPONSE BY BLOCKING TYPE III INTERFERON EXPRESSION (PAPER I)

The arms race between viruses and their hosts has led to co-evolution and an intricate balance in the interactions between pathogen and host, with each attempting to counter the actions of the other. For enteroviruses, as previously discussed in the introduction (chapter 1.3.2), immune evasion is multifaceted and occurs at multiple points during the virus replication cycle. In an excellent study from 2011 by Mukherjee *et al.* (*140*), it was demonstrated that CVB3 attenuates the host response by cleaving MAVS (also known as IPS-1) and TRIF via $3C^{pro}$ activity during infection, thereby perturbing innate immune signalling. Ultimately this perturbation of key adaptor molecules blocked the type I interferon response rendering the cells susceptible to infection.

The type III interferon (IFN-λ) family have been shown to exert broad antiviral activity, especially at mucosal surfaces where the IFN-λs are predominantly expressed (*194*, *195*). The primary infection site for enteroviruses is the gastrointestinal tract and we therefore set out to investigate if, in a similar manner to the type I IFNs, CVB3 was capable of blocking type III response.

HeLa cells were chosen as a model system since the cell line is readily infected by enteroviruses. Detection of CVBs in mammalian cells is accomplished by PRRs such as MDA5, RIG-I and TLR3, that sense single or double stranded viral RNA. To simulate the uncoating of endocytosed virions, Polyinosinic:polycytidylic acid (poly I:C), a RNA double stranded mimic, was used to stimulate the cells. Poly I:C was added directly to the culture media, to be taken up into the cells by endocytosis and thereby allowing detection by TLR3 localised to the endocytic compartment. To facilitate detection by the cytosolic RIG-I and MDA5 the RNA double stranded mimic was also transfected into the cells. Poly I:C

treatment but not CVB3 infection was able to induce the expression of type III IFNs (Figure 1). Indeed, if cells were infected prior to treatment with poly I:C there was a significant reduction in the IFNλ1, IFNλ2 and IFNβ mRNA expression (Figure 2), indicating that CVB3 actively inhibited the induction of type I and type III IFNs.

Earlier studies reported that the expression of type III interferons is dependent on IRF3 activation due to hyperphosporylation of the protein at multiple sites (*196*). By western blot analysis it was observed that CVB3 infection blocks this hyperphosphorylation event (Figure 3B). Due to the alterations in the phosphorylation state or IRF3, further upstream signalling proteins were evaluated by western blot. These experiments showed clear reduction in the protein levels of TRIF and IPS-1 (Figure 4A, B), and interestingly smaller sized protein fragments were also detected. These fragments were thought to be cleavage products of the full length TRIF and IPS-1 proteins. Moreover, the appearance of the cleavage fragments was independent of caspase and proteasome activity during CVB3 infection, as treatment of cells with chemical inhibitors of apoptosis and proteasome degradation did not rescue the expression of TRIF or IPS-1 (Figure 5).

Overexpression of $2A^{pro}$ or $3C^{pro}$ by cloned constructs transfected into HeLa cells showed that $2A^{pro}$ robustly dampened the mRNA expression of IFNλ1, IFNλ2 and IFNβ following poly I:C treatment whereas no effect was observed when overexpressing $3C^{pro}$ (Figure 6A ,B ,C). To further support the hypothesis that $2A^{pro}$ was the protease responsible for the cleavage and subsequent decrease of TRIF and IPS-1 during CVB3 infection, lysates derived from HeLa cells was incubated with $2A^{pro}$ active recombinant protease. Western blot analysis of these lysates detected similar sized cleavage fragments as those observed in during CVB3 infection of HeLa cells (Figure 7). Suggesting that it is indeed $2A^{pro}$ mediated cleavage, that is the underlying cause for TRIF and IPS-1 reduction during CVB3 infection.

It has been shown that perturbation of interferon I production is due to $3C^{pro}$ mediated cleavage of IPS-1 and TRIF during CVB3 infection in HEK293 cells (*140*). Although, our observations are more in line with those made by Feng *et al*. 2014 (*139*) who also observed that IPS-1 degradation is $2A^{pro}$ mediated. In EV71 infection, IPS-1 cleavage also seems to result from $2A^{pro}$, as shown by Wang *et al.* 2013 (*197*). The discrepancies between these studies could be due to multiple reasons, for instance differing time-points, cell lines and virus strains, or potential overlap between $2A^{pro}$ and $3C^{pro}$ cleavage specificity.

In **paper I**, we can conclude that CVBs target the innate signalling pathway and this is in part achieved by $2A^{pro}$ mediated cleavage of IPS-1 and TRIF during CVB3 infection. This leads to a block in the expression of type III interferons, which in turn is likely to promote both the replication and spread of the virus in an infected host.

4.2 NEW COXSACKIEVIRUS 2A^{PRO} AND 3C^{PRO} PROTEASE ANTIBODIES FOR VIRUS DETECTION AND DISCOVERY OF PATHOGENIC MECHANISMS (PAPER II)

Epidemiological studies have suggested a link between type 1 diabetes (T1D) and enterovirus infections (reviewed in section 1.3.1). Many of these studies relied on immunohistochemistry (IHC) staining of pancreatic tissue from T1D organ donors using an antibody directed against the enteroviral capsid protein VP1 (clone 5D8/1, Dako). Concerns have been raised that this widely used antibody may have cross-reactivity with two endogenous cellular proteins, namely a creatine kinase and mitochondrial ATP synthase (*124*). Further characterisations have shown that this VP1 antibody binds to a linearized epitope consisting of a conserved sequence that is present in many enteroviruses, but may also occur in host proteins (*198*). As such, we identified the need for new tools for use in detecting enteroviruses in serum and tissue samples and for reagents that could be used in furthering our understanding of enterovirus biology. We developed three new monoclonal antibodies directed against the enterovirus encoded proteases 2A^{pro} and 3C^{pro}.

Using the genetic sequence derived from CVB3 Nancy (UniProtKB/Swiss-Prot: P03313.4), his-tagged 2A^{pro} and 3C^{pro} were cloned into bacterial expression vectors and recombinant his-tagged protein was produced *in E. coli* and purified. Recombinant proteins were used to immunise Wistar rats (*199*) and to generate hybridomas for monoclonal antibody production. The clones were selected based on their detection of CVB3 in infected HeLa cell lysates by wb and through their reactivity in IHC when screening pancreas derived from CVB3 infected mice. Two clones for 2A^{pro} denoted as 2A-1 and 2A-2 and one for 3C^{pro} denoted 3C were identified and selected for large-scale antibody production.

The earliest time point that the proteases were detected was at 4 hours post infection (hpi) and their expression remained stable until cell lysis occurred at 9 hpi (Figure 1C). Time-course experiment showed that 2A^{pro} and 3C^{pro} expression kinetics mirrored those of the capsid protein VP1. In western blot both 2A^{pro} (2A-I) and 3C^{pro} (3C) antibodies were able to detect all six of the different CVB serotypes (CVB1-6; Figure 1D) in HeLa cells 6 hpi. Analysis of CVB3 infected HeLa cells by flow cytometry confirmed the efficacy of the antibodies for detection of viral proteases with an orthologous methodology (Figure 3A-C), and once again, staining for the proteases was in concordance with that for VP1 expression.

When the antibodies were employed in IHC using a cell array containing cells infected with virus species from enterovirus group EV-A and EV-B, both 2A-II and 3C were able to detect the proteases of all species that belong to the EV-B but not EV-A (Table 1). The only exception to this was 2A-2 which showed no staining for Echovirus 9 (an EV-B virus) in infected cells. Moreover, excellent specificity of the antibodies for the EV-B species was demonstrated as the new 2A^{pro} and 3C^{pro} antibodies did not detect any viruses belonging to the enterovirus A group. When the antibodies were used to probe pancreas tissue from CVB3 or CVB5 infected NOD mice, it was possible to confirm expression of the viral proteases on days 3 and 5 post infection (dpi) (Figure 2).

Translocation of cellular proteins from the nucleolus to the cytoplasm during infection with CVBs is known to occur during infection (reviewed in chapter 1.3.2). Also viral proteins have been known to enter the cell nucleus, localisation of $3C^{pro}$ has previously been described (*200*, *201*), although the protease itself lacks the required nuclear localisation signal (NLS). However in poliovirus, rhinovirus and FMDV, the viral RNA polymerase 3D harbours such a signal (*202–204*). It has been suggested that during infection $3C^{pro}$ localises to the nucleus as a fusion protein (3CD), but that this process is dependent on the activity of $2A^{pro}$ (*205*). Confocal microscopy analysis of time course infections of HeLa cells with CVB3, using the $2A^{pro}$ and $3C^{pro}$ revealed robust expression of the viral proteases at 4 hpi (Figure 4A-E). In our cell system, nuclear localisation of either of the viral proteases was negligible between 2-6 hpi. This raises the question of whether such a nuclear localisation of $3C^{pro}$ occurs during CVB3 infection. More detailed examination of this is required, although it should be noted that a NLS sequence in CVB3 3D has not been reported.

To summarise, **paper II** describes the development of $2A^{pro}$ and $3C^{pro}$ specific antibodies. The antibodies generated can be used in different applications for the detection of EV-B species and will be of value when examining infection kinetics or cross-validating other detection methodologies.

4.3 COXSACKIEVIRUS AFFECTS MULTIPLE STEPS IN INSULIN SECRETION PATHWAY LEADING TO IMPAIRED INSULIN RELEASE BY INFECTED BETA CELLS (PAPER III)

Enteroviruses have been found in the pancreatic islets of patients with T1D, suggesting that direct infection of the β-cells might contribute to diabetes development. This hypothesis is supported by *in vitro* studies demonstrating that β-cells infected with CVBs have impairments in glucose stimulated insulin release (*206*, *207*).

The virus encoded proteases $2A^{pro}$ and $3C^{pro}$ are known to cleave an extensive range of host proteins during infection (*93*). In **paper III** we wanted to explore the hypothesis that β-cell dysfunction observed during CVB infection is due to the expression of $2A^{pro}$ and $3C^{pro}$. Additionally, we also wanted to determine whether the expression of the non-structural protein 3A has an effect on β-cells, as this viral protein has previously been reported to block the secretion of the cytokines IL-6 and IL-8 (*208*) .

We first infected human pancreatic islets with CVB3 or CVB4, and validated that the CVBs replicate successfully in islets and increased in number through monitoring the release of infectious virions in the islet culture media at 48 hpi (Figure 1A). Upon glucose challenge, infected islets showed a trend towards a reduced GSIS (Figure 1B). Under the same experimental conditions islets were also evaluated for their insulin content (Figure 1C), and insulin mRNA expression (Figure 1D), however no statistically significant decreases in insulin mRNA or content were observed. These results supported the conclusion that lowered insulin responses upon glucose challenge were due to a defect in insulin secretion rather than alterations in the production of insulin during infection.

Next, we wanted to reproduce these findings in a pancreatic β-cell line of rat origin (INS-1-832/13 cells), which supports CVB replication (*209*). Several studies have also demonstrated that INS-1-832/13 cells are a valid cell system for studying insulin secretion (*172*). Optimisation experiments examining suitable time-points and multiplicity of infection (MOIs) were conducted, GSIS was then conducted at 5 hpi after infecting cells with CVB4 at 100 MOI. Similar to the previous observations in infected islets, INS1-832/13 cells displayed a trend towards reduced glucose stimulated insulin secretion following infection with CVB4 (Figure 2C). We were also able to validate the presence of the non-structural proteins $2A^{pro}$, $3C^{pro}$ and 3A by PCR or western blot, in samples used in the insulin secretion experiments (Figures 3B and 3C), confirming their expression under our experimental conditions.

To further investigate the effect that the non-structural proteins exerted on β-cell function, $2A^{pro}$, $3C^{pro}$ and 3A derived from the CVB3 sequence were cloned into vectors and used in transfection experiments in INS1-832/13 cells. Initially, low transfection efficiencies of INS1-832/13 cells were observed. Therefore to evaluate the effect of the expression of non-structural proteins on glucose stimulated hormone release, a modified transfection protocol was adopted. In cells overexpressing human growth hormone (hGH), hGH is localised in the same vesicles as insulin and as such, these cells co-secrete hGH with insulin upon

glucose challenge (*178*, *210*). Thus, by co-transfecting hGH constructs with constructs encoding the viral non-structural proteins, hGH can be measured as a proxy for insulin secretion. This experimental design provided a platform to evaluate the effect that the ectopic expression of non-structural viral proteins has on β-cell function, despite low transfection efficiency in INS1-832/13 cells through selectively evaluating cells that have been successfully transfected with both constructs. The viral proteases are known to affect both translation and transcription of host cell mRNA (*92*, *211*), and this was also evident in our experiments as well. The level of hGH was undetectable in $2A^{pro}$ co-transfected INS1-832/13 cells and reduced in those expressing $3C^{pro}$. In contrast cells expressing 3A had higher levels of hGH (Figure 4C). As a result of this, changes in glucose stimulated insulin secretion could only be evaluated in $3C^{pro}$ and 3A transfected cells and overall, the changes in secretory capacity were not significantly different for either protein (Figure 4D), although there was a trend towards a reduction in the fold change differences between basal and glucose stimulated hGH secretion for both.

Monitoring of cell capacitance during patch clamp after depolarising the cell with a train of ten 500-ms fluctuations between 0 to -70mV allows for the degree of exocytosis to be estimated. For all three non-structural proteins ($2A^{pro}$, $3C^{pro}$ and 3A) a significant reduction in membrane capacitance was be recorded (Figures 5B, D, F). By using an independent method for measuring exocytosis, namely total internal reflection fluorescence (TIRF) microscopy, it was possible to confirm the observations made in the patch clamp experiments. Results showed the expression of $2A^{pro}$, $3C^{pro}$ and 3A had a clear negative impact on exocytosis (Figure 6C, D, F, G), while granular density of docked vesicles at plasma membrane was only slightly less for $2A^{pro}$ as compared to control (Figure 6E and H). Together this suggest that the negative effect exerted by viral proteins occur at a late stage of exocytosis.

The process of exocytosis is driven by an influx of extracellular calcium (Ca^{2+}) into the cell through voltage-gated calcium channels (*212*). Therefore, calcium currents were measured by patch clamp in transfected INS1-832/13 cells. Expression of either of the proteases decreased the Ca^{2+} currents while 3A did not have the same effect (7A-C), possibly indicating modulation of the voltage gated Ca^{2+} channels at the plasma membrane or some process regulating Ca^{2} flux.

The results presented in **paper III**, support the previous observations that CVB infection causes β-cell dysfunction *in vitro* and that infection negatively impacts the cells capacity to secrete insulin upon glucose stimulation. Through using different methodologies aimed at evaluating exocytosis in β-cells (patch clamp, TIRF microscopy, hGH secretion assay), we demonstrate that the overexpression of CVB3 non-structural proteins $2A^{pro}$, $3C^{pro}$ and 3A in INS1-832/13 cells inhibits different aspects of insulin secretion in beta cells. As such, viral non-structural proteins could contribute to the β-cell dysfunction observed during CVB infection.

4.4 COXSACKIE B VIRUS PROTINASE SUBSTRATE REPERTOIRE DEFINED IN HUMAN PROTEOME BY N-TERMINAL PROTEOMICS UNCOVERS CLEAVAGE OF TCF7L2 BY $2A^{PRO}$ (PAPER IV)

Endogenous proteases regulate many processes within the cell by cleaving host proteins leading to alterations in their function, rewiring of protein-protein interactions or shifts in subcellular localisation. Aside from processing the viral polyprotein, the viral proteases $2A^{pro}$ and $3C^{pro}$ have essential roles in modulating the host cell to favour virus replication and spread.

To understand enterovirus biology and to find a mechanistic explanation for the observed phenotype of β-cells transfected with $2A^{pro}$ and $3C^{pro}$ (**paper III**) we wanted to define the specific substrate repertoire of the viral proteases. The methodology chosen for this study known as subtiligase labelling (discussed in chapter 1.1.7) (*15*, *67*). Two distinct sample types were used; protein lysates derived from CVB3 (EGFP tagged) infected cells (forward experiment) and protein lysates incubated with recombinant $2A^{pro}$/$3C^{pro}$ or catalytically inactive mutants (reverse experiment) (Figure 2). The following cell lines, HeLa, CaCo-2, EndoC-βH1 (Table 1) were included in the study in order to cover a wide range of potential substrates, as CVB infection kinetics and the abundance of individual proteins might vary between cell lines.

First, the suitability of the samples was validated. Infected cells all had robust replication of CVB3 at 4-6 hpi (Figure 2A, B). The cleavage of the previously defined CVB substrates MAVS ($2A^{pro}$ (*139*, *141*)) and G3BP1 ($3C^{pro}$ (*213*)) were observed under experimental conditions (Figure 2C). This provided us with confidence that the chosen experimental conditions were suitable for detecting substrate cleavage with the subtiligase labelling method. Indeed, analysis of LC-MS/MS raw data and subsequent data cleaning provided us with a number of unique peptides in each data set that were derived from proteolytically processed proteins. Due to the low number of replicates (three) for each sample type and cell line we decided to combine all of the peptides detected in each sample type (grouping data from all cell lines). Then by overlapping the data sets between the forward and reverse experiments we were able to validate a number of substrates that are cleaved both during infection and in lysates digested with recombinant $2A^{pro}$ or $3C^{pro}$ (Figure 3A). The proteins that overlapped between the forward and reverse experiments are subsequently referred to as validated substrates. Detected peptides matched the cleavage specificity of the viral proteases as defined by the cleavage positions in the viral polyprotein (Figure 3B). This gave us confidence that the peptides detected were indeed derived from viral protease substrates and not generated by endogenous cellular proteases.

The validated substrates numbered 81 proteins in total (Figure 4) and the majority of substrate cleavage sites could be assigned to either $2A^{pro}$ or $3C^{pro}$, although there were seven exceptions where the cut positions were found to be overlapping between both proteases (Figure 3A). In addition, some proteins also have multiple cut sites belonging to

the same viral protease, examples of which include AGFG1, CBX8, CEP170, CSTF2 and TAF15. If these substrates with multiple cut sites carry greater significance when compared to those with single cut site, warrants further investigation.

Pathway analysis of the validated substrates showed statistically significant enrichment for multiple biological pathways (Figure 5). Pathways relating to RNA metabolism and SUMOylation of RNA binding proteins contained the largest number of substrates. The interplay between enterovirus replication and cellular RNA metabolism or SUMOlaytion has been described earlier (*160*, *214–217*). However, the majority of substrates uncovered in this study, that belong to the enriched pathways, have not been identified previously. The detailed description of the functional impact of these cleavage events will further refine our understanding of how enterovirus replication modulates cellular RNA metabolism and SUMOylation. The repression of WNT target gene pathway was also enriched, as represented by the proteins transducin like enhancer of split 3 (TLE3) and transcription factor 7 like 2 (TCF7L2). This pathway was of interest to us as mounting evidence suggest that Wnt signalling modulates normal β-cell function (*218*), especially as genetic variants of the transcription factor TCF7L2 have been linked to type 2 diabetes development (*219–221*). Moreover, an essential role for TCFL2 has also been demonstrated for the maintenance of the glucose stimulated secretory capacity of β-cells (*222*, *223*).

To further explore possible disease associations among the validated substrates, our data was cross referenced with the disease association database DisGenNET (*189*). Of the 17 proteins present in the database there were an association with 24 different diseases, with some substrates being associated to multiple conditions. TCF7L2 was associated with diabetes mellitus (non-insulin dependent) with confidence score of (0.38; Figure 6). Interestingly, a diverse array of diseases are represented among the substrate several that have hallmarks of neuronal diseases, amyotrophic lateral sclerosis (ALS ((*224*)), spinocerebellar ataxia type 2 (SCA2 ((*225*)), infection-induced acute encephalopathy 3 (IIAE3 ((*226*)) and these findings are worthy of note as enterovirus infections are known to cause various neurological complications (*227–229*). If these substrates are in some way important for maintaining normal neuronal function or if cleaved, might contribute to neuronal pathology remains to be explored.

By definition essential proteins are crucial for maintaining the proper function of the cellular machinery. If such a protein in compromised, this could severely affect the viability of the cell. In evaluating the impact of a proteolytic cleavage for the phenotype of a cell, identifying essential proteins among the validated substrates is important. An estimate of essentiality can be made based on a protein location (centrality; (*230*) within a protein-protein interaction network (interactome). In order, to identify essential proteins, a human interactome was built using the IntAct database (*190*), and each substrate was mapped to this network and the betweenness centrality measured (Supplementary 1; (*231*). Ranking of substrates based on centrality indicated that TCF7L2 was among the substrates that scored above the 90^{th} percentile, and is thus likely to be a relatively essential protein.

Additionally, when measuring the distance between validated substrates, it became evident that the average distance between validated substrates was significantly shorter compared to the control substrates (Supplementary 3). In reviewing the nearest neighbours within the network, many of the validated substrates were direct interaction partners (Supplementary 2), supporting the idea that proteolysis by the viral proteases is targeted in the human interactome.

Since there is substantial evidence in the scientific literature describing the importance of TCF7L2 for β-cell function and the likelihood that it is an essential protein based on its betweenness centrality score, we wished to validate its cleavage by an independent methodology. Samples used in subtiligase labelling experiments (HeLa and EndoC-βH1) were analysed by western blot and TCF7L2 was detected. Full-length protein disappeared and potential cleavage products appeared upon CVB3 infection and after protein lysates were incubated with recombinant $2A^{pro}$ (Figure 7A). In control samples or those incubated with catalytically inactive proteases, full-length TCF7L2 remained intact and no cleavage products were observed indicating that TCF7L2 is exclusively cleaved by $2A^{pro}$.

Based on the peptide detected by mass spectrometry the amino acid position of the proteolytic cleavage could be determined and was found to occur between threonine (T) and glycine (G) at position 308 (Supplementary 5B). Wild type (wt) TCF7L2 and cleavage resistant mutants in which glycine had been changed to alanine (A) or glutamic acid (E) at position 308, were stably expressed in HeLa cells by lentivirus transduction. In each TCF7L2 construct a V5-tag at the N-terminus and HA-tag in the C-terminus were introduced. Upon infection with CVB3, each of the terminally labelled smaller cleavage fragments of wt TCF7L2 could be detected, while mutants G308A and G308E were resistant to proteolytic processing (Figure 7B). This confirmed that the 308 position is indeed the site cleaved by the protease. Overexpression of each individual cleavage fragment also indicated that degradation of the N-terminal cleavage fragment occurs during CVB3 infection while the C-terminal fragment remains stable. It is possible that the C-terminal fragment can impact virus replication or cellular function, although this needs to be tested experimentally.

In summary, **paper IV** is an exploration of the viral protease substrate repertoire where 81 high confidence substrates targeted by $2A^{pro}$ / $3C^{pro}$ during CVB3 infection were identified. A combination of pathway and network analysis revealed enrichment of validated substrates in different biological pathways with multiple proteins in close proximity to each other within pathways. Integration of a database containing disease associated genes highlighted TCF7L2 as a target of interest due to its role in maintaining normal β-cell function. Moreover, we validated TCF7L2 as a bona fide $2A^{pro}$ substrate. To our knowledge this is the most extensive survey to date of the CVB3, $2A^{pro}$ and $3C^{pro}$ substrate repertoire. Future studies will focus on defining the relevance of proteolytic cleavage both for virus replication and for the host cell.

5 CONCLUDING REMARKS

This thesis has focused on exploring the biology of enteroviruses and more specifically the virally encoded proteases $2A^{pro}$ and $3C^{pro}$, first and foremost by defining the host cellular proteins that are targeted by the viral proteases during infection. By doing this, we are furthering our knowledge regarding the types of cellular functions that these viruses have evolved to modulate, and this will allow us to start exploring the phenotypic relevance of these cleavage events for both the virus and the host cell. Hopefully some of these observations may provide insights into the pathological processes or diseases that are caused by or associated with enterovirus infections.

In writing this thesis, articles have been cited that span more than 50 years of research. In surveying this scientific literature, it becomes evident that enteroviruses will continue to be of importance to human health both in the near and distant futures. The disease burden attributed to acute and chronic infections is substantial and there is a need for the development of curative and preventive therapies. Success stories regarding the treatment of a diverse number of diseases, such as those caused by human immunodeficiency virus (HIV), hepatitis C virus (HCV) and human papillomavirus (HPV) teaches us that understanding basic molecular virology is essential for attaining efficient treatments.

The American novelist Paul Auster once wrote, “The truth of the story lies in the details”, which is at the heart of a scientist’s purpose, namely to uncover new details. It is my hope that the results presented here will contribute to the collection of details related to the ongoing story of enterovirus biology, allowing us one day to discover the true nature of these fascinating viruses.

Based on the original research articles and manuscripts presented in this thesis, the following findings have been made:

Paper I, describes how CVBs are capable of blocking the type III interferon response during infection by the activity of a viral encoded protease.

Paper II, discusses the development of $2A^{pro}$ and $3C^{pro}$ antibodies, which are highly specific for viruses belonging to EV-B species.

Paper III, investigates the impact of CVB infection on the capacity of β-cell to undergo glucose stimulated insulin release. Showing that some of the negative impact on exocytosis might be due to the expression of the viral non-structural proteins $2A^{pro}$, $3C^{pro}$ and 3A during infection.

Paper IV, is a comprehensive survey of the $2A^{pro}$ and $3C^{pro}$ substrate repertoire, identifying 81 substrates and further validating the protein TCF7L2 as a $2A^{pro}$ substrate. This work provides a potential mechanistic explanation for β-cell dysfunction observed during CVB infection.

6 ACKNOWLEDGEMENTS

In science, the most substantial professional development one can experience is thru the honest and open discussion among colleagues and friends. Regardless if the topic is methods, experiments or projects, it is only by sharing knowledge that it can ever grow and it is in teaching others that we ourselves become real scientists. I was very lucky in that I had good mentors and colleagues during my time as a student, and it is because of this that I was able to learn. I thank you all.

I would like to thank my main supervisor **Malin Flodström-Tullberg**, for providing a place in your group for me as a PhD student and giving me the opportunity to work independently on my projects.

I would like to thank my co-supervisor **Jonas Klingström** for always being available. You are one of the few people I know that can keep a consistent positive outlook regardless of circumstances, and that has been so helpful.

I would also like to thank my co-supervisor **Lena Eliasson** for her feedback related to β-cell physiology.

I would express my deep gratitude to **Charles Rice** for inviting me to his lab and allowing me to work there during my PhD studies. It is difficult to express how the time in your lab meant. Your encouragement and support (and forwarded degradomics papers) the last 3 years have been invaluable.

I would like to thank **Mohsan Saeed**. I can honestly say I have learnt more practical laboratory skills from you then anyone else. You have been so generous with your resources, knowledge and time and I think you have set an example for how all should act in instructing those more junior than ourselves.

My mentor, **Niklas Arnberg,** thank you for giving me the opportunity to do my first virology project in your group all those years back in Umeå. It really promoted an interest in this area of research and it gave me direction for my future studies that have lasted so far.

In spending a number of years in a research group one comes to the realisation of how dynamic this environment is. I would like to thank the members in our group both past and current. **Magdalena Mazur,** for being an incredible colleague and friend, organised, efficient and always looking after everyone else. If human cloning ever becomes a thing I think we should start with cloning you! Every group needs their own **Magda**, such a breakthrough would be truly transformative to science. **Miss Virginia Stone (aka Ginny),** I would so much like to thank you for all the help over the years, dedicated, positive and always so helpful reading over rows and rows of my swinglish. Your valuable input has made me a much better writer then any English writing course ever could. **Isabel Diaz Lozano**, passionate scientist spreading good vibes all over the place. When joining the

group, it was like that scene in Pulp Fiction when Uma Thurman gets stabbed in the heart with an adrenalin shot. Such a great energy boost to the group. **Helena Sork**, from the first time we had a skype regarding western blot I realised you were the real deal, utterly confident and pragmatic researcher with no time to waste, the best kind of researcher in my opinion. **Renata Utrova**, the go to person for when a diagnosis is needed. I got to know you during your studies and I immediately felt that you are the kind of doctor I would like to have, genuine, caring and competent. **Marta Butrym**, newest member to our team, I wish you the best of luck in your projects. I have no doubt that you will do great things. **Soile Tuomela**, people like you make me miss Finland so much. Direct no-nonsense scientist with humour so dark it could start its own heavy metal band. The band would be called Staurosporine and it would be absolutely lethal! I know I speak for everyone when I say that we miss you in the lab. **Erna Domsgen**, thank you for all the fun memories and science discussions in those early days. **Emma Svedin**, as I understand it you were the one reading the applications for the PhD position in the group those many years ago, so thank you for promoting the application. You were such an essential part of our group and it was strange when you defended and moved on. **Pär Larsson**, someone not shying away from an intense discussion about the role of CVB in T1D development and a great snowboarder to boot! **Katarina Lind,** for all that western blot magic and fika.

I would also like to thank our collaborators in Finland at Tampere University, **Olli, Vesa, Juha, Anssi** for reagents, meetings and nice discussions about virology and proteases.

I would like to thank all the people at the Rice lab, for the great science environment, lab meetings, lectures, social activities and after work at faculty club. **Santa, Glen, Aileen, Ed, Joe "The Juice", Eliana, Alison, Mariel, Stephanie, Brandon, Yingpu, Bill, Hachung, Xianfang, Heinrich, Leftaris, Kate, Inna** all of you made me feel so welcome in the lab and contributed to me having great time in New York.

I would like to thank **Johan Sandberg** and his group. **Jean-Baptiste** for all the python support and great discussions for anything related to data analysis. One day "crêpe analysis" will be a reality! **Tiphaine** for being an awesome co-pilot, if I ever need to tackle the Parisian traffic again I want you yelling directions! I would also like to thank **Caroline, Michal, Edwin, Joana, Johanna and David.**

I would like to thank **Anna Norrby-Teglund** and her group. First like to thank some former members, **Nikolai** for interesting discussion regarding human reproduction and for sharing both protocols and reagents. **Julia** for being a great flatmate and enthusiastic colleague. I would also like to thank **Srikanth, Bhavya, Johanna and Helena.**

I would like to thank **Jenny Mjösberg** and her group. Both former member and current members, **Marianne, Luca, Anna R, Themis and Anna C.**

I would like to thank **Kalle Malmberg** and his group. **Aline,** you are one of the most hard-working person I know (and most organised), also thank you for teaching me how to use FlowJo. **Ebba** for sharing some great indie music and letting me know when Finnish/Swedish bands are playing in Stockholm**. Alvaro** for your enthusiasm for science

and after works playing video games. **Quirin** for mixing a great gin and tonic and having interesting science related arguments.

I would like to thank **Niklas Björkström** and his group. **Otto** for all the recommended podcasts**. Christine** for advice on the thesis application process, fun that we both defended this spring. **Natalie** for sharing your hilarious everyday observations from the lab and clinic and being someone to intensely discuss the second world war with. I would also like to thank **Martin C, Martin I, Laura, Ben, Jonna, Iva, Lena and Isabella.**

I would like to thank **Tim Willinger** and his group. Especially, **Natalie Sleier** for sharing your passion for sports. I would also like to thank **Emma, Elza** and **Hana**.

I would like to thank **Yenan Bryceson** and his group. **Mattias** someone who shares the interest in proteases and good beer, an excellent mix for lengthy discussions about experimental science. As well as **Irene, Heinrich, Lamberto, Giovanna, Hongya, Martha, Jelve** and of course the onion knight **Tim**.

I would like to thank, **Jacob Michelsson**, for those dope rap lyrics booming in the red lab and to **Nicole** for her great laughter and keeping it real with some legit virology work. I would also like to thank **Hans-Gustaf Ljunggren** and his group, **Kim, Angelica, Tyler,** as well as the rest of the **Klingström** crew, **Carles, Kimia, Johanna, Marina**, **Shawon** and **Moni**.

I would like to thank **Susanna Brighenti** and her group. **Jagadees,** you are one remarkable doctor. I wish you god luck back home, but I know you will not need it. Also rest of the old office at CIM, **Mily** and **Senait**.

Shout goes out to **Mattias Svensson** for his amazing baking skills and keeping us motivated to present our work at SAP. I would like to thank the rest of his group, **Magda**, **Sam**, **Egle** and **Puran**.

I also would like to thank the people behind the scene, keeping the lights on, pipettes working, storage stocked and salaries rolling in. All making it possible for us to actually do science. **Lena, Elisabeth** and **Margit,** we would end up in a very bad place without you.

I would want to thank **Su** for showing me Boston and for being a constant supportive friend, with the absolute best career advice!

Over the years studying you get to meet so many different people and it is exciting to see the paths people take, I would like to thank the people I studied with in Umeå and Uppsala; **Hanna L, Hanna D, Cecilia, Rickard, Lars, Siri, Emma, Feria, Lillian, Gad, Antonio, Nusa, Emelie, Emmy** and so many more. Special thanks to **Sophia** for everything over the years, parties and dinners and all the virology! One day we will go to Smögen at the same time and it will be the best conference.

Finally, I would like to thank my family. My father **Leif** and mother **Gunilla** for always being there and looking after us. I could not have had better parents, you represent all that I would ever want or hope to be, I love you both. The same is true for sisters **Sabina** and **Daniela,** I can honestly say I do not know what I would do without you**.**

7 REFERENCES

1. F. Rohwer, Global phage diversity. *Cell*. **113**, 141 (2003).
2. A. Nasir, G. Caetano-Anollés, A phylogenomic data-driven exploration of viral origins and evolution. *Sci. Adv.* **1**, e1500527 (2015).
3. J. T. Marques, R. W. Carthew, A call to arms: coevolution of animal viruses and host innate immune responses. *Trends Genet.* **23**, 359–364 (2007).
4. F. Broecker, K. Moelling, Evolution of immune systems from viruses and transposable elements. *Front. Microbiol.* **10**, 51 (2019).
5. C. López-Otín, J. S. Bond, Proteases: multifunctional enzymes in life and disease. *J. Biol. Chem.* **283**, 30433–30437 (2008).
6. M. S. Brown, J. Ye, R. B. Rawson, J. L. Goldstein, Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans. *Cell*. **100**, 391–398 (2000).
7. K. Ray, C. S. Hines, J. Coll-Rodriguez, D. W. Rodgers, Crystal structure of human thimet oligopeptidase provides insight into substrate recognition, regulation, and localization. *J. Biol. Chem.* **279**, 20480–20489 (2004).
8. I. Schechter, A. Berger, On the active site of proteases. 3. Mapping the active site of papain; specific peptide inhibitors of papain. *Biochem. Biophys. Res. Commun.* **32**, 898–902 (1968).
9. A. M. Jabaiah, J. A. Getz, W. A. Witkowski, J. A. Hardy, P. S. Daugherty, Identification of protease exosite-interacting peptides that enhance substrate cleavage kinetics. *Biol. Chem.* **393**, 933–941 (2012).
10. N. D. Rawlings, in *Proteases: structure and function*, K. Brix, W. Stöcker, Eds. (Springer Vienna, Vienna, 2013), pp. 1–36.
11. N. D. Rawlings, A. J. Barrett, Evolutionary families of peptidases. *Biochem. J.* **290 (Pt 1)**, 205–218 (1993).
12. N. D. Rawlings, A. J. Barrett, A. Bateman, MEROPS: the peptidase database. *Nucleic Acids Res.* **38**, D227-33 (2010).
13. G. K. T. Nguyen *et al.*, Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat. Chem. Biol.* **10**, 732–738 (2014).
14. R. Yang *et al.*, Engineering a catalytically efficient recombinant protein ligase. *J. Am. Chem. Soc.* **139**, 5351–5358 (2017).
15. T. K. Chang, D. Y. Jackson, J. P. Burnier, J. A. Wells, Subtiligase: a tool for semisynthesis of proteins. *Proc. Natl. Acad. Sci. USA*. **91**, 12544–12548 (1994).
16. A. Shen, Allosteric regulation of protease activity by small molecules. *Mol. Biosyst.* **6**, 1431–1443 (2010).
17. S. Elmore, Apoptosis: a review of programmed cell death. *Toxicol. Pathol.* **35**, 495–516 (2007).
18. E. W. Davie, K. Fujikawa, W. Kisiel, The coagulation cascade: initiation, maintenance, and regulation. *Biochemistry*. **30**, 10363–10370 (1991).
19. R. B. Sim, S. A. Tsiftsoglou, Proteases of the complement system. *Biochem. Soc. Trans.* **32**, 21–27 (2004).
20. N. Fortelny *et al.*, Network analyses reveal pervasive functional regulation between proteases in the human protease web. *PLoS Biol.* **12**, e1001869 (2014).

21. D. W. Nicholson, Caspase structure, proteolytic substrates, and function during apoptotic cell death. *Cell Death Differ.* **6**, 1028–1042 (1999).
22. J. A. Trapani, V. R. Sutton, Granzyme B: pro-apoptotic, antiviral and antitumor functions. *Curr. Opin. Immunol.* **15**, 533–543 (2003).
23. B. Turk *et al.*, Apoptotic pathways: involvement of lysosomal proteases. *Biol. Chem.* **383**, 1035–1044 (2002).
24. S. Orrenius, B. Zhivotovsky, P. Nicotera, Regulation of cell death: the calcium-apoptosis link. *Nat. Rev. Mol. Cell Biol.* **4**, 552–565 (2003).
25. J. C. Timmer, G. S. Salvesen, Caspase substrates. *Cell Death Differ.* **14**, 66–72 (2007).
26. S. Fulda, K. M. Debatin, Extrinsic versus intrinsic apoptosis pathways in anticancer chemotherapy. *Oncogene*. **25**, 4798–4811 (2006).
27. S. W. G. Tait, D. R. Green, Mitochondria and cell death: outer membrane permeabilization and beyond. *Nat. Rev. Mol. Cell Biol.* **11**, 621–632 (2010).
28. B. Farhood, M. Najafi, K. Mortezaee, CD8+ cytotoxic T lymphocytes in cancer immunotherapy: A review. *J. Cell Physiol.* **234**, 8509–8521 (2019).
29. A. M. Abel, C. Yang, M. S. Thakar, S. Malarkannan, Natural killer cells: development, maturation, and clinical utilization. *Front. Immunol.* **9**, 1869 (2018).
30. L. Le Bourhis *et al.*, Antimicrobial activity of mucosal-associated invariant T cells. *Nat. Immunol.* **11**, 701–708 (2010).
31. E. Treiner *et al.*, Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1. *Nature*. **422**, 164–169 (2003).
32. D. Chowdhury, J. Lieberman, Death by a thousand cuts: granzyme pathways of programmed cell death. *Annu. Rev. Immunol.* **26**, 389–420 (2008).
33. S. Halle, O. Halle, R. Förster, Mechanisms and Dynamics of T Cell-Mediated Cytotoxicity In Vivo. *Trends Immunol.* **38**, 432–443 (2017).
34. W. J. Grossman *et al.*, The orphan granzymes of humans and mice. *Curr. Opin. Immunol.* **15**, 544–552 (2003).
35. M. Bots, J. P. Medema, Granzymes at a glance. *J. Cell Sci.* **119**, 5011–5014 (2006).
36. D. Martinvalet, D. M. Dykxhoorn, R. Ferrini, J. Lieberman, Granzyme A cleaves a mitochondrial complex I protein to initiate caspase-independent cell death. *Cell*. **133**, 681–692 (2008).
37. J. Lieberman, Granzyme A activates another way to die. *Immunol. Rev.* **235**, 93–104 (2010).
38. A. J. Darmon, D. W. Nicholson, R. C. Bleackley, Activation of the apoptotic protease CPP32 by cytotoxic T-cell-derived granzyme B. *Nature*. **377**, 446–448 (1995).
39. V. R. Sutton *et al.*, Initiation of apoptosis by granzyme B requires direct cleavage of bid, but not direct granzyme B-mediated caspase activation. *J. Exp. Med.* **192**, 1403–1414 (2000).
40. J. A. Heibein *et al.*, Granzyme B-mediated cytochrome c release is regulated by the Bcl-2 family members bid and Bax. *J. Exp. Med.* **192**, 1391–1402 (2000).
41. S. Aits, M. Jäättelä, Lysosomal cell death at a glance. *J. Cell Sci.* **126**, 1905–1912 (2013).
42. B. Turk, D. Turk, V. Turk, Lysosomal cysteine proteases: more than scavengers. *Biochim. Biophys. Acta*. **1477**, 98–111 (2000).

43. H. Kirschke, B. Wiederanders, D. Brömme, A. Rinne, Cathepsin S from bovine spleen. Purification, distribution, intracellular localization and action on proteins. *Biochem. J.* **264**, 467–473 (1989).
44. V. Stoka *et al.*, Lysosomal protease pathways to apoptosis. Cleavage of bid, not pro-caspases, is the most likely route. *J. Biol. Chem.* **276**, 3149–3157 (2001).
45. B. T. Chua, K. Guo, P. Li, Direct cleavage by the calcium-activated protease calpain can lead to inactivation of caspases. *J. Biol. Chem.* **275**, 5131–5135 (2000).
46. X. S. Puente, L. M. Sánchez, C. M. Overall, C. López-Otín, Human and mouse proteases: a comparative genomic approach. *Nat. Rev. Genet.* **4**, 544–558 (2003).
47. S. Chakraborti, N. S. Dhalla, Eds., *Proteases in health and disease* (Springer New York, New York, NY, 2013).
48. A. Luttun, M. Dewerchin, D. Collen, P. Carmeliet, The role of proteinases in angiogenesis, heart development, restenosis, atherosclerosis, myocardial ischemia, and stroke: insights from genetic studies. *Curr. Atheroscler. Rep.* **2**, 407–416 (2000).
49. J. Hardy, D. J. Selkoe, The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science*. **297**, 353–356 (2002).
50. R. Vassar *et al.*, Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science*. **286**, 735–741 (1999).
51. F. S. Esch *et al.*, Cleavage of amyloid beta peptide during constitutive processing of its precursor. *Science*. **248**, 1122–1124 (1990).
52. M. F. Egan *et al.*, Randomized Trial of Verubecestat for Mild-to-Moderate Alzheimer's Disease. *N. Engl. J. Med.* **378**, 1691–1703 (2018).
53. G. Guido, M. Scaglione, L. Fabbri, M. J. Ceglia, The "osteoporosis disease". *Clin. Cases Miner. Bone Metab.* **6**, 114–116 (2009).
54. D. Brömme, F. Lecaille, Cathepsin K inhibitors for osteoporosis and potential off-target effects. *Expert Opin. Investig. Drugs*. **18**, 585–600 (2009).
55. S. Boonen, E. Rosenberg, F. Claessens, D. Vanderschueren, S. Papapoulos, Inhibition of cathepsin K for treatment of osteoporosis. *Curr. Osteoporos. Rep.* **10**, 73–79 (2012).
56. K. Brixen *et al.*, Bone density, turnover, and estimated strength in postmenopausal women treated with odanacatib: a randomized trial. *J. Clin. Endocrinol. Metab.* **98**, 571–580 (2013).
57. M. T. Drake, B. L. Clarke, M. J. Oursler, S. Khosla, Cathepsin K inhibitors for osteoporosis: biology, potential clinical utility, and lessons learned. *Endocr. Rev.* **38**, 325–350 (2017).
58. J. E. Koblinski, M. Ahram, B. F. Sloane, Unraveling the role of proteases in cancer. *Clin. Chim. Acta*. **291**, 113–135 (2000).
59. L. Sevenich, J. A. Joyce, Pericellular proteolysis in cancer. *Genes Dev.* **28**, 2331–2347 (2014).
60. D. J. Matthews, J. A. Wells, Substrate phage: selection of protease substrates by monovalent phage display. *Science*. **260**, 1113–1117 (1993).
61. B. J. Backes, J. L. Harris, F. Leonetti, C. S. Craik, J. A. Ellman, Synthesis of positional-scanning libraries of fluorogenic peptide substrates to define the extended substrate specificity of plasmin and thrombin. *Nat. Biotechnol.* **18**, 187–193 (2000).
62. W. Shao, G. Yeretssian, K. Doiron, S. N. Hussain, M. Saleh, The caspase-1 digestome identifies the glycolysis pathway as a target during infection and septic

shock. *J. Biol. Chem.* **282**, 36321–36329 (2007).
63. N. J. Agard, J. A. Wells, Methods for the proteomic identification of protease substrates. *Curr. Opin. Chem. Biol.* **13**, 503–509 (2009).
64. P. Van Damme *et al.*, Caspase-specific and nonspecific in vivo protein processing during Fas-induced apoptosis. *Nat. Methods*. **2**, 771–777 (2005).
65. A. Doucet, O. Kleifeld, J. N. Kizhakkedathu, C. M. Overall, Identification of proteolytic products and natural protein N-termini by Terminal Amine Isotopic Labeling of Substrates (TAILS). *Methods Mol. Biol.* **753**, 273–287 (2011).
66. O. Kleifeld *et al.*, Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. *Nat. Biotechnol.* **28**, 281–288 (2010).
67. A. P. Wiita, J. E. Seaman, J. A. Wells, Global analysis of cellular proteolysis by selective enzymatic labeling of protein N-termini. *Meth. Enzymol.* **544**, 327–358 (2014).
68. K. Shimbo *et al.*, Quantitative profiling of caspase-cleaved substrates reveals different drug-induced and cell-type patterns in apoptosis. *Proc. Natl. Acad. Sci. USA*. **109**, 12432–12437 (2012).
69. L. Abrahmsén *et al.*, Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution. *Biochemistry*. **30**, 4151–4159 (1991).
70. S. Mahrus *et al.*, Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. *Cell*. **134**, 866–876 (2008).
71. A. M. Weeks, J. A. Wells, Engineering peptide ligase specificity by proteomic identification of ligation sites. *Nat. Chem. Biol.* **14**, 50–57 (2018).
72. E. D. Crawford *et al.*, The DegraBase: a database of proteolysis in healthy and apoptotic human cells. *Mol. Cell Proteomics*. **12**, 813–824 (2013).
73. N. Fortelny, S. Yang, P. Pavlidis, P. F. Lange, C. M. Overall, Proteome TopFIND 3.0 with TopFINDer and PathFINDer: database and analysis tools for the association of protein termini to pre- and post-translational events. *Nucleic Acids Res.* **43**, D290-7 (2015).
74. M. G. Rossmann, Y. He, R. J. Kuhn, Picornavirus-receptor interactions. *Trends Microbiol.* **10**, 324–331 (2002).
75. S. Ohka, A. Nomoto, Recent insights into poliovirus pathogenesis. *Trends Microbiol.* **9**, 501–506 (2001).
76. A. Iwasaki *et al.*, Immunofluorescence analysis of poliovirus receptor expression in Peyer's patches of humans, primates, and CD155 transgenic mice: implications for poliovirus infection. *J. Infect. Dis.* **186**, 585–592 (2002).
77. G. J. Belsham, N. Sonenberg, RNA-protein interactions in regulation of picornavirus RNA translation. *Microbiol Rev*. **60**, 499–511 (1996).
78. S. Guest, E. Pilipenko, K. Sharma, K. Chumakov, R. P. Roos, Molecular mechanisms of attenuation of the Sabin strain of poliovirus type 3. *J. Virol.* **78**, 11097–11107 (2004).
79. K. Lind, M. H. Hühn, M. Flodström-Tullberg, Immunology in the clinic review series; focus on type 1 diabetes and viruses: the innate immune response to enteroviruses and its possible role in regulating type 1 diabetes. *Clin. Exp. Immunol.* **168**, 30–38 (2012).
80. J. L. Whitton, C. T. Cornell, R. Feuer, Host and virus determinants of picornavirus

pathogenesis and tropism. *Nat. Rev. Microbiol.* **3**, 765–776 (2005).
81. S. E. Bakker *et al.*, Limits of structural plasticity in a picornavirus capsid revealed by a massively expanded equine rhinitis A virus particle. *J. Virol.* **88**, 6093–6099 (2014).
82. M. S. Smyth, J. H. Martin, Picornavirus uncoating. *Mol. Pathol.* **55**, 214–219 (2002).
83. A. Panjwani *et al.*, Capsid protein VP4 of human rhinovirus induces membrane permeability by the formation of a size-selective multimeric pore. *PLoS Pathog.* **10**, e1004294 (2014).
84. J. Ren *et al.*, Picornavirus uncoating intermediate captured in atomic detail. *Nat. Commun.* **4**, 1929 (2013).
85. C. E. Fricks, J. M. Hogle, Cell-induced conformational change in poliovirus: externalization of the amino terminus of VP1 is responsible for liposome binding. *J. Virol.* **64**, 1934–1945 (1990).
86. Z. Liu *et al.*, Structural and functional analysis of the 5' untranslated region of coxsackievirus B3 RNA: In vivo translational and infectivity studies of full-length mutants. *Virology*. **265**, 206–217 (1999).
87. V. Lulla *et al.*, An upstream protein-coding region in enteroviruses modulates virus infection in gut epithelial cells. *Nat. Microbiol.* **4**, 280–292 (2019).
88. A. C. Palmenberg, Proteolytic processing of picornaviral polyprotein. *Annu. Rev. Microbiol.* **44**, 603–623 (1990).
89. M. D. Ryan, J. Drew, Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. *EMBO J.* **13**, 928–933 (1994).
90. M. D. Ryan, A. M. King, G. P. Thomas, Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence. *J. Gen. Virol.* **72 (Pt 11)**, 2727–2732 (1991).
91. G. A. Luke *et al.*, Occurrence, function and evolutionary origins of "2A-like" sequences in virus genomes. *J. Gen. Virol.* **89**, 1036–1042 (2008).
92. A. J. Chase, B. L. Semler, Viral subversion of host functions for picornavirus translation and RNA replication. *Future Virol.* **7**, 179–191 (2012).
93. O. H. Laitinen *et al.*, Enteroviral proteases: structure, host interactions and pathogenicity. *Rev Med Virol*. **26**, 251–267 (2016).
94. P. Jiang, Y. Liu, H.-C. Ma, A. V. Paul, E. Wimmer, Picornavirus morphogenesis. *Microbiol. Mol. Biol. Rev.* **78**, 418–437 (2014).
95. N. Altan-Bonnet, Y.-H. Chen, Intercellular Transmission of Viral Populations with Vesicles. *J. Virol.* **89**, 12242–12244 (2015).
96. Y.-H. Chen *et al.*, Phosphatidylserine vesicles enable efficient en bloc transmission of enteroviruses. *Cell*. **160**, 619–630 (2015).
97. W. Chehadeh *et al.*, Persistent infection of human pancreatic islets by coxsackievirus B is associated with alpha interferon synthesis in beta cells. *J. Virol.* **74**, 10153–10164 (2000).
98. T. Jartti, P. Lehtinen, T. Vuorinen, M. Koskenvuo, O. Ruuskanen, Persistence of rhinovirus and enterovirus RNA after acute respiratory illness in children. *J. Med. Virol.* **72**, 695–699 (2004).
99. K. Klingel *et al.*, Ongoing enterovirus-induced myocarditis is associated with persistent heart muscle infection: quantitative analysis of virus replication, tissue

damage, and inflammation. *Proc. Natl. Acad. Sci. USA*. **89**, 314–318 (1992).
100. R. Basavappa *et al.*, Role and mechanism of the maturation cleavage of VP0 in poliovirus assembly: structure of the empty capsid assembly intermediate at 2.9 A resolution. *Protein Sci.* **3**, 1651–1669 (1994).
101. W. Lee *et al.*, Solution structure of the 2A protease from a common cold agent, human rhinovirus C2, strain W12. *PLoS One*. **9**, e97198 (2014).
102. J. Tan *et al.*, 3C protease of enterovirus 68: structure-based design of Michael acceptor inhibitors and their broad-spectrum antiviral effects against picornaviruses. *J. Virol.* **87**, 4339–4351 (2013).
103. H. M. Berman *et al.*, The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
104. N. Blom, J. Hansen, D. Blaas, S. Brunak, Cleavage site analysis in picornaviral polyproteins: discovering cellular targets by neural networks. *Protein Sci.* **5**, 2203–2216 (1996).
105. S. Kumar, B. J. van Raam, G. S. Salvesen, P. Cieplak, Caspase cleavage sites in the human proteome: CaspDB, a database of predicted substrates. *PLoS One*. **9**, e110539 (2014).
106. J. Song *et al.*, iProt-Sub: a comprehensive package for accurately mapping and predicting protease-specific substrates and cleavage sites. *Brief. Bioinformatics* (2018), doi:10.1093/bib/bby028.
107. E. Gasteiger *et al.*, in *The proteomics protocols handbook*, J. M. Walker, Ed. (Humana Press, Totowa, NJ, 2005), pp. 571–607.
108. M. F. Jacobson, D. Baltimore, Polypeptide cleavages in the formation of poliovirus proteins. *Proc. Natl. Acad. Sci. USA*. **61**, 77–84 (1968).
109. M. F. Jacobson, J. Asso, D. Baltimore, Further evidence on the formation of poliovirus proteins. *J. Mol. Biol.* **49**, 657–669 (1970).
110. D. S. Shih *et al.*, Cell-free synthesis and processing of the proteins of poliovirus. *Proc. Natl. Acad. Sci. USA*. **75**, 5807–5811 (1978).
111. M. F. Jacobson, D. Baltimore, Morphogenesis of poliovirus. I. Association of the viral RNA with coat protein. *J. Mol. Biol.* **33**, 369–378 (1968).
112. G. R. Larsen, C. W. Anderson, A. J. Dorner, B. L. Semler, E. Wimmer, Cleavage sites within the poliovirus capsid protein precursors. *J. Virol.* **41**, 340–344 (1982).
113. M. J. Grubman, D. O. Morgan, J. Kendall, B. Baxt, Capsid intermediates assembled in a foot-and-mouth disease virus genome RNA-programmed cell-free translation system and in infected cells. *J. Virol.* **56**, 120–126 (1985).
114. A. C. Palmenberg, In vitro synthesis and assembly of picornaviral capsid intermediate structures. *J. Virol.* **44**, 900–906 (1982).
115. D. Lugo, P. Krogstad, Enteroviruses in the early 21st century: new manifestations and challenges. *Curr. Opin. Pediatr.* **28**, 107–113 (2016).
116. C. Tapparel, F. Siegrist, T. J. Petty, L. Kaiser, Picornavirus and enterovirus diversity with associated human diseases. *Infect. Genet. Evol.* **14**, 282–293 (2013).
117. D. R. Gamble, K. W. Taylor, Seasonal incidence of diabetes mellitus. *Br. Med. J.* **3**, 631–633 (1969).
118. C. M. Filippi, M. G. von Herrath, Viral trigger for type 1 diabetes: pros and cons. *Diabetes*. **57**, 2863–2871 (2008).
119. J. Green, D. Casabonne, R. Newton, Coxsackie B virus serology and Type 1 diabetes mellitus: a systematic review of published case-control studies. *Diabet.*

Med. **21**, 507–514 (2004).
120. W.-C. G. Yeung, W. D. Rawlinson, M. E. Craig, Enterovirus infection and type 1 diabetes mellitus: systematic review and meta-analysis of observational molecular studies. *BMJ*. **342**, d35 (2011).
121. L. Krogvold *et al.*, Pancreatic biopsy by minimal tail resection in live adult patients at the onset of type 1 diabetes: experiences from the DiViD study. *Diabetologia*. **57**, 841–843 (2014).
122. L. Krogvold *et al.*, Detection of a low-grade enteroviral infection in the islets of langerhans of living patients newly diagnosed with type 1 diabetes. *Diabetes*. **64**, 1682–1687 (2015).
123. O. Skog, K. Klingel, M. Roivainen, O. Korsgren, Large enteroviral vaccination studies to prevent type 1 diabetes should be well founded and rely on scientific evidence. *Diabetologia* (2019), doi:10.1007/s00125-019-4841-1.
124. S. F. Hansson, S. Korsgren, F. Pontén, O. Korsgren, Enteroviruses and the pathogenesis of type 1 diabetes revisited: cross-reactivity of enterovirus capsid protein (VP1) antibodies with human mitochondrial proteins. *J. Pathol.* **229**, 719–728 (2013).
125. N. Ettischer-Schmid *et al.*, A new monoclonal antibody (Cox mAB 31A2) detects VP1 protein of coxsackievirus B3 with high sensitivity and specificity. *Virchows Arch*. **469**, 553–562 (2016).
126. S. J. Richardson *et al.*, Evaluation of the fidelity of immunolabelling obtained with clone 5D8/1, a monoclonal antibody directed against the enteroviral capsid protein, VP1, in human pancreas. *Diabetologia*. **57**, 392–401 (2014).
127. S. J. Richardson, A. Willcox, A. J. Bone, A. K. Foulis, N. G. Morgan, The prevalence of enteroviral capsid protein vp1 immunostaining in pancreatic islets in human type 1 diabetes. *Diabetologia*. **52**, 1143–1151 (2009).
128. N. van der Werf, F. G. M. Kroese, J. Rozing, J.-L. Hillebrands, Viral infections as potential triggers of type 1 diabetes. *Diabetes Metab Res Rev*. **23**, 169–183 (2007).
129. A. L. Notkins, A. Lernmark, Autoimmune type 1 diabetes: resolved and unresolved issues. *J. Clin. Invest.* **108**, 1247–1252 (2001).
130. E. Kawasaki, Type 1 diabetes and autoimmunity. *Clin. Pediatr. Endocrinol.* **23**, 99–105 (2014).
131. E. K. Alidjinou *et al.*, Persistence of Coxsackievirus B4 in pancreatic ductal-like cells results in cellular and viral changes. *Virulence*. **8**, 1229–1244 (2017).
132. M. Oikarinen *et al.*, Type 1 diabetes is associated with enterovirus infection in gut mucosa. *Diabetes*. **61**, 687–691 (2012).
133. N. M. Chapman, K. S. Kim, Persistent coxsackievirus infection: enterovirus persistence in chronic myocarditis and dilated cardiomyopathy. *Curr. Top. Microbiol. Immunol.* **323**, 275–292 (2008).
134. R. S. Fujinami, M. G. von Herrath, U. Christen, J. L. Whitton, Molecular mimicry, bystander activation, or viral persistence: infections and autoimmune disease. *Clin. Microbiol. Rev.* **19**, 80–94 (2006).
135. H. Kato *et al.*, Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature*. **441**, 101–105 (2006).
136. L. G. Guidotti, F. V. Chisari, Noncytolytic control of viral infections by the innate and adaptive immune response. *Annu. Rev. Immunol.* **19**, 65–91 (2001).

137. A. J. Lee, A. A. Ashkar, The dual nature of type I and type II interferons. *Front. Immunol.* **9**, 2061 (2018).
138. P. M. Barral, D. Sarkar, P. B. Fisher, V. R. Racaniello, RIG-I is cleaved during picornavirus infection. *Virology*. **391**, 171–176 (2009).
139. Q. Feng *et al.*, Enterovirus 2Apro targets MDA5 and MAVS in infected cells. *J. Virol.* **88**, 3369–3378 (2014).
140. A. Mukherjee *et al.*, The coxsackievirus B 3C protease cleaves MAVS and TRIF to attenuate host type I interferon and apoptotic signaling. *PLoS Pathog.* **7**, e1001311 (2011).
141. K. Lind *et al.*, Coxsackievirus counters the host innate immune response by blocking type III interferon expression. *J. Gen. Virol.* **97**, 1368–1380 (2016).
142. X. Lei *et al.*, Cleavage of interferon regulatory factor 7 by enterovirus 71 3C suppresses cellular responses. *J. Virol.* **87**, 1690–1698 (2013).
143. M. Sato *et al.*, Distinct and essential roles of transcription factors IRF-3 and IRF-7 in response to viruses for IFN-alpha/beta gene induction. *Immunity*. **13**, 539–548 (2000).
144. H.-C. Hung *et al.*, Synergistic inhibition of enterovirus 71 replication by interferon and rupintrivir. *J. Infect. Dis.* **203**, 1784–1790 (2011).
145. M. K. Weidman *et al.*, The interaction of cytoplasmic RNA viruses with the nucleus. *Virus Res.* **95**, 75–85 (2003).
146. P. Yalamanchili, K. Harris, E. Wimmer, A. Dasgupta, Inhibition of basal transcription by poliovirus: a virus- encoded protease (3Cpro) inhibits formation of TBP-TATA box complex in vitro. *J. Virol.* **70**, 2922–2929 (1996).
147. M. E. Clark, T. Hämmerle, E. Wimmer, A. Dasgupta, Poliovirus proteinase 3C converts an active form of transcription factor IIIC to an inactive form: a mechanism for inhibition of host cell polymerase III transcription by poliovirus. *EMBO J.* **10**, 2941–2947 (1991).
148. P. S. Ray, S. Das, La autoantigen is required for the internal ribosome entry site-mediated translation of Coxsackievirus B3 RNA. *Nucleic Acids Res.* **30**, 4500–4508 (2002).
149. Y. V. Svitkin *et al.*, Internal translation initiation on poliovirus RNA: further characterization of La function in poliovirus translation in vitro. *J. Virol.* **68**, 1544–1550 (1994).
150. K. Meerovitch *et al.*, La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. *J. Virol.* **67**, 3798–3807 (1993).
151. A. W. Craig, Y. V. Svitkin, H. S. Lee, G. J. Belsham, N. Sonenberg, The La autoantigen contains a dimerization domain that is essential for enhancing translation. *Mol. Cell. Biol.* **17**, 163–169 (1997).
152. B. L. Walter, T. B. Parsley, E. Ehrenfeld, B. L. Semler, Distinct poly(rC) binding protein KH domain determinants for poliovirus translation initiation and viral RNA replication. *J. Virol.* **76**, 12008–12022 (2002).
153. L. B. Blyn, J. S. Towner, B. L. Semler, E. Ehrenfeld, Requirement of poly(rC) binding protein 2 for translation of poliovirus RNA. *J. Virol.* **71**, 6243–6246 (1997).
154. L. B. Blyn *et al.*, Poly(rC) binding protein 2 binds to stem-loop IV of the poliovirus RNA 5' noncoding region: identification by automated liquid chromatography-tandem mass spectrometry. *Proc. Natl. Acad. Sci. USA*. **93**, 11115–11120 (1996).

155. C. U. Hellen *et al.*, A cytoplasmic 57-kDa protein that is required for translation of picornavirus RNA by internal ribosomal entry is identical to the nuclear pyrimidine tract-binding protein. *Proc. Natl. Acad. Sci. USA*. **90**, 7642–7646 (1993).
156. R. Gosert *et al.*, Transient expression of cellular polypyrimidine-tract binding protein stimulates cap-independent translation directed by both picornaviral and flaviviral internal ribosome entry sites In vivo. *Mol. Cell. Biol.* **20**, 1583–1595 (2000).
157. K. D. Fitzgerald, B. L. Semler, Re-localization of cellular protein SRp20 during poliovirus infection: bridging a viral IRES to the host cell translation apparatus. *PLoS Pathog.* **7**, e1002127 (2011).
158. K. Watters, A. C. Palmenberg, Differential processing of nuclear pore complex proteins by rhinovirus 2A proteases from different species and serotypes. *J. Virol.* **85**, 10874–10883 (2011).
159. N. Park, T. Skern, K. E. Gustin, Specific cleavage of the nuclear pore complex protein Nup62 by a viral protease. *J. Biol. Chem.* **285**, 28796–28805 (2010).
160. A. Castelló, J. M. Izquierdo, E. Welnowska, L. Carrasco, RNA nuclear export is blocked by poliovirus 2A protease and is concomitant with nucleoporin cleavage. *J. Cell Sci.* **122**, 3799–3809 (2009).
161. R. E. Lloyd, Enterovirus control of translation and RNA granule stress responses. *Viruses*. **8**, 93 (2016).
162. R. J. Jackson, C. U. T. Hellen, T. V. Pestova, The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* **11**, 113–127 (2010).
163. M. Joachims, P. C. Van Breugel, R. E. Lloyd, Cleavage of poly(A)-binding protein by enterovirus proteases concurrent with inhibition of translation in vitro. *J. Virol.* **73**, 718–727 (1999).
164. B.-C. Ho *et al.*, Enterovirus-induced miR-141 contributes to shutoff of host protein translation by targeting the translation initiation factor eIF4E. *Cell Host Microbe*. **9**, 58–69 (2011).
165. Y. Fu *et al.*, Exosome-mediated miR-146a transfer suppresses type I interferon response and facilitates EV71 infection. *PLoS Pathog.* **13**, e1006611 (2017).
166. S. M. Robinson *et al.*, Coxsackievirus B exits the host cell in shed microvesicles displaying autophagosomal markers. *PLoS Pathog.* **10**, e1004045 (2014).
167. S. Makino *et al.*, Breeding of a non-obese, diabetic strain of mice. *Jikken Dobutsu.* **29**, 1–13 (1980).
168. R. Feuer, I. Mena, R. Pagarigan, M. K. Slifka, J. L. Whitton, Cell cycle status affects coxsackievirus replication, persistence, and reactivation in vitro. *J. Virol.* **76**, 4430–4440 (2002).
169. K. P. Dalton, J. K. Rose, Vesicular stomatitis virus glycoprotein containing the entire green fluorescent protein on its cytoplasmic domain is incorporated efficiently into virus particles. *Virology*. **279**, 414–421 (2001).
170. J. E. Laiho *et al.*, Application of bioinformatics in probe design enables detection of enteroviruses on different taxonomic levels by advanced in situ hybridization technology. *J. Clin. Virol.* **69**, 165–171 (2015).
171. N. Ank *et al.*, Lambda interferon (IFN-lambda), a type III IFN, is induced by viruses and IFNs and displays potent antiviral activity against select virus infections in vivo.

J. Virol. **80**, 4501–4509 (2006).
172. H. E. Hohmeier *et al.*, Isolation of INS-1-derived cell lines with robust ATP-sensitive K+ channel-dependent and -independent glucose-stimulated insulin secretion. *Diabetes*. **49**, 424–430 (2000).
173. P. Ravassard *et al.*, A genetically engineered human pancreatic β cell line exhibiting glucose-inducible insulin secretion. *J. Clin. Invest.* **121**, 3589–3597 (2011).
174. K. Lind *et al.*, Induction of an antiviral state and attenuated coxsackievirus replication in type III interferon-treated primary human pancreatic islets. *J. Virol.* **87**, 7646–7654 (2013).
175. W. N. Burnette, "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* **112**, 195–203 (1981).
176. J. M. Silva, M. McMahon, The fastest Western in town: a contemporary twist on the classic Western blot analysis. *J. Vis. Exp.*, e51149 (2014).
177. S. L. Eaton *et al.*, A guide to modern quantitative fluorescent western blotting with troubleshooting strategies. *J. Vis. Exp.*, e52099 (2014).
178. J. Vikman *et al.*, Truncation of SNAP-25 reduces the stimulatory action of cAMP on rapid exocytosis in insulin-secreting cells. *Am. J. Physiol. Endocrinol. Metab.* **297**, E452-61 (2009).
179. B. G. Kornreich, The patch clamp technique: principles and technical considerations. *J. Vet. Cardiol.* **9**, 25–37 (2007).
180. K. N. Fish, *Curr. Protoc. Cytom.*, in press, doi:10.1002/0471142956.cy1218s50.
181. S. Durand, A. Cimarelli, The inside out of lentiviral vectors. *Viruses*. **3**, 132–159 (2011).
182. T. Sakuma, M. A. Barry, Y. Ikeda, Lentiviral vectors: basic to translational. *Biochem. J.* **443**, 603–618 (2012).
183. W. McKinney, Data Structures for Statistical Computing in Python (2010).
184. J. D. Hunter, Matplotlib: A 2D Graphics Environment. *Comput Sci Eng*. **9**, 90–95 (2007).
185. S. van der Walt, S. C. Colbert, G. Varoquaux, The NumPy Array: A Structure for Efficient Numerical Computation. *Comput Sci Eng*. **13**, 22–30 (2011).
186. A. Hagberg, P. Swart, D. S. Chult, Exploring Network Structure, Dynamics, and Function Using NetworkX (2008).
187. G. Yu, L.-G. Wang, G.-R. Yan, Q.-Y. He, DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics*. **31**, 608–609 (2015).
188. G. Yu, Q.-Y. He, ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol. Biosyst.* **12**, 477–479 (2016).
189. J. Piñero *et al.*, DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res.* **45**, D833–D839 (2017).
190. H. Hermjakob *et al.*, IntAct: an open source molecular interaction database. *Nucleic Acids Res.* **32**, D452-5 (2004).
191. D. Croft *et al.*, Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* **39**, D691-7 (2011).

192. O. Wagih, ggseqlogo: a versatile R package for drawing sequence logos. *Bioinformatics*. **33**, 3645–3647 (2017).
193. E. Jones, T. Oliphant, P. Peterson, SciPy: Open Source Scientific Tools for Python (2001).
194. P. Sheppard *et al.*, IL-28, IL-29 and their class II cytokine receptor IL-28R. *Nat. Immunol.* **4**, 63–68 (2003).
195. S. V. Kotenko *et al.*, IFN-lambdas mediate antiviral protection through a distinct class II cytokine receptor complex. *Nat. Immunol.* **4**, 69–77 (2003).
196. H. M. Lazear, T. J. Nice, M. S. Diamond, Interferon-λ: Immune Functions at Barrier Surfaces and Beyond. *Immunity*. **43**, 15–28 (2015).
197. B. Wang *et al.*, Enterovirus 71 protease 2Apro targets MAVS to inhibit anti-viral type I interferon responses. *PLoS Pathog.* **9**, e1003231 (2013).
198. G. Maccari, A. Genoni, S. Sansonno, A. Toniolo, Properties of Two Enterovirus Antibodies that are Utilized in Diabetes Research. *Sci. Rep.* **6**, 24757 (2016).
199. B. T. Clause, The Wistar Rat as a right choice: establishing mammalian standards and the ideal of a standardized mammal. *J Hist Biol*. **26**, 329–349 (1993).
200. S. P. Amineva, A. G. Aminev, A. C. Palmenberg, J. E. Gern, Rhinovirus 3C protease precursors 3CD and 3CD' localize to the nuclei of infected cells. *J. Gen. Virol.* **85**, 2969–2979 (2004).
201. R. Ghildyal *et al.*, Rhinovirus 3C protease can localize in the nucleus and alter active and passive nucleocytoplasmic transport. *J. Virol.* **83**, 7349–7352 (2009).
202. R. Sharma, S. Raychaudhuri, A. Dasgupta, Nuclear entry of poliovirus protease-polymerase precursor 3CD: implications for host cell transcription shut-off. *Virology*. **320**, 195–205 (2004).
203. E. Walker *et al.*, Rhinovirus 16 2A Protease Affects Nuclear Localization of 3CD during Infection. *J. Virol.* **90**, 11032–11042 (2016).
204. M. T. Sanchez-Aparicio, M. F. Rosas, F. Sobrino, Characterization of a nuclear localization signal in the foot-and-mouth disease virus polymerase. *Virology*. **444**, 203–210 (2013).
205. W. Tian, Z. Cui, Z. Zhang, H. Wei, X. Zhang, Poliovirus 2A(pro) induces the nucleic translocation of poliovirus 3CD and 3C' proteins. *Acta Biochim. Biophys. Sin. (Shanghai)*. **43**, 38–44 (2011).
206. G. Frisk, H. Diderholm, Tissue culture of isolated human pancreatic islets infected with different strains of coxsackievirus B4: assessment of virus replication and effects on islet morphology and insulin release. *Int. J. Exp. Diabetes Res.* **1**, 165–175 (2000).
207. M. Roivainen *et al.*, Mechanisms of coxsackievirus-induced damage to human pancreatic beta-cells. *J. Clin. Endocrinol. Metab.* **85**, 432–440 (2000).
208. D. A. Dodd, T. H. Giddings, K. Kirkegaard, Poliovirus 3A protein limits interleukin-6 (IL-6), IL-8, and beta interferon secretion during viral infection. *J. Virol.* **75**, 8158–8165 (2001).
209. S. Nair, K.-C. Leung, W. D. Rawlinson, Z. Naing, M. E. Craig, Enterovirus infection induces cytokine and chemokine expression in insulin-producing cells. *J. Med. Virol.* **82**, 1950–1957 (2010).
210. P. F. Wick, R. A. Senter, L. A. Parsels, M. D. Uhler, R. W. Holz, Transient transfection studies of secretion in bovine chromaffin cells and PC12 cells.

Generation of kainate-sensitive chromaffin cells. *J. Biol. Chem.* **268**, 10983–10989 (1993).
211. R. E. Lloyd, Translational control by viral proteinases. *Virus Res.* **119**, 76–88 (2006).
212. P. Rorsman, F. M. Ashcroft, Pancreatic β-Cell Electrical Activity and Insulin Secretion: Of Mice and Men. *Physiol. Rev.* **98**, 117–214 (2018).
213. G. Fung *et al.*, Production of a dominant-negative fragment due to G3BP1 cleavage contributes to the disruption of mitochondria-associated protective stress granules during CVB3 infection. *PLoS One*. **8**, e79546 (2013).
214. V. G. Wilson, Viral Interplay with the Host Sumoylation System. *Adv. Exp. Med. Biol.* **963**, 359–388 (2017).
215. S.-C. Chen *et al.*, Sumoylation-promoted enterovirus 71 3C degradation correlates with a reduction in viral replication and cell apoptosis. *J. Biol. Chem.* **286**, 31373–31384 (2011).
216. J. P. White, A. M. Cardenas, W. E. Marissen, R. E. Lloyd, Inhibition of cytoplasmic mRNA stress granule formation by a viral proteinase. *Cell Host Microbe*. **2**, 295–305 (2007).
217. W. Ullmer, B. L. Semler, Diverse strategies used by picornaviruses to escape host RNA decay pathways. *Viruses*. **8** (2016), doi:10.3390/v8120335.
218. H. J. Welters, R. N. Kulkarni, Wnt signaling: relevance to beta-cell biology and diabetes. *Trends Endocrinol. Metab.* **19**, 349–355 (2008).
219. S. F. A. Grant *et al.*, Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat. Genet.* **38**, 320–323 (2006).
220. S. Mayans *et al.*, TCF7L2 polymorphisms are associated with type 2 diabetes in northern Sweden. *Eur. J. Hum. Genet.* **15**, 342–346 (2007).
221. C. Cropano *et al.*, The rs7903146 Variant in the TCF7L2 Gene Increases the Risk of Prediabetes/Type 2 Diabetes in Obese Adolescents by Impairing β-Cell Function and Hepatic Insulin Sensitivity. *Diabetes Care*. **40**, 1082–1089 (2017).
222. G. da Silva Xavier *et al.*, TCF7L2 regulates late events in insulin secretion from pancreatic islet beta-cells. *Diabetes*. **58**, 894–905 (2009).
223. Y. Zhou *et al.*, TCF7L2 is a master regulator of insulin production and processing. *Hum. Mol. Genet.* **23**, 6419–6431 (2014).
224. S. Zarei *et al.*, A comprehensive review of amyotrophic lateral sclerosis. *Surg Neurol Int*. **6**, 171 (2015).
225. S.-M. Pulst, in *Genetic instabilities and neurological diseases* (Elsevier, 2006), pp. 351–361.
226. D. Neilson, in *GeneReviews(®)*, R. A. Pagon et al., Eds. (University of Washington, Seattle, Seattle (WA), 1993).
227. K. Messacar *et al.*, A cluster of acute flaccid paralysis and cranial nerve dysfunction temporally associated with an outbreak of enterovirus D68 in children in Colorado, USA. *Lancet*. **385**, 1662–1671 (2015).
228. K. Messacar *et al.*, Enterovirus D68 and acute flaccid myelitis-evaluating the evidence for causality. *Lancet Infect. Dis.* **18**, e239–e247 (2018).
229. I.-J. Chen, S.-C. Hu, K.-L. Hung, C.-W. Lo, Acute flaccid myelitis associated with enterovirus D68 infection: A case report. *Medicine*. **97**, e11831 (2018).
230. H. Jeong, S. P. Mason, A. L. Barabási, Z. N. Oltvai, Lethality and centrality in

protein networks. *Nature*. **411**, 41–42 (2001).
231. M. P. Joy, A. Brock, D. E. Ingber, S. Huang, High-betweenness proteins in the yeast protein interaction network. *J. Biomed. Biotechnol.* **2005**, 96–103 (2005).